

**Proteomic profiling of phosphoproteins and secreted proteins from mammalian cell lines in order to gain insights into factors affecting cellular growth and recombinant protein production.**

A thesis submitted for the degree of PhD  
by Martin Power, BSc.

April 2015

The experimental work in this thesis was performed  
under the supervision of

Dr. Paula Meleady

&

Prof. Martin Clynes

School of Biotechnology

Dublin City University

*I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.*

Signed: \_\_\_\_\_ ID No.: \_\_\_\_\_

Date: \_\_\_\_\_

*This thesis is dedicated to my mother, Doris*

*and*

*the memory of my father, Frank.*

## **Acknowledgements**

Firstly I would like to thank my supervisor Dr. Paula Meleady for taking me on and guiding me through the PhD process with endless positivity, encouragement and support.

I would also like to thank Prof. Martin Clynes for giving me the opportunity to carry out my work in the National Institute for Cellular Biotechnology. I am sincerely grateful for your unwavering support and guidance over the course of the long journey towards my PhD.

A special thank you to Michael Henry for all his help with the mass spec work (of which there was no shortage in my project!). Thank you for all the time you gave me and your endless patience.

Thank you to Dr. Annemarie Larkin for all her help with the hybridoma work and to Dr. Colin Clarke for his help with the bioinformatics. Also thank you to Carol, Yvonne and Mairead for all their help in the background to keep the show on the road.

A special thank you to John, Kate and Bernadette in NIBRT for all the support they showed me during the final stages of my PhD.

Thank you to Noelia, Shane, Mark, Paul, Alan, Gemma, Erica, Deidre and Damian and all my friends in the NICB for making it such a great place to come to work every day.

Finally to my mum, Doris, thank you for being so patient for so long. *Now*, I'm finally finished!



	<b>Table of Contents</b>	
	<b>Abstract</b>	<b>1</b>
<b>1.0</b>	<b>Introduction</b>	<b>3</b>
<b>1.1</b>	<b>Chinese Hamster Ovary cells</b>	<b>3</b>
<b>1.2</b>	<b>Temperature shift</b>	<b>7</b>
1.2.1	Mechanisms controlling Cell growth rates in temperature shifted CHO cells	7
1.2.2	Mechanisms controlling specific productivity in temperature shifted CHO cells	8
1.2.3	Proteomic profiling of temperature shifted CHO cells	12
<b>1.3</b>	<b>Proteomic profiling of desirable cell phenotypes</b>	<b>13</b>
<b>1.4</b>	<b>Host cell proteins (HCPs)</b>	<b>17</b>
1.4.1	Proteases as part of the HCP matrix	19
1.4.2	Growth Factors and Conditioned Media	21
<b>1.5</b>	<b>Introduction to Proteomics</b>	<b>24</b>
1.5.1	Proteomic sample preparation	24
1.5.2	Sample separation	25
1.5.2.1	Protein separation by 2D-PAGE	26
1.5.2.2	Protein separation by Liquid chromatography	27
1.5.3	Mass spectrometry for protein identification and characterisation	28
1.5.3.1	Sample preparation for Mass Spectrometry	28
1.5.3.2	Nanoflow Liquid Chromatography coupled to Mass Spectrometers	29
1.5.3.3	The basic components of a mass spectrometer	30
1.5.3.4	Determining the peptide sequence from mass spectrometry data	34
1.5.4	Mass spectrometry data output	36
1.5.5	Bioinformatic analysis of mass spectral data	36
1.5.5.1	CHO specific proteomic databases	38
1.5.6	Post Translational Modifications (PTM) Analysis using Mass Spectrometry	39
1.5.7	Identification of phosphopeptides by mass spectrometry	40
1.5.8	Phosphoproteomic databases	43
1.5.9	Quantitative label free LC/MS	45
<b>1.6</b>	<b>Analysis of protein phosphorylation</b>	<b>47</b>
1.6.1	Protein arrays to study phosphorylation	50
1.6.2	Staining of phosphoproteins for gel based approaches	51
1.6.3	Chromatography for the enrichment of phosphoproteins and phosphopeptides	52
1.6.3.1	Titanium dioxide phosphopeptide enrichment	53
1.6.3.2	IMAC for phosphopeptide enrichment	55
1.6.3.3	Strong cation exchange (SCX)	56
	<b>Aims</b>	<b>58</b>
<b>2.0</b>	<b>Materials &amp; Method</b>	<b>61</b>
<b>2.1</b>	<b>Preparation for Cell Culture</b>	<b>61</b>
2.1.1	Ultrapure water	61
2.1.2	Glassware	61
2.1.3	Sterilisation	61

2.1.4	Subculturing of cell lines suspension cells	61
2.1.5	Assessment of cell number and viability	63
2.1.6	Cryopreservation of cells	63
2.1.7	Thawing of cryopreserved cells	63
2.1.8	Sterility checks	64
2.1.9	Mycoplasma testing	64
<b>2.2</b>	<b>Protein Quantification</b>	<b>64</b>
2.2.1	Bradford Assay	64
2.2.2	ELISA Assay	65
2.2.3	SEAP Assay	65
<b>2.3</b>	<b>Proteomic Sample Collection and Preparation</b>	<b>66</b>
2.3.1	Cell Lysate Sample Preparation	66
2.3.2	Conditioned Media Collection	67
2.3.3	Conditioned Media Sample Preparation	67
2.3.4	Sample clean-up for Mass Spectral Analysis	67
2.3.5	In solution digestion for preparation of conditioned media samples	67
2.3.6	In solution digestion for preparation of whole cell lysate samples	68
2.3.7	Peptide sample clean-up using C18 spin column	68
2.3.8	Peptide sample clean-up using Graphite spin column	69
<b>2.4</b>	<b>Proteomic Enrichment Methods</b>	<b>69</b>
2.4.1	Phosphoprotein enrichment Pierce	69
2.4.2	Phosphoprotein Label Free	70
2.4.3	Gallium Phosphopeptide enrichment	70
2.4.4	TiO <sub>2</sub> Phosphopeptide enrichment	70
2.4.5	Iron Phosphopeptide enrichment	70
2.4.6	LC/MS analysis	71
2.4.7	Mass Spectrometry Protein Identifications	72
2.4.8	Label-free LC-MS data analysis	73
<b>2.5</b>	<b>Other Proteomic Techniques</b>	<b>74</b>
2.5.1	Gel electrophoresis	74
2.5.2	Western blotting	74
2.5.3	Enhanced chemiluminescence (ECL) detection	75
2.5.4	MAP Kinase array	76
2.5.5	Depletion of IgG from conditioned media collected from DP12 cell line	77
<b>3.0</b>	<b>Host Cell Protein Analysis</b>	<b>79</b>
<b>3.1</b>	<b>Characterisation of the host cell protein profile of a non-producing hybridoma cell line</b>	<b>80</b>
3.1.1	Growth profiling of non-producing hybridoma cell line	80
3.1.2	Quantitative label free LC/MS analysis of conditioned media from a non-producing hybridoma cell line	82
3.1.3	Data analysis of differentially regulated proteins identified by label free LC/MS profiling of conditioned media from a hybridoma cell line	82
3.1.3.1	Differential regulation of HCPs 24hrs vs 96hrs at 37°C	83
3.1.3.2	Differential regulation of HCPs 24hrs vs 96hrs at 31°C	85
3.1.3.3	Differential regulation of HCPs at 24hrs 37°C vs 31°C	87
3.1.3.4	Differential regulation of HCPs at 96hrs 37°C vs 31°C	89

3.1.3.5	Comparison of differentially regulated HCPs between 24hrs and 96hrs post temperature shift at 37°C and 31°C in a non-producing hybridoma cell line	91
3.1.4	Qualitative analysis of proteins identified by label free LC/MS profiling of conditioned media from a non-producing hybridoma cell line	94
3.1.5	Summary	96
<b>3.2</b>	<b>Characterisation of the Conditional Medium host cell protein profile of a non-producing CHO K1 cell line</b>	<b>98</b>
3.2.1	Growth profiling of non-producing CHO K1 cell line	98
3.2.2	Quantitative Label free LC/MS analysis of conditioned media from a CHO K1 non-producing cell line	100
3.2.3	Data analysis of differentially regulated proteins identified by quantitative label free LC/MS profiling of conditioned media from a CHO K1 non-producing cell line	101
3.2.3.1	Differential regulation of HCPs 24hrs vs 120hrs at 37°C	102
3.2.3.2	Differential regulation of HCPs 24hrs vs 120hrs at 31°C	105
3.2.3.3	Differential regulation of HCPs at 24hrs 37°C vs 31°C	108
3.2.3.4	Differential regulation of HCPs at 120hrs 37°C vs 31°C	111
3.2.3.5	Pearson correlation analysis of differentially regulated HCPs over 24, 72, 96 and 120 hours at 37°C	114
3.2.3.6	Pearson correlation analysis of differentially regulated HCPs over 24, 72, 96 and 120 hours at 31°C	116
3.2.3.7	Comparison of differentially regulated HCPs over time in culture at 37°C and 31°C in a non-producing CHO K1 cell line	118
3.2.3.8	Comparison of quantitative label-free LC/MS CHO-K1 HCP data normalised to cell number compared to normalisation by protein concentration	125
3.2.4	Bioinformatic analysis of differentially regulated protein lists	126
3.2.5	Qualitative analysis of proteins identified by quantitative label free LC/MS profiling of conditioned media from a non-producing CHO K1 cell line	132
3.2.5.1	Qualitative comparison of HCPs identified in temperature shifted and non-temperature shifted CHO K1 cultures	133
3.2.6	Summary	135
<b>3.3</b>	<b>Characterisation of the host cell protein profile of an IgG producing DP12 cell line</b>	<b>136</b>
3.3.1	Growth curves at 31°C and 37°C of IgG producing DP12 cell line	136
3.3.2	IgG depletion of conditioned media from IgG producing DP12 CHO cell line	138
3.3.3	Quantitative label free LC/MS analysis of conditioned media from a DP12 CHO cell line	140
3.3.4	Data analysis of differentially regulated proteins identified by quantitative label free LC/MS profiling of conditioned media from a DP12 CHO cell line	141
3.3.4.1	Differential regulation of HCPs 24hrs vs 144hrs at 37°C	145
3.3.4.2	Differential regulation of HCPs 24hrs vs 144hrs at 31°C	148
3.3.4.3	Differential regulation of HCPs at 24hrs 37°C vs 31°C	151

3.3.4.4	Differential regulation of HCPs at 144hrs 37°C vs 31°C	154
3.3.4.5	Western blot validation of HCPs identified as differentially-expressed following quantitative label-free LC/MS analysis of the cell culture supernatant from a CHO-K1 and DP12 cell line	157
3.3.4.6	Pearson correlation analysis of differentially regulated HCPs over 24, 72, 96, 120 and 144 hours at 37°C	159
3.3.4.7	Pearson correlation analysis of differentially regulated HCPs over 24, 72, 96, 120 and 144 hours at 31°C	161
3.3.4.8	Comparison of differentially regulated HCPs over time in culture at 37°C and 31°C in a DP12 cell line	163
3.3.4.9	Comparison of Pearson correlation analyses of differentially regulated HCPs over time in culture at 37°C and 31°C in a DP12 cell line	167
3.3.4.10	Comparison of differentially-regulated HCPs in the cell culture supernatant of a CHO-K1 and DP12 cell line.	171
3.3.4.11	Comparison of quantitative label-free LC/MS DP12 HCP data normalised to cell number compared to normalisation by protein concentration	173
3.3.5	Bioinformatic analysis of differentially regulated proteins	174
3.3.6	Qualitative analysis of proteins identified by label free LC/MS profiling of conditioned media from a DP12 CHO cell line	179
3.3.6.1	Qualitative comparison of HCPs identified in temperature shifted and non-temperature shifted DP12 cultures	179
3.3.7	Qualitative comparison of HCPs identified in a CHO K1 and DP12 cell line	181
3.3.7.1	Qualitative comparison of HCPs identified in a CHO K1 and a DP12 cell line under temperature shifted and non-temperature shifted conditions	183
3.3.8	Utilisation of CHO specific proteomic databases in the identification of HCPs in conditioned media samples from a non-producing CHO K1 cell line and an IgG secreting DP12 cell line.	186
3.3.8.1	Comparison of CHO NCBI and CHO BB proteomic databases in the identification of HCPs from a non-producing CHO K1 cell line	187
3.3.8.2	Comparison of CHO NCBI and CHO BB proteomic databases in the identification of HCPs from an IgG secreting DP12 cell line	189
3.3.9	Summary	191
<b>4.0</b>	<b>Phosphoproteomic Analysis of CHO cells</b>	<b>192</b>
<b>4.1</b>	<b>Phospho-Kinase Array Analysis of CHO SEAP cells subject to temperature shift</b>	<b>194</b>
4.1.1	Western blot validation of Total and phosphorylated ERK	198
4.1.2	Western blot validation of Total and phosphorylated MEK6	204
4.1.3	Western blot validation of Total and phosphorylated JNK	205
4.1.4	Western blot validation of Total and phosphorylated AKT1	207
4.1.5	Summary	209
<b>4.2</b>	<b>Phosphoproteomic profiling of CHO SEAP cells subject to temperature shift</b>	<b>210</b>
4.2.1	Growth of CHO SEAP cells subject to temperature shift	210
4.2.2	Qualitative analysis of phosphopeptide enrichment of CHO SEAP	213

4.2.2.1	Comparison of enrichment methods - Qualitative analysis of all phosphopeptides and phosphosites identified	214
4.2.2.2	Utilisation of CHO specific proteomic databases in the analysis of LC/MS data in a phosphoproteomic study	218
4.2.2.3	Phosphopeptides uniquely identified in temperature shifted or non-temperature shifted cultures	225
4.2.2.4	Summary	233
4.2.3	Quantitative Phosphoproteomic label free LC/MS analysis of CHO SEAP cells subject to temperature shift	234
4.2.4	Quantitative label free LC/MS analysis of CHO SEAP whole cell lysate	234
4.2.4.1	Bioinformatic analysis of differentially regulated proteins identified in the whole cell lysate of CHO SEAP cells subject to temperature shift	242
4.2.4.2	Summary	245
4.2.5	Quantitative label free LC/MS analysis of phosphoprotein enriched CHO SEAP lysate	246
4.2.5.1	Bioinformatic analysis of differentially regulated phosphoproteins	254
4.2.5.2	Summary	257
4.2.6	Quantitative label free LC/MS analysis of a Gallium oxide phosphopeptide enrichment of a phosphoprotein enriched CHO SEAP lysate	258
4.2.6.1	Summary	266
4.2.7	Quantitative label free LC/MS analysis of titanium dioxide phosphopeptide enriched CHO SEAP whole cell lysate	267
4.2.7.1	Bioinformatic analysis of differentially expressed phosphopeptides enriched using titanium dioxide	275
4.2.7.2	Summary	276
4.2.8	Quantitative label free LC/MS analysis of an Iron oxide phosphopeptide enriched CHO SEAP whole cell lysate	277
4.2.8.1	Bioinformatic analysis of differentially expressed phosphopeptides enriched using iron oxide	286
4.2.8.2	Bioinformatic analysis of all differentially expressed phosphopeptides enriched using gallium oxide, titanium dioxide and iron oxide	289
4.2.9	Western blot validation of differentially expressed phosphopeptides identified by quantitative label free LC/MS analysis of temperature shifted CHO SEAP cells	292
4.2.9.1	Summary	295
4.2.10	Overall summary of quantitative label free LC/MS analysis of CHO SEAP cells subjected to temperature shift.	295
<b>4.3</b>	<b>Overall summary of results from phosphoproteomic and host cell protein analysis of temperature shifted CHO cell lines</b>	<b>296</b>
<b>5.0</b>	<b>Host Cell Protein (HCP) Analysis Overview</b>	<b>298</b>
<b>5.1</b>	<b>Characterisation of HCP profiles from a non-producing hybridoma cell line using Mass Spectrometry</b>	<b>299</b>
5.1.1	Quantitative comparison of HCPs identified between early (24hrs) and late (96hrs) in temperature shifted and non-temperature shifted cultures	299
5.1.2	Quantitative comparison of HCPs identified between Temperature shifted and non-temperature shifted culture at 24hrs and 96hrs post temperature shift	301

5.1.3	Qualitative assessment of the HCP profile of a non-producing hybridoma cell line.	302
<b>5.2</b>	<b>Characterisation of HCP profiles from CHO K1 and DP12 cell lines using Mass Spectrometry</b>	<b>303</b>
5.2.1	Quantitative Label free LC/MS analysis of conditioned media from CHO K1 and DP12 cell lines	307
5.2.1.1	Growth factors	307
5.2.1.2	Apoptosis	311
5.2.1.3	Proteases	312
5.2.1.4	Reductases	318
5.2.1.5	Glycosidases	319
5.2.2	Qualitative analysis of CHO K1 and DP12 cell lines under temperature shifted conditions	324
5.2.2.1	Effect of temperature shift on the HCP profile	328
5.2.2.2	Qualitative analysis of secreted proteins	329
5.2.3	Utilisation of CHO specific proteomic databases in the analysis of LC/MS data in a HCP study	333
5.2.4	Overall Summary	334
<b>6.0</b>	<b>Phosphoproteomic Analysis of CHO Cells Overview</b>	<b>335</b>
<b>6.1</b>	<b>Phospho-Kinase Array of SEAP cells subject to temperature shift</b>	<b>337</b>
6.1.1	Phospho-ERK1/2 in Temperature shifted CHO cells	338
6.1.2	Phospho-AKT in Temperature shifted CHO cells	341
6.1.3	Phospho-JNK and MEK6 in Temperature shifted CHO cells	342
6.1.4	Summary	343
<b>6.2</b>	<b>Quantitative Phosphoproteomic label free LC/MS analysis of CHO SEAP cells subject to temperature shift</b>	<b>344</b>
6.2.1	Quantitative label free LC/MS analysis of CHO SEAP cells subject to temperature shift	344
6.2.1.1	Differential expression of proteins involved in translation	344
6.2.1.2	Comparison of this quantitative label free LC/MS analysis of temperature shifted CHO cells with other proteomic profiling of temperature shifted CHO cells in the literature	346
6.2.2	Label free LC/MS analysis of phosphoprotein-enriched lysates from CHO SEAP cells subject to temperature shift	348
6.2.2.1	Differential expression of phosphoproteins involved in cell structure	348
6.2.2.1.1	Ezrin	350
6.2.2.1.2	Annexin A6	350
6.2.2.1.3	GAPDH	350
6.2.2.2	Differential expression of phosphoproteins involved in translation	351
6.2.3	Label free LC/MS analysis of a Gallium oxide phosphopeptide enrichment of phosphoprotein enriched CHO SEAP cells subject to temperature shift	352
6.2.3.1	PITSLRE serine/threonine protein kinase CDC2L1	353
6.2.3.2	Nucleolar phosphoprotein p130	354
6.2.4	Qualitative assessment of titanium dioxide and iron oxide phosphopeptide enrichment methods	356

6.2.5	Differentially expressed phosphopeptides common to titanium dioxide and iron oxide phosphopeptide enrichment methods	358
6.2.5.1	eIF4g3	358
6.2.5.2	Protein NDRG1	359
6.2.5.3	GO analysis of differentially expressed phosphopeptides common to titanium dioxide and iron oxide phosphopeptide enrichment methods	360
6.2.5.3.1	Serine/threonine-protein kinase Nek1	360
6.2.5.3.2	Anaphase-promoting complex subunit 1	361
6.2.5.3.3	Differential expression of phosphopeptides in proteins involved in transcription	362
6.2.5.3.4	TRIM28 and ZKSCAN3	362
6.2.6	Differentially expressed phosphopeptides unique to titanium dioxide enrichment	364
6.2.6.1	Programmed cell death 4 (PDCD4)	365
6.2.6.2	Zyxin	365
6.2.7	Differentially expressed phosphopeptides unique to iron oxide enrichment	366
6.2.7.1	ATF-2 and ATF-7	366
6.2.7.2	DNA excision repair protein ERCC-6	367
6.2.7.3	Differential expression of phosphopeptides in proteins involved in cell division	367
6.2.8	Utilisation of CHO specific proteomic databases in the analysis of LC/MS data in a phosphoproteomic study	368
6.2.9	Summary	369
	<b>Conclusions</b>	<b>371</b>
	<b>Future work</b>	<b>374</b>
	<b>Bibliography</b>	<b>375</b>

## List of Figures

<b>Figure 1.2.1</b> Diagram of the key responses to reduced culture temperature conditions seen in CHO cells.	9
<b>Figure 1.5.1</b> Schematic of the basic components of a mass spectrometer	30
<b>Figure 1.5.2</b> The proposal for a common nomenclature for sequence ions in mass spectra of peptides as put forward by Roepstorff and Fohlman (1984).	31
<b>Figure 1.5.3</b> Full MS of singly charged peptide as denoted by difference of 1Da between the isotopic peaks 842.51 and 843.51.	33
<b>Figure 1.5.4</b> Example of the peptide VATVASLPR and the ion ladder made up of b-ions and y-ions resulting from its fragmentation.	35
<b>Figure 1.5.5</b> Diagram of a phosphorylated MS/MS sequence in red overlapped with the unphosphorylated sequence in black.	41
<b>Figure 1.5.6</b> MS/MS showing a predominant peak corresponding to the neutral loss of H <sub>3</sub> PO <sub>4</sub> from a phosphorylated Serine.	42
<b>Figure 1.6.1</b> Overview of the process workflow for a protein array.	51
<b>Figure 1.6.2</b> Schematic of phosphopeptide enrichment strategies.	57
<b>Figure 3.1.1</b> Growth curve of a non-producing hybridoma cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions.	81
<b>Figure 3.1.2</b> PCA plot for the non-producing hybridoma cell line sample groups 24hrs (pink) and 96hrs (blue) at 37°C generated during quantitative label-free LC/MS data analysis using Progenesis software.	83
<b>Figure 3.1.3</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as STIP1 and Cathepsin D.	84
<b>Figure 3.1.4</b> PCA plot for non-producing hybridoma cell line sample groups compared at 24hrs (pink) and 96hrs (blue) at 31°C generated during label-free LC/MS data analysis using Progenesis software.	85
<b>Figure 3.1.5</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as VCP and Calumenin.	86
<b>Figure 3.1.6</b> PCA plot for non-producing hybridoma cell line sample groups 37°C (pink) and 31°C (blue) after 24hrs in culture generated during label-free LC/MS data analysis using Progenesis software.	87
<b>Figure 3.1.7</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as Laminin subunit gamma-1 and Nidogen 2.	88
<b>Figure 3.1.8</b> PCA plot for non-producing hybridoma cell line sample groups 37°C (pink) and 31°C (blue) after 120hrs in culture generated during label-free LC/MS data analysis using Progenesis software.	89
<b>Figure 3.1.9</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as Endoplasmic reticulum aminopeptidase 1 and Cathepsin B.	90
<b>Figure 3.1.10</b> Overlap of the list of proteins identified in the conditioned media of a non-producing hybridoma cell line grown at 37°C and 31°C over 144 hours in stationary suspension culture.	95
<b>Figure 3.2.1</b> Growth curve of CHO-K1 non-producing cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions.	99
<b>Figure 3.2.2</b> PCA plot for CHO-K1 non-producing cell line sample groups 24hrs (pink) and 120hrs (blue) at 37°C generated during quantitative label-free LC/MS data analysis using Progenesis software.	102



<b>Figure 3.2.3</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as Guanine nucleotide-binding protein subunit beta-2-like 1 and Beta-galactosidase.	104
<b>Figure 3.2.4</b> PCA plot for non-producing CHO-K1 cell line sample groups 24hrs (pink) and 120hrs (blue) at 31°C generated during quantitative label-free LC/MS data analysis using Progenesis software.	105
<b>Figure 3.2.5</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as Cathepsin D and Pentraxin-related protein PTX3.	107
<b>Figure 3.2.6</b> PCA plot for CHO-K1 non-producing cell line sample groups 37°C (pink) and 31°C (blue) after 24hrs in culture generated during quantitative label-free LC/MS data analysis using Progenesis software.	108
<b>Figure 3.2.7</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as GRP78 and the glycosidase Lysosomal alpha-glucosidase.	110
<b>Figure 3.2.8</b> PCA plot for CHO-K1 non-producing cell line sample groups 37°C (pink) and 31°C (blue) after 120hrs in culture generated during quantitative label-free LC/MS data analysis using Progenesis software	111
<b>Figure 3.2.9</b> Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the apoptotic protein Guanine nucleotide-binding protein subunit beta-2-like 1 and the protease Matrix metalloproteinase-19.	113
<b>Figure 3.2.10</b> Progenesis label-free output of the tumour suppressor semaphorin-3B showing its increase in expression from 24hrs to 120hrs at 37°C and the output from excel on which Pearson correlation calculation is based.	115
<b>Figure 3.2.11</b> Progenesis label-free output of the secreted protein Suprabasin showing its increase in expression from 24hrs to 120hrs at 37°C and the output from excel on which Pearson correlation calculation is based.	115
<b>Figure 3.2.12</b> Progenesis label-free output of the secreted protein Lactadherin showing its increase in expression from 24hrs to 120hrs at 31°C and the output from excel on which Pearson correlation calculation is based.	117
<b>Figure 3.2.13</b> Progenesis label-free output of the secreted protein Neuroblast differentiation-associated protein AHNAK showing its increase in expression from 24hrs to 120hrs at 31°C and the output from excel on which Pearson correlation calculation is based.	117
<b>Figure 3.2.14</b> Progenesis label-free software outputs and a table of proteins common to 24hrs vs 120hrs 37°C and 24hrs vs 120hrs 31°C differentially-expressed protein lists from non-producing CHO-K1 cell line.	119
<b>Figure 3.2.15</b> Shows the differential expression of the HCPs Endoplasmin and Complement C1r-A subcomponent between early and late stages in culture at 37°C and 31°C. The differential expression between 37°C and 31°C at 120hrs is also shown.	121
<b>Figure 3.2.16</b> Shows the differential expression of the HCPs Semaphorin-3B and HSPG2 between early and late stages in culture at 37°C and 31°C. The differential expression between 37°C and 31°C at 120hrs is also shown.	122
<b>Figure 3.2.17</b> Pearson correlation plots and the outputs from Progenesis label-free software of the intracellular proteins Pyruvate kinase isozymes M1/M2 and Vimentin as they increase over time in culture at 31°C and 37°C.	124

<b>Figure 3.2.18</b> Graphical representation of Table 3.2.20 depicting GO cellular component enrichment for differentially-expressed HCPs identified in the conditioned media of a CHO-K1 non-producing cell line at 37°C when 24hrs was compared to 120hrs.	129
<b>Figure 3.2.19</b> Graphical representation of Table 3.2.21 depicting GO cellular component enrichment for differentially-expressed HCPs identified in the conditioned media of a CHO-K1 non-producing cell line at 31°C when 24hrs was compared to 120hrs.	131
<b>Figure 3.2.20</b> Overlap of the list of proteins identified in the conditioned media of a non-producing CHO-K1 cell line grown at 37°C and 31°C over 168 hours in shake flask culture.	134
<b>Figure 3.3.1</b> Growth curve of an IgG producing DP12 cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions.	137
<b>Figure 3.3.2</b> Western blot image of IgG secreted by a DP12 cell line.	138
<b>Figure 3.3.3</b> Concentration of IgG produced by a DP12 cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions was determined by ELISA assay.	139
<b>Figure 3.3.4</b> Concentration of IgG remaining in conditioned media sample from a DP12 cell line grown at 31°C and 37°C shows post IgG depletion by Protein A chromatography as determined by ELISA assay.	139
<b>Figure 3.3.5</b> PCA plot for the IgG secreting DP12 conditioned media sample groups 24hrs (pink) and 144hrs (blue) at 37°C generated during quantitative label-free LC/MS data analysis using Progenesis software.	145
<b>Figure 3.3.6</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the secreted protein Suprabasin and the growth factor Bone morphogenetic protein 1.	147
<b>Figure 3.3.7</b> PCA plot for the IgG secreting DP12 conditioned media sample groups 24hrs (pink) and 144hrs (blue) at 31°C generated during quantitative label-free LC/MS data analysis using Progenesis software.	148
<b>Figure 3.3.8</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the glycosidase Alpha-L-iduronidase and the apoptotic protein Amyloid beta A4 protein.	150
<b>Figure 3.3.9</b> PCA plot for the IgG secreting DP12 conditioned media sample groups 37°C (pink) and 31°C (blue) after 24hrs in culture generated during quantitative label-free LC/MS data analysis using Progenesis software	151
<b>Figure 3.3.10</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as Beta-galactosidase and the growth factor Vascular endothelial growth factor C.	153
<b>Figure 3.3.11</b> PCA plot for the IgG secreting DP12 conditioned media sample groups 37°C (pink) and 31°C (blue) after 144hrs in culture generated during label-free LC/MS data analysis using Progenesis software	154
<b>Figure 3.3.12</b> Examples of Progenesis label-free outputs of differentially-expressed HCPs, such as Bone morphogenetic protein 1 and Vascular endothelial growth factor A.	156
<b>Figure 3.3.13</b> Western blot confirming the down-regulation of the intracellular protein Sulfated Glycoprotein 1 at 31°C when temperature-shifted culture is compared to non-temperature-shifted culture at 24hrs and 144hrs in the conditioned media of the DP12 cell line.	157

<b>Figure 3.3.14</b> Western blot confirming the up-regulation of the intracellular protein GRP78 at 31°C when temperature-shifted culture is compared to non-temperature-shifted culture at 24hrs and 120hrs in the conditioned media of the CHO-K1 cell line.	158
<b>Figure 3.3.15</b> Western blot confirming the up-regulation of the intracellular protein GRP78 at 144hrs when 144hrs is compared to 24hrs at 37°C and at 31°C in the conditioned media of the DP12 cell line.	158
<b>Figure 3.3.16</b> Western blot confirming the up-regulation of the protease BMP1 at 120hrs when 120hrs is compared to 24hrs at 37°C in the conditioned media of the CHO-K1 cell line.	158
<b>Figure 3.3.17</b> Western blot confirming the up-regulation of the protease MMP19 at 31°C when temperature-shifted culture is compared to non-temperature-shifted culture in the conditioned media of the CHO-K1 cell line.	159
<b>Figure 3.3.18</b> Progenesis label-free output of the secreted protein Nidogen-1 showing its decrease in expression from 24hrs to 144hrs at 37°C. Output from excel on which Pearson correlation calculation is based.	160
<b>Figure 3.3.19</b> Progenesis label-free output of the protein Elongation factor 2 showing its increase in expression from 24hrs to 144hrs at 37°C. Output from excel on which Pearson correlation calculation is based.	160
<b>Figure 3.3.20</b> Progenesis label-free output of the secreted protein Complement C3 showing its decrease in expression from 24hrs to 144hrs at 31°C. Output from excel on which Pearson correlation calculation is based.	162
<b>Figure 3.3.21</b> Progenesis label-free output of the protease Cathepsin D showing its increase in expression from 24hrs to 144hrs at 31°C. Output from excel on which Pearson correlation calculation is based.	162
<b>Figure 3.3.22</b> Shows changes in the expression of the proteases Cathepsin B and Complement C1r-A subcomponent and the glycosidase Glucosylceramidase between early and late stages in culture at both 37°C and 31°C.	164
<b>Figure 3.3.23</b> Progenesis label-free outputs and information on the differential expression of the secreted proteins MMP9, Sulfated glycoprotein 1 and Suprabasin as they change in relative abundance between early and late stages in culture.	166
<b>Figure 3.3.24</b> Pearson correlation plots and the outputs from Progenesis label-free software of the protease ADAMTS1 (A) and the reductase TXNDC5 (B) show how they change over time in culture at both 37°C and 31°C.	168
<b>Figure 3.3.25</b> Pearson correlation plots and the outputs from Progenesis label-free software of the secreted protein Tissue alpha-L-fucosidase shows how it changes over time in culture at both 37°C and 31°C.	170
<b>Figure 3.3.26</b> Progenesis label-free outputs and information on HCPs identified as being differentially-expressed in both CHO-K1 and DP12 cell lines.	172
<b>Figure 3.3.27</b> Graphical representation of Table 3.3.20 depicting GO cellular component enrichment for differentially-expressed HCPs identified in the conditioned media of a DP12 cell line at 37°C when 24hrs was compared to 144hrs.	176
<b>Figure 3.3.28</b> Graphical representation of Table 3.3.32 depicting GO cellular component enrichment for differentially-expressed HCPs identified in the conditioned media of a DP12 cell line at 31°C when 24hrs was compared to 144hrs.	178

<b>Figure 3.3.29</b> Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line grown at 37°C and 31°C over 216 hours in shake flask culture.	180
<b>Figure 3.3.30</b> Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line and a non-producing CHO-K1 cell line grown at 37°C and 31°C in shake flask culture.	182
<b>Figure 3.3.31</b> Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line and a non-producing CHO-K1 cell line grown at 37°C in shake flask culture.	184
<b>Figure 3.3.32</b> Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line and a non-producing CHO-K1 cell line grown at 31°C in shake flask culture.	185
<b>Figure 3.3.33</b> Overlap of the list of proteins identified in the conditioned media of a non-producing CHO-K1 cell line.	188
<b>Figure 3.3.34</b> Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line.	190
<b>Figure 4.1.1</b> Image of kinase array containing 26 different kinases including 9 MAP kinases.	195
<b>Figure 4.1.2</b> Growth curve of CHO SEAP cell line under both temperature-shifted and non-temperature shifted conditions.	196
<b>Figure 4.1.3</b> Growth curve of DP12 cell line under both temperature-shifted and non-temperature shifted conditions.	196
<b>Figure 4.1.4</b> Graph showing the final titre of SEAP secreting CHO cells grown under temperature-shifted (31°C) and non-temperature shifted (37°C) conditions.	197
<b>Figure 4.1.5</b> Western blot image of Phosphorylated ERK and total ERK in a CHO SEAP cell line over time in culture at 31°C and 37°C.	200
<b>Figure 4.1.6</b> Western blot image of Phosphorylated ERK1/2 and total ERK1/2 in a DP12 cell line over time in culture at 31°C and 37°C.	201
<b>Figure 4.1.7</b> Western blot image of Phosphorylated Mek1/2 and total Mek1/2 in a CHO SEAP cell line over time in culture at 31°C and 37°C.	203
<b>Figure 4.1.8</b> Western blot image of Phosphorylated Mek6 in a CHO SEAP cell line over time in culture at 31°C and 37°C.	204
<b>Figure 4.1.9</b> Western blot image of Phosphorylated JNK and total JNK in a CHO SEAP cell line over time in culture at 31°C and 37°C.	206
<b>Figure 4.1.10</b> Western blot image of Phosphorylated AKT1 and total AKT1 in a CHO SEAP cell line over time in culture at 31°C and 37°C.	208
<b>Figure 4.1.11</b> Western blot image of Phosphorylated AKT1 in a DP12 cell line over time in culture at 31°C and 37°C.	209
<b>Figure 4.2.1</b> Growth curve of SEAP-secreting CHO cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions.	211
<b>Figure 4.2.2</b> Graph showing the final titre of SEAP-secreting CHO cells grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions.	212
<b>Figure 4.2.3</b> Venn diagram of the number of phosphopeptides that are common and unique to each phosphopeptide enrichment method following enrichment of CHO SEAP cells grown under temperature-shifted and non-temperature-shifted conditions.	216

<b>Figure 4.2.4</b> Graph of the number of phosphorylation sites per peptide as identified by titanium, iron and gallium phosphopeptide enrichment methods following LC/MS analysis of CHO SEAP cells subject to temperature-shift.	217
<b>Figure 4.2.5</b> Venn diagram showing the number of phosphopeptides identified when raw files from LC/MS analysis of a phosphopeptide enriched CHO SEAP sample were searched against CHO BB, CHO NCBI and Swiss-Prot human, mouse, rat and CHO databases.	219
<b>Figure 4.2.6</b> Example of the use of BLAST analysis to identify a protein based on sequence homology in a related species with a well annotated database.	221
<b>Figure 4.2.7</b> Overlap of the list of phosphopeptides identified after the iron oxide enrichment of a whole cell lysate from a CHO SEAP cell line.	223
<b>Figure 4.2.8</b> Overlap of the list of phosphopeptides identified after the titanium dioxide enrichment of a whole cell lysate from a CHO SEAP cell line.	223
<b>Figure 4.2.9</b> Overlap of the list of phosphopeptides identified after the gallium oxide enrichment of a whole cell lysate pre-enriched for phosphoproteins from a CHO SEAP cell line.	224
<b>Figure 4.2.10</b> PCA plot for CHO SEAP whole cell lysate sample groups 37°C (pink) and 31°C (blue) generated during label-free LC/MS data analysis using Progenesis software.	235
<b>Figure 4.2.11</b> GO biological process enrichment for differentially-expressed proteins identified in quantitative label-free LC/MS analysis of whole cell lysate from CHO SEAP cells sampled at 36hrs post-temperature-shift.	244
<b>Figure 4.2.12</b> PCA plot for CHO SEAP phosphoprotein-enriched whole cell lysate sample groups 37°C (pink) and 31°C (blue) generated during quantitative label-free LC/MS data analysis using Progenesis software	247
<b>Figure 4.2.13</b> GO biological process enrichment for differentially-expressed phosphoproteins identified in quantitative label-free LC/MS analysis of phosphoprotein-enriched samples from CHO SEAP cells sampled at 36hrs post temperature-shift.	255
<b>Figure 4.2.14</b> PCA plot for CHO SEAP gallium phosphopeptide enriched whole cell lysate, phosphoprotein pre-enriched, sample groups 37°C (pink) and 31°C (blue) generated during quantitative label-free LC/MS data analysis using Progenesis software.	258
<b>Figure 4.2.15</b> MS/MS fragmentation spectrum output from a MASCOT search resulting in the identification of the phosphopeptide “DLLSDLQDISDSER” belonging to the protein PITSLRE serine/threonine-protein kinase CDC2L1.	263
<b>Figure 4.2.16</b> MS/MS fragmentation spectrum output from a SEQUEST search resulting in the identification of the phosphopeptide “DLLSDLQDISDSER” belonging to the protein PITSLRE serine/threonine-protein kinase CDC2L1.	264
<b>Figure 4.2.17</b> Progenesis label-free LC/MS output showing normalised abundance of the phosphopeptide “DLLSDLQDISDSER”.	265
<b>Figure 4.2.18</b> PCA plot for CHO SEAP titanium dioxide Phosphopeptide enriched whole cell lysate sample groups 37°C (pink) and 31°C (blue) generated during quantitative label-free LC/MS data analysis using Progenesis software.	267
<b>Figure 4.2.19</b> MS/MS fragmentation spectrum output from a MASCOT search resulting in the identification of the phosphopeptide “TASGSSVTSLEGPR” belonging to Protein NDRG1.	272

<b>Figure 4.2.20</b> MS/MS fragmentation spectrum output from a SEQUEST search resulting in the identification of the phosphopeptide “TASGSSVTSLEGPR” belonging to Protein NDRG1.	273
<b>Figure 4.2.21</b> Progenesis label-free LC/MS output showing normalised abundance of the phosphopeptide “TASGSSVTSLEGPR” from the Protein NDRG1.	274
<b>Figure 4.2.22</b> PCA plot for CHO SEAP Iron phosphopeptide enriched whole cell lysate sample groups 37°C (pink) and 31°C (blue) generated during quantitative label-free LC/MS data analysis using Progenesis software.	278
<b>Figure 4.2.23</b> MS/MS fragmentation spectrum output from a MASCOT search resulting in the identification of the phosphopeptide “SPVAATVVQR” from the protein Eukaryotic translation initiation factor 4 gamma 3.	283
<b>Figure 4.2.24</b> MS/MS fragmentation spectrum output from a SEQUEST search resulting in the identification of the phosphopeptide “SPVAATVVQR” belonging to Eukaryotic translation initiation factor 4 gamma 3.	284
<b>Figure 4.2.25</b> Progenesis label-free LC/MS output showing normalised abundance of the phosphopeptide “SPVAATVVQR”.	285
<b>Figure 4.2.26</b> GO biological process enrichment for differentially-expressed phosphopeptides identified using quantitative label-free LC/MS analysis of gallium oxide, titanium dioxide and iron oxide enriched samples from CHO SEAP cells sampled 36hrs post-temperature-shift.	291
<b>Figure 4.2.27</b> Western blot validation of the differential expression of a phosphopeptide from the protein NDRG1 identified as being up-regulated at 31°C compared to 37°C.	293
<b>Figure 4.2.28</b> Western blot showing that no bands for the total protein NDRG1 were detected in CHO despite NDRG1.	294
<b>Figure 4.2.29</b> Western blot validation of the differential expression of a phosphopeptide from the protein ATF2 identified as being down-regulated at 31°C compared to 37°C.	294
<b>Figure 6.0.1</b> Schematic of the different approaches used to profile the phosphoproteome of CHO cells under temperature-shift conditions using mass spectrometry.	336
<b>Figure 6.1.1</b> Images shows simplified schematic of MEK – ERK pathway.	339
<b>Figure 6.1.2</b> Image showing the variety of cellular processes that Erk is involved in controlling and the complexity of the related pathways.	340
<b>Figure 6.1.3</b> Schematic of pathways and cellular processes regulated by AKT.	342

## List of Tables

<b>Table 1.2.1</b> Effect of temperature-shift on specific productivity on CHO cells.	9
<b>Table 1.3.1</b> Overview of some 'desirable' characteristics in a recombinant protein producing cell line.	14
<b>Table 1.3.2</b> Table summarising some of the proteomic profiling investigations conducted date in CHO cells that have furthered our understanding of some of the cellular mechanisms involved in important cell attributes such as growth rate, productivity, apoptosis and metabolic shift.	15
<b>Table 1.5.1</b> List of amino acids, their 1 and 3 letter codes, and associated monoisotopic masses.	34
<b>Table 1.5.2</b> List of modifications for amino acids and the mass difference associated with the modification.	39
<b>Table 1.6.1</b> Table summarising the advantages and disadvantages of phosphoprotein / phosphopeptide enrichment strategies used in the analysis of the phosphoproteome. Table adapted from (Harsha and Pandey 2010).	48
<b>Table 2.1.1</b> Table of cell lines used in experiments presented in this thesis.	62
<b>Table 2.5.1</b> List of primary antibodies used in the experiments used in this thesis including those used to detect phosphorylated and non-phosphorylated versions of the proteins studied.	75
<b>Table 2.5.2</b> Shows list of kinases and the specific site of phosphorylation detected by the kinase array.	76
<b>Table 3.1.9:</b> Table of differentially-expressed proteins common to 24hrs vs 96hrs 37°C and 24hrs vs 96hrs 31°C differentially expressed protein lists from non-producing hybridoma cell line.	92
<b>Table 3.2.1:</b> Number of differentially expressed proteins at 24hrs compared to 120hrs at 37°C in a CHO K1 non-producing cell culture supernatant.	103
<b>Table 3.2.4:</b> Number of differentially expressed proteins at 24hrs compared to 120hrs at 31°C in a CHO K1 non-producing cell culture supernatant.	106
<b>Table 3.2.7:</b> Number of differentially expressed proteins between 31°C and 37°C at 24hrs post temperature shift in a CHO K1 non-producing cell culture supernatant.	109
<b>Table 3.2.10:</b> Number of differentially expressed proteins between 31°C and 37°C at 120hrs post temperature shift in a CHO K1 non-producing cell culture supernatant.	112
<b>Table 3.2.18:</b> The use of a correction factor was investigated to facilitate normalisation of the data by cell number (as opposed to normalisation by protein concentration). This table show a comparison of the number of differentially-expressed HCPs identified in the conditioned media of a CHO-K1 cell line with and without the correction factor applied.	126
<b>Table 3.2.19:</b> List of tables containing quantitative information on proteins classified as secreted by the Uniprot protein database.	127
<b>Table 3.2.20:</b> GO cellular component enrichment for differentially expressed host cell proteins identified in the conditioned media of a CHO K1 non-producing cell line at 37°C when 24hrs was compared to 120hrs (post temperature shift for a culture grown in parallel) analysed using quantitative label free LC/MS. An adjusted p-value of $\leq 0.05$ (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.	128

<b>Table 3.2.21:</b> GO cellular component enrichment for differentially expressed host cell proteins identified in the conditioned media of a CHO K1 non-producing cell line at 31°C when 24hrs was compared to 120hrs (post temperature shift) analysed using quantitative label free LC/MS. An adjusted p-value of $\leq 0.05$ (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.	130
<b>Table 3.3.1:</b> Top 10 differentially regulated proteins, based on peptide number, when all features of the protein profile from 24hrs in a DP12 culture was compared to 144hrs and normalised to an internal standard, Cytochrome C. These proteins were then identified in the same comparison but where all features were normalised to proteins that did not change in expression. This allowed for direct comparison to be made between Cytochrome C Normalisation (C) and the Normal means of Normalisation (N), based on peptide count, ion score, ANOVA, fold change and what condition the proteins is most highly expressed in.	143
<b>Table 3.3.2:</b> Top 10 differentially regulated proteins, based on peptide number, when all features of the protein profile from 24hrs in a DP12 culture was compared to 144hrs and normalised to proteins that did not change in expression. These proteins were then identified in the same comparison but where all features were normalised to an internal standard, Cytochrome C. This allowed for direct comparison to be made between the Normal means of Normalisation (N) and Cytochrome C Normalisation (C), based on peptide count, ion score, ANOVA, fold change and what condition the proteins is most highly expressed in.	144
<b>Table 3.3.3:</b> Number of differentially expressed proteins at 24hrs compared to 144hrs at 37°C in an IgG secreting DP12 cell culture supernatant.	146
<b>Table 3.3.6:</b> Number of differentially expressed proteins at 24hrs compared to 144hrs at 31°C in an IgG secreting DP12 cell culture supernatant.	149
<b>Table 3.3.9</b> Number of differentially expressed proteins between 31°C and 37°C at 24hrs post temperature shift in the culture supernatant of an IgG producing DP12 cell line.	152
<b>Table 3.3.12:</b> Number of differentially expressed proteins between 31°C and 37°C at 144hrs post temperature shift in the culture supernatant of an IgG producing DP12 cell line.	155
<b>Table 3.3.23:</b> Table indicating the number of differentially expressed HCPs identified in a CHO K1 and DP12 cell line compared to the number of differentially expressed HCPs common to both cell lines.	171
<b>Table 3.3.28:</b> The use of a correction factor was investigated to facilitate normalisation of the data by cell number (as opposed to normalisation by protein concentration). This table show a comparison of the number of differentially-expressed HCPs identified in the conditioned media of a DP12 cell line with and without the correction factor applied.	173
<b>Table 3.3.29:</b> List of tables containing quantitative information on proteins classified as secreted by the Uniprot protein database.	174
<b>Table 3.3.30:</b> GO cellular component enrichment for differentially expressed host cell proteins identified in the conditioned media of a DP12 CHO cell line at 37°C when 24hrs was compared to 144hrs (post temperature shift for a culture grown in parallel) analysed using quantitative label free LC/MS. An adjusted p-value of $\leq 0.05$ (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.	175



<b>Table 3.3.31:</b> GO cellular component enrichment for differentially expressed host cell proteins identified in the conditioned media of a DP12 cell line at 31°C when 24hrs was compared to 144hrs (post temperature shift) analysed using quantitative label free LC/MS. An adjusted p-value of $\leq 0.05$ (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.	177
<b>Table 3.3.56</b> Table displaying the number of proteins identified in the conditioned media of a CHO K1 cell line grown under temperature shift and non-temperature shift cultures with samples taken at 24, 72, 96 and 120 hours post temperature shift. RAW files from the mass spectral analysis were searched against the CHO NCBI and CHO BB protein databases. Columns are divided into those proteins identified by 1 peptide and proteins identified by 2 or more peptides.	188
<b>Table 3.3.60:</b> Table displaying the number of proteins identified in the conditioned media of a DP12 cell line grown under temperature shift and non-temperature shift cultures with samples taken at 24, 72, 96, 120 and 144 hours post temperature shift. RAW files from the mass spectral analysis were searched against the CHO NCBI and CHO BB protein databases. Columns are divided into those proteins identified by 1 peptide and proteins identified by 2 or more peptides.	190
<b>Table 4.1.1</b> 12 Kinases identified using a Phosphokinase array as having a potential change in phosphorylation status in CHO SEAP cell 24hrs post-temperature shift. This table represents the average fold change observed from duplicate biological samples analysed on the array (n=2).	195
<b>Table 4.2.1:</b> Table indicating the total number of phosphopeptides identified using each phosphopeptide enrichment method following enrichment of CHO SEAP cells grown under temperature shifted and non-temperature shifted conditions	216
<b>Table 4.2.16</b> Symbols denoting the amino acids Serine, Tyrosine and Threonine.	220
<b>Table 4.2.26</b> This table shows the number of phosphopeptides that are considered to be ‘potentially unique’ to temperature-shifted and non-temperature-shifted cultures using gallium, titanium and iron enrichment methods. This table also contains the table number for the list of potentially unique phosphopeptides from each enrichment method and each culture temperature.	226
<b>Table 4.2.27</b> List of phosphopeptides that were only identified in CHO SEAP cells at 37°C using titanium dioxide phosphopeptide enrichment.	227
<b>Table 4.2.28</b> List of phosphopeptides that were only identified in CHO SEAP cells at 31°C using titanium dioxide phosphopeptide enrichment.	228
<b>Table 4.2.29</b> List of phosphopeptides that were only identified in CHO SEAP cells at 37°C using iron oxide phosphopeptide enrichment.	229
<b>Table 4.2.30</b> List of phosphopeptides that were only identified in CHO SEAP cells at 31°C using iron oxide phosphopeptide enrichment.	230
<b>Table 4.2.31</b> List of phosphopeptides that were only identified in CHO SEAP cells at 37°C using gallium oxide phosphopeptide enrichment of phosphoprotein-enriched CHO SEAP whole cell lysate samples subject to temperature-shift.	231
<b>Table 4.2.32</b> List of phosphopeptides that were only identified in CHO SEAP cells at 31°C using gallium oxide phosphopeptide enrichment of phosphoprotein-enriched CHO SEAP whole cell lysate samples subject to temperature-shift.	232

<b>Table 4.2.33</b> List of differentially-expressed proteins, down-regulated at 31°C compared to 37°C, identified in quantitative label-free LC/MS analysis of CHO SEAP whole cell lysates from cells sampled at 36hrs post-temperature-shift.	237
<b>Table 4.2.34</b> List of differentially-expressed proteins, up-regulated at 31°C compared to 37°C, identified in quantitative label-free LC/MS analysis of CHO SEAP whole cell lysates from cells sampled at 36hrs post-temperature-shift.	239
<b>Table 4.2.35</b> GO biological process (BP) enrichment for differentially-expressed proteins identified in quantitative label-free LC/MS analysis of whole cell lysate from CHO SEAP cells sampled at 36hrs post-temperature-shift.	243
<b>Table 4.2.36</b> GO biological process (BP) enrichment for differentially-expressed proteins identified in quantitative label-free LC/MS analysis of whole cell lysate from CHO SEAP cells sampled at 36hrs post-temperature-shift.	244
<b>Table 4.2.37</b> List of phosphoproteins, the expression of which decreased at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift.	249
<b>Table 4.2.38</b> List of phosphoproteins, the expression of which increased at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift.	251
<b>Table 4.2.39</b> List of proteins common to two lists of differentially-expressed proteins, down-regulated at 31°C compared to 37 °C, identified in quantitative label-free LC/MS analysis of CHO SEAP whole cell lysate and phosphoprotein-enriched CHO SEAP whole cell lysate.	252
<b>Table 4.2.40</b> List of proteins common to two lists of differentially-expressed proteins, up-regulated at 31°C, identified in quantitative label-free LC/MS analysis of CHO SEAP whole cell lysate and phosphoprotein-enriched CHO SEAP whole cell lysate.	253
<b>Table 4.2.41</b> GO biological process enrichment for differentially-expressed phosphoproteins identified in quantitative label-free LC/MS analysis of phosphoprotein-enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift.	255
<b>Table 4.2.42</b> GO biological process enrichment for phosphoproteins identified as up-regulated at 31°C compared to 37°C in quantitative label-free LC/MS analysis of phosphoprotein-enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift.	256
<b>Table 4.2.43</b> GO biological process enrichment for phosphoproteins identified as down-regulated at 31°C compared to 37°C in quantitative label-free LC/MS analysis of phosphoprotein-enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift.	256
<b>Table 4.2.44</b> List of differentially-expressed phosphopeptides, down-regulated at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift.	261
<b>Table 4.2.45</b> List of differentially-expressed phosphopeptides, up-regulated at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift.	261

<b>Table 4.2.46</b> List of differentially-expressed phosphopeptides, down-regulated at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift.	269
<b>Table 4.2.47</b> List of differentially-expressed phosphopeptides, up-regulated at 31°C compared to 37°C, obtained from label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift.	270
<b>Table 4.2.48</b> List of differentially-expressed phosphopeptides, down-regulated at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift.	279
<b>Table 4.2.49</b> List of differentially-expressed phosphopeptides, up-regulated at 31°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift.	281
<b>Table 4.2.50</b> List of proteins and associated differentially-expressed phosphopeptides common to quantitative label-free LC/MS analysis of titanium dioxide phosphopeptide enriched (TiO <sub>2</sub> ) and iron oxide phosphopeptide enriched (Fe <sup>+</sup> ) CHO SEAP whole cell lysates subject to temperature-shift. This table shows the phosphopeptides that were down-regulated at 31°C.	287
<b>Table 4.2.51</b> List of proteins and associated differentially-expressed phosphopeptides common to quantitative label-free LC/MS analysis of titanium dioxide phosphopeptide enriched (TiO <sub>2</sub> ) and iron oxide phosphopeptide enriched (Fe <sup>+</sup> ) CHO SEAP whole cell lysates subject to temperature-shift. This table shows the phosphopeptides that were up-regulated at 31°C.	288
<b>Table 4.2.52</b> GO biological process enrichment for differentially-expressed phosphopeptides identified using quantitative label-free LC/MS analysis of gallium oxide, titanium dioxide and iron oxide enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of a p-value ≤0.05.	290
<b>Table 5.2.1</b> Summary table of differentially-expressed growth factors identified following quantitative label-free LC/MS analysis of conditioned media from CHO-K1 and DP12 cells.	309
<b>Table 5.2.2</b> List of differentially-expressed proteases identified in the cell culture supernatant of a CHO-K1 and DP12 cell line. (Protein function was obtained following a search of the protein names against the Uniprot protein database.)	314
<b>Table 5.2.3</b> Summary table of differentially-expressed proteases identified following quantitative label-free LC/MS analysis of conditioned media from CHO-K1 and DP12 cells.	315
<b>Table 5.2.4</b> Table summarising the potential role and effect of the glycocomponent of glycoproteins (adapted from (Walsh and Jefferis 2006)).	320
<b>Table 5.2.5</b> List of 23 differentially-expressed glycosidases identified in the cell culture supernatant of a CHO-K1 and DP12 cell line. (Protein function was obtained following a search of the protein names against the Uniprot protein database.)	321
<b>Table 5.2.6</b> Summary of differentially-expressed glycosidases identified following quantitative label-free LC/MS analysis of conditioned media from CHO-K1 and DP12 cells.	323

**Table 6.2.1** Summary table comparing quantitative label-free LC/MS analysis of temperature-shifted CHO cells conducted in this thesis with other proteomic profiling of temperature-shifted CHO cells in the literature.

347

## Appendix Table of Contents (Enclosed on CD)

<b>Appendix A</b>	1
<b>Table 3.1.1:</b> List of host cell proteins ( $\geq 1$ Peptides), the expression of which increased after 96hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from a non-producing hybridoma cell line.	2
<b>Table 3.1.2:</b> List of host cell proteins ( $\geq 1$ Peptides), the expression of which decreased after 96hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from a non-producing hybridoma cell line.	3
<b>Table 3.1.3:</b> List of host cell proteins ( $\geq 1$ Peptides), the expression of which increased after 96hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from a non-producing hybridoma cell line.	5
<b>Table 3.1.4:</b> List of host cell proteins ( $\geq 1$ Peptides), the expression of which decreased after 96hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from a non-producing hybridoma cell line.	6
<b>Table 3.1.5:</b> List of host cell proteins ( $\geq 1$ Peptides), the expression of which increased at 31°C when compared to 37°C after 24hrs in temperature shifted culture, identified during quantitative label free LC/MS analysis of conditioned media from a non-producing hybridoma cell line.	7
<b>Table 3.1.6:</b> List of host cell proteins ( $\geq 1$ Peptides), the expression of which decreased at 31°C when compared to 37°C after 24hrs in temperature shifted culture, identified during quantitative label free LC/MS analysis of conditioned media from a non-producing hybridoma cell line.	8
<b>Table 3.1.7:</b> List of host cell proteins ( $\geq 1$ Peptides), the expression of which decreased at 31°C when compared to 37°C after 96hrs in temperature shifted culture, identified during quantitative label free LC/MS analysis of cell culture supernatant from a non-producing hybridoma cell line.	9
<b>Table 3.1.8:</b> List of host cell proteins ( $\geq 1$ Peptides), the expression of which increased at 31°C when compared to 37°C after 96hrs in temperature shifted culture, identified during quantitative label free LC/MS analysis of cell culture supernatant from a non-producing hybridoma cell line.	10
<b>Table 3.1.9:</b> Table of proteins common to 24hrs vs 96hrs 37°C and 24hrs vs 96hrs 31°C differentially expressed protein lists from non-producing hybridoma cell line.	11
<b>Table 3.1.10:</b> HCPs from Hybridoma cell culture supernatant identified with one or more peptides as being common to 37°C and 31°C cultures. Proteins identified with two or more peptides as being common to 37°C and 31°C cultures are also indicated.	13
<b>Table 3.1.11:</b> HCPs from Hybridoma cell culture supernatant identified with one or more peptides as being unique to the 37°C culture.	24
<b>Table 3.1.12:</b> HCPs from Hybridoma cell culture supernatant identified with one or more peptides as being unique to the 31°C culture.	27

<b>Table 3.1.13:</b> HCPs from Hybridoma cell culture supernatant identified with two or more peptides as being unique to the 37°C culture.	29
<b>Table 3.1.14:</b> HCPs from Hybridoma cell culture supernatant identified with two or more peptides as being unique to the 31°C culture.	31
<b>Appendix B</b>	<b>32</b>
<b>Table 3.2.1:</b> Number of differentially expressed proteins at 24hrs compared to 120hrs at 37°C in a CHO K1 non-producing cell culture supernatant.	33
<b>Table 3.2.2:</b> List of host cell proteins, the expression of which increased after 120hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from a CHO K1 non-producing cell line.	34
<b>Table 3.2.3:</b> List of host cell proteins, the expression of which decreased after 120hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from a CHO K1 non-producing cell line.	43
<b>Table 3.2.4:</b> Number of differentially expressed proteins at 24hrs compared to 120hrs at 31°C in a CHO K1 non-producing cell culture supernatant.	45
<b>Table 3.2.5:</b> List of host cell proteins, the expression of which increased after 120hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from a CHO K1 non-producing cell line.	46
<b>Table 3.2.6:</b> List of host cell proteins, the expression of which decreased after 120hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from a CHO K1 non-producing cell line.	49
<b>Table 3.2.7:</b> Number of differentially expressed proteins between 31°C and 37°C at 24hrs post temperature shift in a CHO K1 non-producing cell culture supernatant.	51
<b>Table 3.2.8:</b> List of host cell proteins, the expression of which decreased at 31°C when compared to 37°C, 24hrs post temperature shift, identified during quantitative label free LC/MS analysis of cell culture supernatant from a CHO K1 non-producing cell line.	52
<b>Table 3.2.9:</b> List of host cell proteins, the expression of which increased at 31°C when compared to 37°C, 24hrs post temperature shift, identified during quantitative label free LC/MS analysis of cell culture supernatant from a CHO K1 non-producing cell line.	53
<b>Table 3.2.10:</b> Number of differentially expressed proteins between 31°C and 37°C at 120hrs post temperature shift in a CHO K1 non-producing cell culture supernatant.	59
<b>Table 3.2.11:</b> List of host cell proteins, the expression of which decreased at 31°C when compared to 37°C, 120hrs post temperature shift, identified during quantitative label free LC/MS analysis of cell culture supernatant from a CHO K1 non-producing cell line.	60

<b>Table 3.2.12:</b> List of host cell proteins, the expression of which increased at 31°C when compared to 37°C, 120hrs post temperature shift, identified during quantitative label free LC/MS analysis of cell culture supernatant from a CHO K1 non-producing cell line.	65
<b>Table 3.2.13:</b> Pearson's correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 37°C culture. The list below displays those proteins that were found to correlate negatively (decrease in expression) with time in a CHO K1 culture.	68
<b>Table 3.2.14:</b> Pearson's correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 37°C culture. The list below displays those proteins that were found to correlate positively (increase in expression) with time in a CHO K1 culture.	69
<b>Table 3.2.15:</b> Pearson's correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 31°C culture. The list below displays those proteins that were found to correlate positively (increase expression) with time in a CHO K1 culture.	72
<b>Table 3.2.16:</b> Table of proteins common to 24hrs vs 120hrs 37°C and 24hrs vs 120hrs 31°C differentially expressed protein lists from non-producing CHO K1 cell line. This table also includes information from 37°C vs 31°C at 120hrs so as to provide an indication of which culture condition the relative abundance of the given HCP is greatest.	73
<b>Table 3.2.17:</b> Table of proteins common to Pearson correlation analysis of protein expression over time at 24, 72, 96 and 120 hours at 37°C and 24, 72, 96 and 120 hours at 31°C in a non-producing CHO K1 cell line.	75
<b>Table 3.2.18:</b> The use of a correction factor was investigated to facilitate normalisation of the data by cell number (as opposed to normalisation by protein concentration). This table show a comparison of the number of differentially-expressed HCPs identified in the conditioned media of a CHO-K1 cell line with and without the correction factor applied.	76
<b>Table 3.2.19:</b> List of tables containing quantitative information on proteins classified as secreted by the Uniprot protein database.	77
<b>Table 3.2.20:</b> GO cellular component enrichment for differentially expressed host cell proteins identified in the conditioned media of a CHO K1 non-producing cell line at 37°C when 24hrs was compared to 120hrs (post temperature shift for a culture grown in parallel) analysed using quantitative label free LC/MS. An adjusted p-value of $\leq 0.05$ (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.	78
<b>Table 3.2.21:</b> GO cellular component enrichment for differentially expressed host cell proteins identified in the conditioned media of a CHO K1 non-producing cell line at 31°C when 24hrs was compared to 120hrs (post temperature shift) analysed using quantitative label free LC/MS. An adjusted p-value of $\leq 0.05$ (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.	79

<b>Table 3.2.22:</b> List of host cell proteins, classified as secreted by uniprot protein database, the expression of which increased at 120hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of conditioned media from a CHO K1 cell line.	80
<b>Table 3.2.23:</b> List of host cell proteins, classified as secreted by uniprot protein database, the expression of which decreased at 120hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of conditioned media from a CHO K1 cell line.	82
<b>Table 3.2.24:</b> List of host cell proteins, classified as secreted by uniprot protein database, the expression of which increased at 120hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of conditioned media from a CHO K1 cell line.	83
<b>Table 3.2.25:</b> List of host cell proteins, classified as secreted by uniprot protein database, the expression of which decreased at 120hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of conditioned media from a CHO K1 cell line.	84
<b>Table 3.2.26:</b> HCPs from non-producing CHO K1 cell culture supernatant identified with one or more peptides as being common to 37°C and 31°C cultures. Proteins identified with two or more peptides as being common to 37°C and 31°C cultures are also indicated.	85
<b>Table 3.2.27:</b> HCPs from non-producing CHO K1 cell culture supernatant identified with one or more peptides as being unique to the 37°C culture.	102
<b>Table 3.2.28:</b> HCPs from non-producing CHO K1 cell culture supernatant identified with one or more peptides as being unique to the 31°C culture.	106
<b>Table 3.2.29:</b> HCPs from non-producing CHO K1 cell culture supernatant identified with two or more peptides as being unique to the 37°C culture.	111
<b>Table 3.2.30:</b> HCPs from non-producing CHO K1 cell culture supernatant identified with two or more peptides as being unique to the 31°C culture.	115

**Appendix C** **118**

<b>Table 3.3.1:</b> Top 10 differentially regulated proteins, based on peptide number, when all features of the protein profile from 24hrs in a DP12 culture was compared to 144hrs and normalised to an internal standard, Cytochrome C. These proteins were then identified in the same comparison but where all features were normalised to proteins that did not change in expression. This allowed for direct comparison to be made between Cytochrome C Normalisation (C) and the Normal means of Normalisation (N), based on peptide count, ion score, ANOVA, fold change and what condition the proteins is most highly expressed in.	119
---	-----



<b>Table 3.3.2:</b> Top 10 differentially regulated proteins, based on peptide number, when all features of the protein profile from 24hrs in a DP12 culture was compared to 144hrs and normalised to proteins that did not change in expression. These proteins were then identified in the same comparison but where all features were normalised to an internal standard, Cytochrome C. This allowed for direct comparison to be made between the Normal means of Normalisation (N) and Cytochrome C Normalisation (C), based on peptide count, ion score, ANOVA, fold change and what condition the proteins is most highly expressed in.	120
<b>Table 3.3.3:</b> Number of differentially expressed proteins at 24hrs compared to 144hrs at 37°C in an IgG secreting DP12 cell culture supernatant.	121
<b>Table 3.3.4:</b> List of host cell proteins, the expression of which increased after 144hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from an IgG secreting DP12 cell line.	122
<b>Table 3.3.5:</b> List of host cell proteins, the expression of which decreased after 144hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from an IgG secreting DP12 cell line.	132
<b>Table 3.3.6:</b> Number of differentially expressed proteins at 24hrs compared to 144hrs at 31°C in an IgG secreting DP12 cell culture supernatant.	134
<b>Table 3.3.7:</b> List of host cell proteins, the expression of which increased after 144hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from an IgG secreting DP12 cell line.	135
<b>Table 3.3.8:</b> List of host cell proteins, the expression of which decreased after 144hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of cell culture supernatant from an IgG secreting DP12 cell line.	146
<b>Table 3.3.9</b> Number of differentially expressed proteins between 31°C and 37°C at 24hrs post temperature shift in the culture supernatant of an IgG producing DP12 cell line.	147
<b>Table 3.3.10:</b> List of host cell proteins, the expression of which decreased at 31°C when compared to 37°C, 24hrs post temperature shift, identified during quantitative label free LC/MS analysis of cell culture supernatant from an IgG secreting DP12 cell line.	148
<b>Table 3.3.11:</b> List of host cell proteins, the expression of which increased at 31°C when compared to 37°C, 24hrs post temperature shift, identified during quantitative label free LC/MS analysis of cell culture supernatant from an IgG secreting DP12 cell line.	152
<b>Table 3.3.12:</b> Number of differentially expressed proteins between 31°C and 37°C at 144hrs post temperature shift in the culture supernatant of an IgG producing DP12 cell line.	154
<b>Table 3.3.13:</b> List of host cell proteins, the expression of which decreased at 31°C when compared to 37°C, 144hrs post temperature shift, identified during quantitative label free LC/MS analysis of cell culture supernatant from an IgG secreting DP12 cell line.	155

<b>Table 3.3.14:</b> List of host cell proteins, the expression of which increased at 31°C when compared to 37°C, 144hrs post temperature shift, identified during quantitative label free LC/MS analysis of cell culture supernatant from an IgG secreting DP12 cell line.	158
<b>Table 3.3.15:</b> Pearson’s correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 37°C culture. The list below displays host cell proteins that were found to correlate negatively (decrease in expression) with time in a DP12 culture.	162
<b>Table 3.3.16:</b> Pearson’s correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 37°C culture. . The list below displays host cell proteins that were found to correlate positively (increase in expression) with time in a DP12 culture.	165
<b>Table 3.3.17:</b> Pearson’s correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 31°C culture. The list below displays host cell proteins that were found to correlate positively (increase in expression) with time in a DP12 culture.	170
<b>Table 3.3.18:</b> Pearson’s correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 31°C culture. The list below displays host cell proteins that were found to correlate negatively (decrease in expression) with time in a DP12 culture.	183
<b>Table 3.3.19:</b> Table of proteins common to 24hrs vs 120hrs 37°C and 24hrs vs 120hrs 31°C differentially expressed protein lists from a DP12 cell line.	184
<b>Table 3.3.20:</b> Table of proteins common to 24hrs vs 144hrs 37°C and 24hrs vs 144hrs 31°C differentially expressed protein lists from a DP12 cell line. This table also includes information from 37°C vs 31°C at 144hrs so as to provide an indication of which culture condition the relative abundance of the given HCP is greatest.	188
<b>Table 3.3.21:</b> Table of proteins common to Pearson correlation analysis of protein expression over time at 24, 72, 96,120 and 144 hours at 37°C and 24, 72, 96,120 and 144 hours at 31°C in a DP12 cell line.	191
<b>Table 3.3.22:</b> Table of proteins common to 24, 72, 96,120 and 144 hours at 37°C and 24, 72, 96,120 and 144 hours at 31°C Pearson correlation lists from a DP12 cell line. This table also includes information from 37°C vs 31°C at 144hrs so as to provide an indication of which culture condition the relative abundance of the given HCP is greatest.	196
<b>Table 3.3.23:</b> Table indicating the number of differentially expressed HCPs identified in a CHO K1 and DP12 cell line compared to the number of differentially expressed HCPs common to both cell lines.	199
<b>Table 3.3.24:</b> Table of differentially expressed proteins common to a CHO K1 and DP12 cell line when 24hrs is compared to 120hrs (CHO K1) / 144hrs (DP12) at 37C.	200
<b>Table 3.3.25:</b> Table of differentially expressed proteins common to a CHO K1 and DP12 cell line when 24hrs is compared to 120hrs (CHO K1) / 144hrs (DP12) at 31C.	211

<b>Table 3.3.26:</b> Table of differentially expressed proteins common to a CHO K1 and DP12 cell line when 37C is compared to 31C at 24hrs.	215
<b>Table 3.3.27:</b> Table of differentially expressed proteins common to a CHO K1 and DP12 cell line when 37C is compared to 31C at 120hrs (CHO K1) / 144hrs (DP12).	217
<b>Table 3.3.28:</b> The use of a correction factor was investigated to facilitate normalisation of the data by cell number (as opposed to normalisation by protein concentration). This table show a comparison of the number of differentially-expressed HCPs identified in the conditioned media of a DP12 cell line with and without the correction factor applied.	221
<b>Table 3.3.29:</b> List of tables containing quantitative information on proteins classified as secreted by the Uniprot protein database.	222
<b>Table 3.3.30:</b> GO cellular component enrichment for differentially expressed host cell proteins identified in the conditioned media of a DP12 CHO cell line at 37°C when 24hrs was compared to 144hrs (post temperature shift for a culture grown in parallel) analysed using quantitative label free LC/MS. An adjusted p-value of $\leq 0.05$ (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.	223
<b>Table 3.3.31:</b> GO cellular component enrichment for differentially expressed host cell proteins identified in the conditioned media of a DP12 cell line at 31°C when 24hrs was compared to 144hrs (post temperature shift) analysed using quantitative label free LC/MS. An adjusted p-value of $\leq 0.05$ (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.	224
<b>Table 3.3.32:</b> List of host cell proteins, classified as secreted by uniprot protein database, the expression of which increased at 144hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of conditioned media from a DP12 CHO cell line.	225
<b>Table 3.3.33:</b> List of host cell proteins, classified as secreted by uniprot protein database, the expression of which decreased at 144hrs when compared to 24hrs at 37°C, identified during quantitative label free LC/MS analysis of conditioned media from a DP12 CHO cell line.	227
<b>Table 3.3.34:</b> List of host cell proteins, classified as secreted by uniprot protein database, the expression of which increased at 144hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of conditioned media from a DP12 CHO cell line.	229
<b>Table 3.3.35:</b> List of host cell proteins, classified as secreted by uniprot protein database, the expression of which decreased at 144hrs when compared to 24hrs at 31°C, identified during quantitative label free LC/MS analysis of conditioned media from a DP12 CHO cell line.	231
<b>Table 3.3.36:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being common to 37°C and 31°C cultures. Proteins identified with two or more peptides as being common to 37°C and 31°C cultures are also indicated.	232

<b>Table 3.3.37:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being unique to the 37°C culture.	250
<b>Table 3.3.38:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being unique to the 31°C culture.	254
<b>Table 3.3.39:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with two or more peptides as being unique to the 37°C culture.	259
<b>Table 3.3.40:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with two or more peptides as being unique to the 31°C culture.	262
<b>Table 3.3.41:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 and an IgG secreting DP12 cell line identified with one or more peptides as being common to both cell lines. Proteins identified with two or more peptides as being common to both cell lines are also indicated.	265
<b>Table 3.3.42:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being unique to that cell line when compared to the HCP content of a non-producing CHO K1 cell line.	284
<b>Table 3.3.43:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with one or more peptides as being unique to that cell line when compared to the HCP content of an IgG secreting DP12 cell line.	290
<b>Table 3.3.44:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with two or more peptides as being unique to that cell line when compared to the HCP content of a non-producing CHO K1 cell line.	299
<b>Table 3.3.45:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with two or more peptides as being unique to that cell line when compared to the HCP content of an IgG secreting DP12 cell line.	304
<b>Table 3.3.46:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 and an IgG secreting DP12 cell line identified with one or more peptides as being common to both cell lines at 37°C. Proteins identified with two or more peptides as being common to both cell lines are also indicated.	309
<b>Table 3.3.47:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being unique to that cell line when compared to the HCP content of a non-producing CHO K1 cell line at 37°C.	324
<b>Table 3.3.48:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with one or more peptides as being unique to that cell line when compared to the HCP content of an IgG secreting DP12 cell line at 37°C.	329
<b>Table 3.3.49:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with two or more peptides as being unique to that cell line when compared to the HCP content of a non-producing CHO K1 cell line at 37°C.	335
<b>Table 3.3.50:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with two or more peptides as being unique to that cell line when compared to the HCP content of an IgG secreting DP12 cell line at 37°C.	339
<b>Table 3.3.51:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 and an IgG secreting DP12 cell line identified with one or more peptides as being common to both cell lines at 31°C. Proteins identified with two or more peptides as being common to both cell lines are also indicated.	343

<b>Table 3.3.52:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being unique to that cell line when compared to the HCP content of a non-producing CHO K1 cell line at 31°C.	358
<b>Table 3.3.53:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with one or more peptides as being unique to that cell line when compared to the HCP content of an IgG secreting DP12 cell line at 31°C.	364
<b>Table 3.3.54:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with two or more peptides as being unique to that cell line when compared to the HCP content of a non-producing CHO K1 cell line at 31°C.	371
<b>Table 3.3.55:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with two or more peptides as being unique to that cell line when compared to the HCP content of an IgG secreting DP12 cell line at 31°C.	376
<b>Table 3.3.56</b> Table displaying the number of proteins identified in the conditioned media of a CHO K1 cell line grown under temperature shift and non-temperature shift cultures with samples taken at 24, 72, 96 and 120 hours post temperature shift. RAW files from the mass spectral analysis were searched against the CHO NCBI and CHO BB protein databases. Columns are divided into those proteins identified by 1 peptide and proteins identified by 2 or more peptides.	380
<b>Table 3.3.57:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with one or more peptides as being common to both NCBI and CHO BB databases.	381
<b>Table 3.3.58:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with one or more peptides as being unique to the NCBI database.	399
<b>Table 3.3.59:</b> HCPs from the cell culture supernatant of a non-producing CHO K1 cell line identified with one or more peptides as being unique to the CHO BB database.	410
<b>Table 3.3.60:</b> Table displaying the number of proteins identified in the conditioned media of a DP12 cell line grown under temperature shift and non-temperature shift cultures with samples taken at 24, 72, 96, 120 and 144 hours post temperature shift. RAW files from the mass spectral analysis were searched against the CHO NCBI and CHO BB protein databases. Columns are divided into those proteins identified by 1 peptide and proteins identified by 2 or more peptides.	415
<b>Table 3.3.61:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being common to both NCBI and CHO BB databases.	416
<b>Table 3.3.62:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being unique to the NCBI database.	431
<b>Table 3.3.63:</b> HCPs from the cell culture supernatant of an IgG secreting DP12 cell line identified with one or more peptides as being unique to the CHO BB database.	439

<b>Appendix D</b>	444
<b>Table 4.2.1:</b> Table indicating the total number of phosphopeptides identified using each phosphopeptide enrichment method following enrichment of CHO SEAP cells grown under temperature shifted and non-temperature shifted conditions	445
<b>Table 4.2.2:</b> List of phosphopeptides identified as unique to Titanium dioxide phosphopeptide enrichment of temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift following analysis by LC/MS.	446
<b>Table 4.2.3:</b> List of phosphopeptides identified as unique to Iron phosphopeptide enrichment of temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift following analysis by LC/MS.	456
<b>Table 4.2.4:</b> List of phosphopeptides identified as unique to Gallium phosphopeptide enrichment of temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift following analysis by LC/MS.	467
<b>Table 4.2.5:</b> List of phosphopeptides identified as common to Titanium dioxide, Iron and Gallium phosphopeptide enrichments of temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift following analysis by LC/MS.	469
<b>Table 4.2.6:</b> List of phosphopeptides identified as common to Titanium dioxide and Iron phosphopeptide enrichments of temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift following analysis by LC/MS.	471
<b>Table 4.2.7:</b> List of phosphopeptides identified as common to Titanium dioxide and Gallium phosphopeptide enrichments of temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift following analysis by LC/MS.	486
<b>Table 4.2.8:</b> List of phosphopeptides identified as common to Iron and Gallium phosphopeptide enrichments of temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift following analysis by LC/MS.	487
<b>Table 4.2.9:</b> List of phosphopeptides identified following LC/MS analysis of an Iron phosphopeptide enriched sample of a temperature shifted CHO SEAP whole cell lysate taken down 36 hours post temperature shift that was found to be unique to the CHO BB protein database.	488
<b>Table 4.2.10:</b> List of phosphopeptides identified following LC/MS analysis of an Iron phosphopeptide enriched sample of a temperature shifted CHO SEAP whole cell lysate taken down 36 hours post temperature shift that was found to be unique to the CHO NCBI protein database.	489
<b>Table 4.2.11:</b> List of phosphopeptides identified following LC/MS analysis of an Iron phosphopeptide enriched sample of a temperature shifted CHO SEAP whole cell lysate taken down 36 hours post temperature shift that was found to be unique to the Swiss-Prot protein database containing sequences for Human, Mouse Rat and CHO (downloaded January 2011).	491

<b>Table 4.2.12:</b> List of phosphopeptides identified following LC/MS analysis of an Iron phosphopeptide enriched sample of a temperature shifted CHO SEAP whole cell lysate taken down 36 hours post temperature shift that were found to be common to CHO BB, CHO NCBI and Swiss-Prot protein databases. (Note: The Swiss-Prot protein database contains sequences for Human, Mouse Rat and CHO, downloaded January 2011)	492
<b>Table 4.2.13:</b> List of phosphopeptides identified following LC/MS analysis of an Iron phosphopeptide enriched sample of a temperature shifted CHO SEAP whole cell lysate taken down 36 hours post temperature shift that were found to be common to CHO BB and CHO NCBI protein databases.	494
<b>Table 4.2.14:</b> List of phosphopeptides identified following LC/MS analysis of an Iron phosphopeptide enriched sample of a temperature shifted CHO SEAP whole cell lysate taken down 36 hours post temperature shift that were found to be common to CHO BB and Swiss-Prot protein databases. (Note: The Swiss-Prot protein database contains sequences for Human, Mouse Rat and CHO, downloaded January 2011)	496
<b>Table 4.2.15:</b> List of phosphopeptides identified following LC/MS analysis of an Iron phosphopeptide enriched sample of a temperature shifted CHO SEAP whole cell lysate taken down 36 hours post temperature shift that were found to be common to CHO NCBI and Swiss-Prot protein databases. (Note: The Swiss-Prot protein database contains sequences for Human, Mouse Rat and CHO, downloaded January 2011)	497
<b>Table 4.2.16</b> Symbols denoting the amino acids Serine, Tyrosine and Threonine.	499
<b>Table 4.2.17:</b> List of phosphopeptides identified following LC/MS analysis of Iron phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be common to CHO BB and CHO NCBI protein databases.	500
<b>Table 4.2.18:</b> List of phosphopeptides identified following LC/MS analysis of Iron phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be unique to the CHO NCBI database.	516
<b>Table 4.2.19:</b> List of phosphopeptides identified following LC/MS analysis of Iron phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be unique to the CHO BB database.	526
<b>Table 4.2.20:</b> List of phosphopeptides identified following LC/MS analysis of Titanium Dioxide phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be common to CHO BB and CHO NCBI protein databases.	528
<b>Table 4.2.21:</b> List of phosphopeptides identified following LC/MS analysis of Titanium Dioxide phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be unique to the CHO NCBI database.	543

<b>Table 4.2.22:</b> List of phosphopeptides identified following LC/MS analysis of Titanium Dioxide phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be unique to the CHO BB database.	552
<b>Table 4.2.23:</b> List of phosphopeptides identified following LC/MS analysis of Gallium phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be common to CHO BB and CHO NCBI protein databases.	554
<b>Table 4.2.24:</b> List of phosphopeptides identified following LC/MS analysis of Gallium phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be unique to the CHO NCBI database.	556
<b>Table 4.2.25:</b> List of phosphopeptides identified following LC/MS analysis of Gallium phosphopeptide enriched samples from temperature shifted and non-temperature shifted CHO SEAP whole cell lysates taken down 36 hours post temperature shift that were found to be unique to the CHO BB database.	558



## Abbreviations

aa	-	Amino acid
AMPK	-	AMPactivated protein kinase
ATCC	-	American Tissue Culture Centre
BSA	-	Bovine Serum Albumin
cDNA	-	Complementary DNA
CHO	-	Chinese Hamster Ovary
Da	-	Daltons
DHFR	-	Dihydrofolate reductase
DMEM	-	Dulbecco's Modified Eagle's Medium
DMSO	-	Dimethyl Sulphoxide
DNA	-	Deoxyribonucleic Acid
DTT	-	Dithiothreitol
EDTA	-	Ethylene Diamine Tetracetic Acid
ELISA	-	Enzyme Linked Immunosorbent Assay
ERK	-	Extracellular SignalRegulated Kinase
FA	-	Formic Acid
FCS	-	Fetal Calf Serum
FWHM	-	Full Width at Half Maximum
GAPDH	-	Glyceraldehyde6phosphate dehydrogenase
HPLC	-	High Pressure Liquid Chromatography
HRP	-	Horseradish Peroxidase
Ig	-	Immunoglobulin
IMS	-	Industrial Methylated Spirits
JNK	-	cJun Nterminal Protein Kinase
kDa	-	kilo Daltons
LC	-	Liquid Chromatography
MAP	-	MitogenActivated Protein
mRNA	-	Messenger RNA

miRNA	-	microRNA
MS	-	Mass Spectrometry
MTX	-	Methotrexate
MWCO	-	Molecular Weight Cut Off
MWM	-	Molecular Weight Marker
NCBI	-	National Centre for Biotechnology Information
PBS	-	Phosphate Buffered Saline
RNA	-	Ribonucleic Acid
rpm	-	Revolutions per minute
SDS	-	Sodium Dodecyl Sulphate
SEAP	-	Secreted Alkaline Phosphatase
SFM	-	SerumFree Medium
TBS	-	Tris Buffered Saline
TFA	-	Trifluoroacetic Acid
Tris	-	Tris(hydroxymethyl)aminomethane
UHP	-	Ultra High Purity

**Title:** Proteomic profiling of phosphoproteins and secreted proteins from mammalian cell lines in order to gain insights into factors affecting cellular growth and recombinant protein production.

**Name:** Martin Power

### **Abstract**

Chinese hamster ovary (CHO) cells are one of the most commonly used cell lines in the production of biopharmaceuticals. The reduction of culture temperature during the exponential phase of a culture is a strategy that is commonly employed in the bioprocessing industry to increase process yield. Lower culture temperature results in a marked reduction in cell growth, increased specific productivity and prolonged cell viability.

In order to understand the mechanisms involved at the post-transcriptional level in the cellular response to temperature-shift a phosphoproteomic analysis of SEAP secreting CHO cells was conducted. Using a liquid-chromatography mass spectrometry-based label-free approach in conjunction with Gallium, Iron and Titanium phosphopeptide enrichment strategies 1,307 unique phosphopeptides (1,480 phosphosites) were identified. Gene Ontology analysis revealed enrichment of pathways involved in protein synthesis and cell cycle progression. 92 phosphopeptides were determined as being significantly differentially-expressed 36hrs post temperature-shift, including translation initiation factors EIF5B, EIF4G3, Transcription activator BRG1 and the tumour suppressor protein, Protein NDRG1. Such proteins are potentially involved in controlling growth and recombinant protein production in CHO cells.

While the effects of temperature-shift can result in an increase in overall product yield, it can also have implications for the extracellular-milieu from which the product must be purified. The second aspect of this thesis identifies changes in the host cell protein (HCP) profile of an IgG-producing and non-producing cell line over time in culture under both temperature-shifted and non-temperature-shifted conditions. This includes the identification of proteases (Cathepsin B, Cathepsin D, MMP9) and glycosidases (beta-galactosidase and  $\alpha$ -N-acetylgalactosaminidase), that could negatively impact product quality. In addition, proteins that could enhance cell growth such as vascular endothelial growth factor isoforms A and C and Hepatoma-derived growth factor were also identified. It was also found that the cell line and culture conditions used both impacted on the HCP profile generated.

# **Introduction**

## **1.0 Introduction**

### **1.1 Chinese Hamster Ovary cells**

Chinese hamster ovary (CHO) cells are a mammalian epithelial cell line that have been genetically engineered to produce recombinant protein products for what is now a multi-billion dollar industry (Beausoleil, Labrie, Dubreuil 2002; Rader 2012). The Chinese Hamster itself was first used for typing pneumococci as a replacement for mice in 1919 (Jayapal et al. 2007). The first CHO cell was isolated in 1957 for use in the study of somatic mammalian cell genetics (Puck, Cieciura, Robinson 1958). The discovery of the relatively low number of chromosomes in Chinese hamsters ( $2n=22$ ) made them ideal for radiation and chemical mutagenesis experiments (Chu 2004; Deaven and Petersen 1973). In 1986, Tissue plasminogen activator (tPA, Activase®), became the first therapeutic product approved for production using CHO cells. This came just four years after Lilly's Humulin became the first recombinant protein therapeutic to obtain regulatory approval from the FDA (Biospectrum 2007).

Used in the production of seven of the top ten selling biologics in 2011 (Table 1.1.1), CHO cells are used to produce nearly 70% of the recombinant therapeutic products on today's market (Huggett and Lähteenmaki 2012; Jayapal et al. 2007; Noh, Sathyamurthy, Lee 2013). CHO cells have been the predominant cell line of choice by the biopharmaceutical industry for a number of reasons; 1) CHO cells can be adapted to grow in suspension which enables high densities of cells to be obtained enabling greater product yield; they have a relatively fast growth rate and are reasonably robust in nature (Jayapal et al. 2007; Kumar, Gammell, Clynes 2007; Rodrigues et al. 2010). 2) The adaptation of CHO cells to grow in serum free media is also a significant advantage given the potential risks associated with the use of serum in cell culture media for the growth of therapeutic protein producing mammalian cell lines (Even, Sandusky, Barnard 2006; Sinacore, Drapeau, Adamson 2000). 3) Having been used for the production of biologics for over 25 years, CHO cells have a well-established record as a safe host for the production of parenteral products (Chu and Robinson 2001).

4) CHO cells have the ability to produce complex protein structures with correct folding patterns and post-translation modifications such as glycosylation which enables proper protein function (Hossler, Khattak, Li 2009; Vergara et al. 2012). 5) In addition to being stable hosts for the expression of heterologous genes, gene amplification techniques such as dihydrofolate reductase (DHFR) or glutamine synthetase (GS) mediated systems may also be used to increase cellular productivity (Barnes and Dickson 2006; Noh, Sathyamurthy, Lee 2013). DHFR negative CHO cells are transfected with a gene for the DHFR enzyme. Through the addition of increasing concentrations of the chemotherapeutic agent methotrexate (MTX) copies of the DHFR gene are also increased. This results in an increase in the attached gene for the therapeutic protein of interest, thus increasing its production (Jayapal et al. 2007). The GS system operates in a similar manner; in this instance however, the GS inhibitor methionine sulfoximine (MSX) is used to suppress endogenous copies of the GS gene. Only cells containing high enough copy number of the GS gene, linked to the desired recombinant protein product, will be able to proliferate. The addition of 20-25 $\mu$ M of MSX to the culture media enables selection of cells containing the GS gene, while increasing the concentration to between 100-1000 $\mu$ M results in amplification of the gene coding for the product of interest (Agrawal and Bal 2012).

### **Top Selling Biologic Drugs**

<b>Name</b>	<b>Product Cell Lines</b>	<b>2011 Worldwide Sales (dollars in billions)</b>
Humira (adalimumab)	Chinese Hamster Ovary	7.932
Enbrel (etanercept)	Chinese Hamster Ovary	7.367
Rituxan (rituxan)	Chinese Hamster Ovary	6.772
Remicade (infliximab)	Mouse Myeloma (NSO)	6.751
Avastin (bevacizumab)	Chinese Hamster Ovary	5.968
Herceptin (trastuzumab)	Chinese Hamster Ovary	5.924
Neulasta (pegfilgrastim)	Escherichia Coli	3.952
Lucentis (ranibizumab)	Escherichia Coli	3.769
Avonex (interferon beta-1a)	Chinese Hamster Ovary	2.687
Rebif (interferon beta-1a)	Chinese Hamster Ovary	2.354

**Table 1.1.1** Table of the top 10 selling biologic drugs in 2011 of which 7 were produced using a CHO host cell line (Noh, Sathyamurthy, Lee 2013).

Product titre has improved significantly over the years, having gone from ~50mg/l in the mid-eighties to ~4.7g/l in 2004. Product titres of >10g/l are now within reach (Huang et al. 2010; Wurm 2004). The increase in product yield has largely come from improvements in media formulation and process design (De Jesus and Wurm 2011). Further improvements are required, however, if the cost of these therapeutic products are to be brought down to a sustainable level (Dietmair, Nielsen, Timmins 2011). Herceptin, used for the treatment of breast cancer, costs US\$60,000 per patient per year (Waltz 2005). Treatment of colon cancer using Avastin costs US\$50,000, and for rare diseases, costs can be in excess of US\$200,000 per patient per year (Herper 2010). While manufacturing can play some role in influencing the cost of the drug i.e., volume of product produced, expression system used (cheaper prokaryotic systems versus more expensive eukaryotic models), such high prices are largely attributed to ever increasing cost of R and D including costly clinical trials. Although cell engineering strategies have made some contribution towards improving cell line characteristics, results from the ‘targeted’ engineering approach have so far fallen somewhat short of expectations (Lim et al. 2010; O’Callaghan and James 2008). Advances in modern technology are currently enabling unprecedented characterisation of CHO cells in the fields of genomics, proteomics and metabolomics (Datta, Linhardt, Sharfstein 2013).

The recent publication of the CHO genome will undoubtedly play a significant role in the quest for true cell line engineering (Lewis et al. 2013; Xu et al. 2011). Through better understanding of the mechanisms involved in controlling different cellular traits we will be one step closer to finally developing the much sought after 'super producer' (Dietmair, Nielsen, Timmins 2011; Seth et al. 2007).



## **1.2 Temperature-shift**

The reduction of culture temperature towards the end of the exponential phase or during the stationary phase of a culture is a strategy that is commonly employed in the bioprocessing industry to increase process yield. Temperature reduction results in a marked reduction in cell growth, increased specific productivity and prolonged cell viability (Bollati-Fogolín et al. 2005; Rodriguez et al. 2005; Sakurai et al. 2005; Schatz et al. 2003). Despite the increases in specific productivity and cell viability, overall product yield may be decreased as a result of reduced growth. The effect of this temperature reduction can also vary greatly according to the temperature, recombinant protein, cell line and even clone (Yee et al. 2008; Yee, Gerdtzen, Hu 2009; Yoon, Hwang, Lee 2004). While the use of temperature-shift to increase the specific productivity of mammalian cells has been investigated (Fox et al. 2004; Yoon, Song, Lee 2003) and the effect of cold shock well documented (Al-Fageeh and Smales 2006; Al-Fageeh et al. 2006), the precise cellular mechanisms involved have only begun to become clear in more recent times. Despite this, there still remains a significant gap in our knowledge of the mechanics of reduced culture temperature, particularly when compared with that of other biological systems such as bacteria or plant. The following section will discuss our current understanding of the mechanisms involved in the response of CHO cells to reduced culture temperature with respect to (i) cell growth and (ii) specific productivity.

### **1.2.1 Mechanisms controlling Cell growth rates in temperature-shifted CHO cells**

As mentioned previously, cell growth is significantly curtailed under reduced culture temperature conditions. This reduction in growth can be attributed to the cell cycle arrest with a shift in the proportion of cells from the S to the G1 phase of the cell cycle observed at reduced temperatures (Hendrick et al. 2001; Marchant et al. 2008; Moore et al. 1997). It should be pointed out that, under mild hypothermic conditions (~32°C) cells may not arrest completely and can continue to proliferate, albeit at a reduced rate compared to 37°C, while under more severe conditions there is almost a complete cessation of cell proliferation (Roobol et al. 2009).

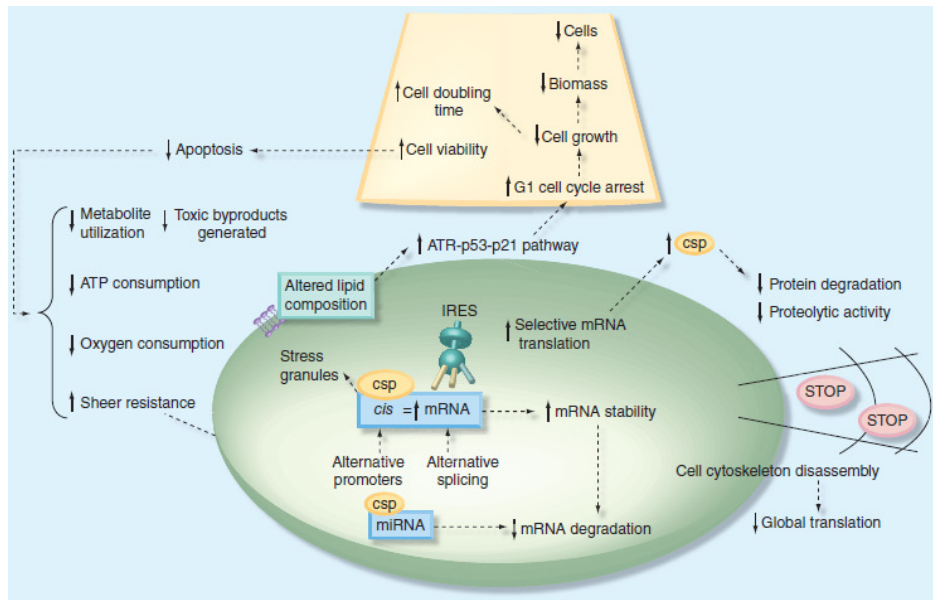
This reduced growth rate results in a reduced requirement for metabolites and in turn decreased production of waste by-products (Wagstaff et al. 2013). It has been shown that the cell cycle arrest of CHO cells occurs via the ATR (ataxia telangiectasia mutated-and Rad3-related kinase) - p53 – p21 pathway (Roobol et al. 2011). The finding that p53 is central to the arrest of temperature-shifted CHO cells is hardly surprising given that p53 deficient mammalian cells experiencing mild hypothermia do not undergo cell cycle arrest (Matijasevic, Snyder, Ludlum 1998; Ohnishi et al. 1998). It was proposed that activation of the ATR-p53–p21 pathway is due to changes in membrane rigidity as a result of alterations in the membrane lipid composition; specifically, an increase in the ratio of polyunsaturated to saturated fatty acids upon temperature-shift (Roobol et al. 2011). It has been suggested that a change in the fluidity of the nuclear envelope is ‘sensed’ by nuclear localised-ATR (Zhang, Zhao, Ma 2007). In CHO, this leads to the intranuclear translocation of ATR and subsequent phosphorylation of p53 at Ser15 (Roobol et al. 2011). The phosphorylation of Ser15 is enhanced by the phosphorylation of Ser33 and Ser46 on p53 by p38MAPK (Bulavin et al. 1999), also activated by the ATR pathway (Roobol et al. 2011). This enhanced Ser15 phosphorylation serves to stabilise and augment binding of p53 to the transcription co-activator p300/CBP (Dumaz and Meek 1999), thus enabling transcription of p21, resulting in cell cycle arrest.

### **1.2.2 Mechanisms controlling specific productivity in temperature-shifted CHO cells**

Although cell growth is significantly reduced under temperature-shifted compared to non-temperature-shifted conditions, extended culture viability coupled with increased specific productivity can lead to an increase in overall product yield. The effect of mild hypothermia on specific productivity can be cell line specific however and as such, may not be a suitable strategy for all cell lines (Table 1.2.1). Despite the recent elucidation of the mechanisms involved in cell cycle arrest of CHO cells under reduced culture temperature conditions (Roobol et al. 2009), our understanding of the pathways controlling specific productivity is a lot less clear. (A diagram summarising our understanding of the effect of temperature-shift on CHO cells is provided in Figure 1.2.1.). From the studies conducted so far it would appear that the factors facilitating the increases in specific productivity are more complex.

Product	T (°C)	Productivity (fold change)	Feed Strategy	Author
Antibody	33	1.7	Batch	(Yee, Gerdtzen, Hu 2009)
Antibody-IL-2 fusion	30	2.8	Batch	(Shi et al. 2005)
Chimeric Fab	28	14	Batch	(Schatz et al. 2003)
EPO	30	5.6	Batch	(Moore et al. 1997)
EPO	33	4	Batch	(Yoon, Song, Lee 2003)
EPO	32	4	Perfusion	(Ahn et al. 2008)
FSH	32	13	Perfusion	(Yoon, Ahn, Jeong 2007)
hGM-CSF	33	2.1	Batch	(Bollati-Fogolin et al. 2005)
IFN- $\gamma$	32	4	Fed-batch	(Fox et al. 2005)
IgG	30	2	Batch	(Kantardjieff et al. 2010)
Pro-urokinase	31	1.3	Batch	(Chen et al. 2004)
Pro-urokinase	34	1.5	Perfusion	(Chen et al. 2004)
SEAP	30	1.7	Batch	(Kaufmann et al. 1999)
SEAP	33	8	Fed-batch	(Nam, Ermonval, Sharfstein 2009)
TIMP	34	4	Batch	(Jenkins and Hovey 1993)
TNFR-Fc	30	5	Batch	(Kou et al. 2011)
tPA	33	1.5	Batch	(Berrios et al. 2009)
tPA	32	1.7	Batch	(Hendrick et al. 2001)
tPA	33	1.7	Perfusion	(Vergara et al. 2009)

**Table 1.2.1** Effect of temperature-shift on specific productivity on CHO cells.



**Figure 1.2.1** Diagram of the key responses to reduced culture temperature conditions seen in CHO cells (Diagram reproduced from Masterson et al. 2014)

Overall, there is a reduction in protein synthesis as a result of the attenuation of transcription and translation processes within the cell. This response appears to be common to both bacterial and mammalian systems experiencing mild hypothermia (Ermolenko and Makhatadze 2002).

While this may seem somewhat contradictory given the increase in specific productivity, it is generally thought that increased mRNA stability/decreased mRNA degradation coupled with the availability of translational and protein folding/secretory machinery, unburdened by task of synthesising endogenous cellular proteins, creates increased capacity for the production of more recombinant protein product. Again however, a consensus is yet to be reached on precisely how this occurs. This is primarily because of a variations in the results reported across different cell lines. For example, a 5.5-fold increase in the specific productivity of tumour necrosis factor receptor-immunoglobulin G1 Fc fusion (TNFR-Fc) in a temperature-shifted CHO cell line was attributed to enhanced mRNA stability, rather than the changes in the gene copy number or an increase in the transcription rate (Kou et al. 2011). It has also been suggested that an increase in EPO mRNA levels under temperature-shift conditions was indicative of an increase in EPO transcription therefore providing the means for an increase in EPO productivity (Yoon, Song, Lee 2003). It is thought that this increased stability is facilitated by the translation of a specific group of proteins termed “cold shock proteins”. Two of the most widely studied cold shock proteins are cold-inducible RNA binding protein (CIRP) and RNA-binding motif protein 3 (RBM3) (Al-Fageeh and Smales 2006). Although structurally very similar (Nishiyama et al. 1997), CIRP specifically binds to the 3′ untranslated region (UTR) of RNAs, improving RNA stability and translation efficiency at times of cell stress (Yang and Carrier 2001), while it is thought that RBM3 binds the 5′ UTR of mRNA to facilitating cap-independent translation during cold shock (Chappell, Owens, Mauro 2001). In addition to this it is also thought that the binding of such cold shock proteins (CSPs) to the cis-element of the promoter region of certain genes may also enable their transcription during temperature-shift (Al-Fageeh and Smales 2006).

To add another layer of complexity, the expression of small non-coding RNAs called microRNA (miR), which largely bind to the 3'-UTR of the target mRNA to negatively regulate their translation, have been shown to modulate under temperature-shift (Barron et al. 2011; Gammell et al. 2007), offering another means by which pathways involved in specific productivity could be effected.

At the post transcriptional level, the decreases in global protein synthesis has been attributed to the phosphorylation of the initiation factor eIF2 $\alpha$  (Masterton et al. 2010; Underhill et al. 2005). It has been shown that upon phosphorylation, eIF2 $\alpha$  forms an inhibitory complex with eIF2B, which in turn prevents the binding of initiator met-tRNA<sub>i</sub> to the 40S ribosomal subunit resulting in the suppression of global protein synthesis (Krishnamoorthy et al. 2001; Underhill et al. 2005). Despite this, it does stand to reason that other eukaryotic initiation translation factors and their associated pathways must in some way be affected by the decrease in culture temperature in order to facilitate an increase in specific productivity. This pathway has, as yet, to be identified.

Remodelling of the cytoskeleton has also been observed during temperature-shift of CHO cells (Kumar et al. 2008a; Yee, Gerdtzen, Hu 2009). As such, it is likely that this may have some impact on translation due to the co-localisation of initiation factors with the cytoskeleton (Stapulionis, Kolli, Deutscher 1997). Disruption to the cytoskeleton through the disassembly of microtubules can be impeded through the expression of stable-tubule-only-polypeptides (STOP) proteins to stabilize the microtubule network (Masterton and Smales 2014). The specific mechanisms by which this impacts on the translational machinery of the cell also has yet to be elucidated.

In summary, the means by which specific productivity is increased appears to be multifaceted; from the expression of CSPs, to the stabilisation of mRNA, the induction of miRNAs to modulate mRNA translation and the phosphorylation of specific eukaryotic initiation translation factors to facilitate translation of specific proteins under reduced culture temperature conditions.

### **1.2.3 Proteomic profiling of temperature-shifted CHO cells**

While targeted approaches have been used to investigate specific pathways / known proteins that may be implicated in the cellular response to temperature-shift such as those described in sections 1.2.1 and 1.2.2, proteomic profiling has been used to identify some of the more global changes within the proteome of temperature-shifted cells. This section briefly discusses these reports and their key findings.

It should be noted that a detailed description of the proteomic techniques mentioned in this section is provided in section 1.6.

One of the earliest profiling studies utilized a two dimensional polyacrylamide gel electrophoresis (2D PAGE) approach to examine extracts from cells grown at 37°C and 31°C (Kaufmann et al. 1999). The gels were stained using a sensitive stain for the detection of proteins called silver stain and the stained protein spot intensity compared between temperature-shifted and non-temperature-shifted samples. The identification of 10 spots with altered intensities provided a proof of concept that temperature-downshift results in the altered expression of proteins in CHO cells. In addition to this, the identification of alterations in the levels of tyrosine phosphorylation on a 80 kDa and 180 kDa protein between 37°C and 31°C, provided the first evidence of CHO cells actively responding at the posttranslational level to changes in culture temperature (Kaufmann et al. 1999).

Also employing silver stained/2D PAGE approach, the proteomic profiling element of a combined transcriptome and proteome analysis of EPO producing CHO cells identified 13 spots as having at least a twofold change in spot intensity upon temperature-shift (Baik et al. 2006). Although spot identification by MALDI-ToF and MS/MS revealed only nine proteins to be differentially-expressed, the increase in a number of proteins (PDI, ERp57, HSP70) implicated changes in the protein folding machinery as attributing factor towards increased EPO productivity.

In a more sophisticated transcriptomic and proteomic investigation, an expansion of the cell's secretory machinery was also identified as a key component of the increase in specific productivity of an IgG producing CHO cell line following temperature-shift and sodium butyrate treatment (Kantardjieff et al. 2010).

Using a metabolic labelling (stable isotope labelling with amino acids in cell culture (SILAC)) approach to study the secretion rate of IgG CHO cells grown (i) at 33°C and (ii) at treated with sodium butyrate (NaBu) at 33°C, it was found that rather than faster processing times, an increase in the availability of the light chain molecule facilitated more efficient production of the IgG product. This was backed up by transcriptome analysis which identified functional classes related to protein secretion to be enriched under temperature-shift and NaBu treatments.

2D difference gel electrophoresis (DIGE) analysis of a non-producing CHO-K1 cell line under temperature-shift revealed that proteins predominantly involved in cell structure and metabolism altered in expression (Kumar et al. 2008a). Given that this study was carried out in a non-producing cell line, few insights into the mechanisms involved in the increase in specific productivity seen in many temperature-shifted cell lines could be obtained. Nonetheless, among the 201 differentially-expressed proteins identified (118 up-regulated at 31°C), a number of those likely to be involved in important cellular pathways were also observed including; Heterogeneous nuclear ribonucleoprotein C (HNRPC) (regulation of growth), EIF4A (cap-independent translation), importin- $\alpha$  (apoptosis), Vimentin (structural) and alpha glucosidase 2 (glycoprotein quality control) (Kumar et al. 2008a).

### **1.3 Proteomic profiling of desirable cell phenotypes**

A wide range of proteomic profiling studies have been carried out on CHO cells in an effort to decipher the mechanics of the recombinant protein producing cell. By profiling cells displaying desirable phenotypic traits such as fast growth rate, high productivity, low lactate/ammoniac production, delayed onset of apoptosis etc, it is hoped that this information can one day be used to engineer a ‘super producer’ (Seth et al. 2007). While our knowledge of the CHO proteome has vastly improved over the last number of years (Baycin-Hizal et al. 2012), a deeper understanding is required in order to select cell engineering targets with a view to improving cellular phenotype in a more efficient manner (Dietmair, Nielsen, Timmins 2011). This section provides a brief overview of some of the ‘desirable’ characteristics that could be engineered into a cell in Table 1.3.1 and summarises some of the proteomic profiling investigations conducted to date in Table 1.3.2

Cellular trait	Overview 'desirable' cellular characteristics
Cell growth	The link between high cell density and final product yield is well established and, in some cases, has been shown to be more important than the intrinsic productivity rate of the cell line (Clarke et al. 2011). Rapidly achieving such a high cell density greatly reduces (1) the time spent scaling up the cell culture process through serial culture passages of the seed train and inoculum train (2) the time taken to obtain a suitable cell density before a growth arrest strategy can be employed to enhance specific productivity (Doolan et al. 2010; Sunley and Butler 2010), increasing the specific growth rate of a cell would therefore greatly benefit the bioprocessing community both in terms of increasing overall product yield and reducing lead times.
Specific productivity	This is perhaps one of the most obvious characteristics to engineer into a cell line. Profiling of different phenotypes has identified various bottlenecks in the secretory machinery, such as protein folding. The overexpression of endoplasmic reticulum chaperones has had mixed results; for example, the over expression of ERp57, Calnexin and Calreticulin produced a two-fold increase in the production of Thrombopoietin (TPO) (Chung et al. 2004; Hwang, Chung, Lee 2003), while the overexpression of PDI did not result in an increase in the specific productivity of TPO, although increases in IgG production were seen (Borth et al. 2005; Mohan et al. 2007). By increasing our understanding of the components required to make a fully functioning recombinant molecule, and how they operate, cell engineering strategies will become more efficient in their approach to making a high producing cell line.
Cell metabolism	Because of metabolic inefficiencies, the high consumption rate of nutrients such as glucose and glutamine can result in the accumulation of lactate and ammonia in a bioreactor, which can be detrimental to the health of the culture (Altamirano et al. 2013). A number of engineering strategies have proven successful in generating cell lines that can circumvent this problem, including the expression of carbamoylphosphate synthetase I and ornithine transcarbamoylase enabling CHO cells to convert $\text{NH}_4$ to citrulline, thereby reducing $\text{NH}_4$ concentration in the cultures up to 1.3-fold (Park et al. 2000). By examining phenomena such as 'metabolic shift' whereby cells have significantly reduced nutrient consumption / waste by-product production compared to cells that have not undergone this shift (Korke et al. 2004), insights may be gained into approaches that would enable cells to metabolise nutrients more efficiently, reducing waste by-product production.
Apoptosis	Extending culture viability by delaying cell apoptosis would be a very desirable characteristic to engineer into a cell; by prolonging the lifetime of the culture product yields can be increased (Krampe and Al-Rubeai 2010). Although there have been a number of publications reporting the use of the overexpression of anti-apoptotic proteins such as Bcl-2 (Kim and Lee 2000), Bcl-xL (Chiang and Sisk 2005), and Mcl-1 (Majors et al. 2009), or the down-regulation of pro-apoptotic proteins such as Bak and Bax (Cost et al. 2010), proteomics studies are still on-going in an attempt to identify alternate means by which apoptosis can be deferred (Wei et al. 2011).

**Table 1.3.1** Overview of some 'desirable' characteristics in a recombinant protein producing cell line.



**Table 1.3.2** Table summarising some of the proteomic profiling investigations conducted date in CHO cells that have furthered our understanding of some of the cellular mechanisms involved in important cell attributes such as growth rate, productivity, apoptosis and metabolic shift.

Cell line	Cellular Trait	Proteomic Technique	Comparison	Product	Key Findings	Reference
CHO	Apoptosis	2D DIGE	Analysis of CHO cells undergoing apoptosis	IgG	Although no direct apoptosis-regulating proteins were identified, the potential importance of UPR and energy metabolism in prolonged cultures and apoptotic cell death was established. In addition to the type of apoptotic pathway triggered, apoptotic proteins associated with ER stress, mitochondrial dysfunction and energy metabolism were identified.	(Wei et al. 2011)
NS0	Productivity	2D PAGE (Ruby stained)	Analysis of four NS0 cell lines with different specific productivities	IgG	Proteins exhibiting a significant increase in abundance with increasing qMab included molecular chaperones known to interact directly with nascent immunoglobulins during their folding and assembly (e.g., BiP, endoplasmin, protein disulfide isomerase).	(Smales et al. 2004)
NS0	Productivity	2D PAGE (Ruby stained)	Analysis of four NS0 cell lines with different specific productivities if particular cell lines have distinct functional capabilities	IgG	Significant increase in the abundance of proteins in the ER chaperone, non-ER chaperone and cytoskeletal categories, while proteins categorised as being involved in protein synthesis, protein degradation and nucleic acid synthesis remained unchanged.	(Dinnis et al. 2006)

CHO	Productivity	2D DIGE	Analysis of two CHO cell line pairs, with each pairing differing in their ability to sustain high productivity over a ten day culture period	IgG	Overlap comparisons between the two sets of cell line pairs identified 12 proteins, involved in processes such as protein folding and protein translation were differentially expressed in the same direction.	(Meleady et al. 2011)
CHO	Growth rate	QLF-LC/MS	miRNA, mRNA and protein expression analysis of 'fast' and 'slow' growing clones	IgG	Biological processes driving proliferation including mRNA processing and translation, in addition, the upregulation of numerous ribosomal proteins (RP) was observed (e.g. RPL14, RPS15 and RPL15).	(Clarke et al. 2012)
CHO	Growth rate	2D DIGE	Genomic and proteomic analysis of 'fast' and 'slow' growing clones	IgG	Genomic/proteomic analysis identified 21 high-priority cell engineering targets. Functional validation found of five selected targets found the knockdown of HSPB1, ENO1 and VCP all impacted on cell growth.	(Doolan et al. 2010)
CHO	Metabolic shift	2D PAGE (Silver stained)	Lactate consuming compared to non-lactate consuming cell line	IgG	Proteins involved in glycolysis, protein processing and cell structure were differentially expressed. How low lactate conditions are generated e.g. through different feed strategies, can impact on the proteome of the cell.	(Pascoe et al. 2007)
CHO	Productivity	iTRAQ (LC/MS)	Transcriptomic and proteomic analysis of high and low producing CHO cells	Fusion protein	Differential expression of genes and proteins include carbohydrate metabolism, signal transduction, and transport were observed. Upregulation of proteins involved in protein metabolism and protein folding in the high producer were also identified.	(Nissom et al. 2006)

## **1.4 Host cell proteins (HCPs)**

Host cell proteins (HCPs) are process related impurities derived from a host cell line that must be removed from the final drug product. As well as impacting on product quality during processing, insufficient removal of these impurities may result in an immunogenic response in the patient (Gao et al. 2011; Shankar, Pendley, Stein 2007). Although no limit has been set by the Regulatory Authorities it is recommended that levels of HCPs in the final drug product are “below detectable levels using a highly sensitive analytical method” (CBER 1997). It has been acknowledged that testing for such impurities is very challenging since the test must be capable of detecting fragments of HCPs in a complex sample that may cover a large dynamic range (Rathore et al. 2003). One popular means of testing for the presence of HCPs is the use of enzyme linked immunosorbent assays (ELISAs) as they are capable of detection between 1-100 ppm (Eaton 1995; Hoffman 2000). Such assays are generated by raising antibodies against a protein mixture from the host cell line. To avoid false positives it is important to ensure that the assay does not contain antibodies that have an affinity for the product of interest. For this reason the protein mixture is generated from a non-producing cell line. It could be argued that proteins that are weakly / non-immunoreactive will not be detected by ELISA. The use of an alternative detection method such as silver staining on a 1D or 2D gel may be useful in such instances (Champion et al. 2005).

ELISAs used to test for the presence of HCPs may be categorised as generic, custom or multiproduct. The type of ELISA used will depend on a number of factors such as product maturity and general suitability. Commercially available generic assays are typically employed during the early stages of product development i.e. phase 1 clinical trials, because they are inexpensive and capable of measuring HCPs present in a given cell line (Schwertner and Kirchner 2010). However, results from generic assays have been shown to be variable and sometimes lack sufficient sensitivity (Schwertner and Kirchner 2010), so once a product reaches phase 3 of clinical trials a more specific i.e. custom assay, may be required by the regulatory authorities.

Usually taking 9-12 months, custom assay development is a time consuming endeavour (Wolter and Richter 2005). It is also costly and may be deemed unsuitable if changes to the final purification process are made at a later date (Champion et al. 2005). A compromise between these two types of assay may be found in the multiproduct assay (Schwertner and Kirchner 2010). This assay is developed by the company to be used across a number of products which are derived from a specific cell line.

Krawitz et al. (2006) supported the use of multi-product immunoassays for HCP analysis in recombinant protein production based on a comparison of the proteomes of three different CHO cell lines. Using 2D PAGE stained using a sensitive staining method (SPYRO Ruby), this study found no difference in the basal expression of proteins in the three cell lines (Krawitz et al. 2006). This is not surprising given that other studies have found the proteome of therapeutic producing cell lines to be very similar (Champion et al. 1999; Hayduk, Choe, Lee 2004). It can be argued that these studies only consider the intracellular proteome and do not take into account subtle differences that may exist in the extracellular HCPs (Jin et al. 2010), such as an increase in the abundance of acidic proteins in the extracellular proteome (Wimmer et al. 1994). Nonetheless the intracellular proteome is regarded as being sufficiently similar to the extracellular proteome to make such studies valid (Hoffman 2000).

In two recent papers, 2D DIGE was used to investigate the HCP profile of IgG producing CHO cell lines (Grzeskowiak et al. 2009; Jin et al. 2010). Interestingly, the HCP profiles of cells exposed to different culture conditions e.g. inoculation density, basal and feed media, temperature-shift strategy, were found to be very similar (Jin et al. 2010). The most notable changes came as a result in cell viability where HCP content was greatest in low viability cultures (Grzeskowiak et al. 2009), while there was minimal change to the HCP profile of cells maintained at high viability where (Jin et al. 2010). As a result of this finding, cell culture viability must now be a consideration for the generation and use of ELISAs to detect HCPs during routine analytic testing.

#### **1.4.1 Proteases as part of the HCP matrix**

Proteases are an integral part of the intracellular proteome. They play an important role in cellular processes such as DNA replication, cell-cycle progression and cell proliferation (López-Otín and Overall 2002). As a result, the presence of proteases in the cell culture broth is inevitable be it as a result of secretion from viable cells (Elliott, Hohmann, Spanos 2003; Robert et al. 2009; Sandberg et al. 2006; Satoh, Hosoi, Sato 1990), dead cells (Satoh, Hosoi, Sato 1990; Teige, Weidemann, Kretzmer 1994) or the mechanical lysis of cells during culture or harvest (Kao et al. 2010; Trexler-Schmidt et al. 2010). The addition of excipients such as media supplements is also a potential source of proteases (Mols et al. 2004). Proteases can continue to be a problem right throughout the purification process. Harvest operations may cause excessive cell shear, causing intracellular proteases to be released into the cell culture fluid (Kao et al. 2010; Trexler-Schmidt et al. 2010).

Concentration of the protein product in the early stages of purification may also lead to a significant increase in protease activity (Sandberg et al. 2006). Changes in the pH of buffers used in the purification process may also lead to the activation of proteases such as acidic or aspartic proteases, such as those identified in the conditioned media of mAb producing cell lines in other studies (Karl, Donovan, Flickinger 1990; Sandberg et al. 2006). HCP proteases can also cleave Protein A ligand (Shukla et al. 2007) having the effect of both reducing column performance and adding an additional impurity that must also be removed.

Protease activity in the cell culture can have a detrimental effect on product quality (Kao et al. 2010). Degradation of the product by protease activity has been reported in some of the most commercially relevant, eukaryotic expression systems, CHO (Elliott, Hohmann, Spanos 2003; Robert et al. 2009; Sandberg et al. 2006) and Hybridoma cell lines (Karl, Donovan, Flickinger 1990).

Results from these studies seem to indicate that product degradation as a result of protease activity is somewhat cell line and product specific. There may be a number of reasons for this; (i) some level of protection may be provided by naturally occurring protease inhibitors produced by the cell line (ii) the protease may be produced in an inactive form, or (iii) the product may not contain amino acids sequences that are sensitive to cleavage by the proteases present in the media (Mols

et al. 2004; Trexler, Bányai, Patthy 2001). Reasons for inhibiting protease activity are two-fold: a decrease in product degradation would not only have the effect of increasing process productivity but a reduction in in the volume of degraded antibody molecules would mean a reduction in the size of the Protein A column required (Reid et al. 2010).

By and large protease inhibition studies have been the mainstay of protease identification and product conservation (Elliott, Hohmann, Spanos 2003; Lindskog, Svensson, Häggström 2006; Sandberg et al. 2006). On an industrial scale the addition of protease inhibitors can be costly and their removal from the final drug substance must also be considered. Up to 2011 the lack of a comprehensive CHO proteome database has hampered efforts to conclusively identify some proteases (Sandberg et al. 2006) and implement potential cell engineering solutions (Robert et al. 2009). However, following the release of two CHO proteomic databases (Meleady et al. 2012; Xu et al. 2011) one would expect our knowledge of such proteases to increase as the host cell proteome comes under greater scrutiny.

Media formulation could also benefit from insights into protease activity since the use of serum free media, as required by the regulatory authorities, may see an increase in protease activity due to the absence of protease inhibitors that would be present in serum (Teige, Weidemann, Kretzmer 1994).

None the less, even our limited understanding of these proteases can contribute toward making key strategic decisions such as when to harvest the product. How long the product spends in the cell culture broth prior to harvest has been shown to impact on product quality both in terms of the presence/absence of crucial PTMs and structural integrity (Hansen et al. 1997; Reid et al. 2010; Teige, Weidemann, Kretzmer 1994).

Protease activity may also affect cell line choice. It has been suggested that a carboxypeptidase B or another basic carboxypeptidase is at least partially responsible for a C-terminal lysine variation in Mabs produced in CHO and B cell hybridoma cells (Dick Jr et al. 2008). Although it had been hypothesized that a reduction in culture temperature could reduce protease activity while at the same time increasing overall productivity (Chuppa et al. 1997; Kaufmann et al. 1999), one study found

that product quality was similar for both conditions despite an increase in a 69kDa protease in the 37°C culture (Clark, Chaplin, Harcum 2004).

In terms of product quality much of the emphasis has been placed on the ability of the upstream process to produce a consistent recombinant protein structure (Zhou et al. 2008). It is only in more recent times that the focus has shifted to the effect of downstream processing on product integrity (Gao et al. 2011; Reid et al. 2010). The importance of thorough purification of the protein product was highlighted by Gao et al (2010) who showed that the presence of residual protease activity in a highly purified monoclonal antibody sample can negatively impact product quality.

#### **1.4.2 Growth Factors and Conditioned Media**

Serum is rich source of nutrients that are known to provide cells with protection from apoptosis (Zanghi, Fussenegger, Bailey 1999) and support cell proliferation (Liew et al. 2010). Its precise constituents, however, remain somewhat unknown. Due to its animal origins the components of serum can vary in concentration thus causing unpredictable growth and productivity patterns (Even, Sandusky, Barnard 2006). There are also regulatory concerns that it may contain viruses making it unsuitable for use given the intended market of the therapeutic product (Even, Sandusky, Barnard 2006). Serum Free Media (SFM) is a defined media composition containing known growth factors and other elements that support cell growth and recombinant protein production (Schroder, Matischak, Friedl 2004). The addition of proteins to media can be expensive (particularly on an industrial scale) and are often obtained from an animal source (Keenan, Pearson, Clynes 2006). Chemically defined media has been the gold standard in terms of media formulation (Grillberger et al. 2009).

One area however that has received very little attention in bioprocessing is the effect of autocrine factors on cell proliferation and productivity. Conditioned media can be a rich source of such factors (Dowell, Johnson, Li 2009). Conditioned media from cancer cell lines is commonly used to identify biomarkers in cancer that correlate with characteristics of interest e.g. invasiveness, growth rate, etc. (Xue, Lu, Lai 2008)(Chenau et al. 2009; Makridakis et al. 2010). It is a source that does not appear to have been so thoroughly investigated for bioprocessing cell lines. Such information could potentially be used to screen for biomarkers of desirable

phenotypes (Woolley and Al-Rubeai 2009) or to engineer a cell line that secretes factors that promote growth and proliferation (Pak et al. 1996).

The addition of conditioned media to cell cultures has been reported to increase growth and productivity of hybridoma cells (Dutton, Scharer, Moo-Young 1999; Spens and Häggström 2005). It has been hypothesised that even in media rich in nutrients, the depletion of autocrine factors can result in cell death (Spens and Häggström 2007).

The mechanism by which proteins secreted into the media affect cell proliferation can be complex. An example of this is the mitogenic effect of IGF-1 on fibroblast cells which is controlled by a family of high affinity IGF-I binding proteins (IGFBPs) (Baxter and Martin 1989). Specifically, IGFBP-5 plays an important role in the release of IGF-I to receptors on the cell surface (Fowlkes et al. 1994; Mohan et al. 1995). It has been shown that only a cleaved form of IGFBP-5 can mediate the IGF-I receptor activation (Imai et al. 1997).

It has been reported that a secreted 88kDa serine protease that is responsible for the cleavage of IGFBP-5 (Busby et al. 2000). A study carried out in CHO cells revealed that while transferrin alone did not promote survival or proliferation and IGF-I helped maintain cell viability for an extended period of time, a combination of IGF and transferrin greatly improved cell survival and proliferation (Chun et al. 2003; Sunstrom et al. 2000).

To date there have been very few studies that have looked at the secretome of CHO cells. One such study however identified 24 proteins as being differentially-regulated at different phases of a 6 day growth cycle (Kumar et al. 2008b). Unfortunately due to the use of SELDI ToF as the sole means of analysis no protein identification is provided. Nonetheless this paper provides an interesting insight into the secretory profile of CHO cells. Some proteins accumulate until day four of the culture and their expression decreases over the following 3 days either due to them being of nutritional value to the cells or proteolytic activity. Other proteins appear only in the final days of culture suggesting that their presence is as the result of cell stress (Woolley and Al-Rubeai 2009) or apoptosis (Arden and Betenbaugh 2004).



The authors also submit the hypothesis that the appearance of these proteins late in the culture could also be as the result of proteolysis of the product or other proteins in the media. Sandberg et al. (2006) identified an MMP pro-enzyme of  $\geq 200$  kDa that was released from the CHO cells during culture, autoproteolysis of which yielded several smaller MMPs with non-specific protease activity. Interestingly, in an insect cell line, *Trichoplusia ni*, a secreted 48kDa metalloproteinase was identified as having growth promoting effects (Eriksson et al. 2005).

Although there have been several reports of different metalloproteinase variants in CHO culture media (Elliott, Hohmann, Spanos 2003; Mols et al. 2004; Sandberg et al. 2006) no studies to date have made a link between their presence and CHO cell proliferation despite it being well established that MMPs are involved in autocrine signalling and growth promotion (Fowlkes and Winkler 2002).

## **1.5 Introduction to Proteomics**

The study of proteins is called ‘Proteomics’, which is defined as the system-wide characterization of all the proteins in an organism in terms of their sequence, localization, abundance, post-translational modifications, and biomolecular interactions (Käll and Vitek 2011). While DNA contains all the genetic information required to construct an organism, a non-linear relationship often exists between the genome and the proteome of a cell as a result of events such as alternative splicing and the presence of non-coding RNAs (e.g. miRNAs) (Altelaar, Munoz, Heck 2013). Given that proteins are intrinsically linked with biological function, a complete knowledge of all the proteins in an organism would provide some of the greatest insights into characterising that system (Cox and Mann 2007). This section provides a brief overview of some of the techniques commonly used to characterise the proteome of some of those biological systems.

### **1.5.1 Proteomic sample preparation**

To prepare samples for analysis, cells and tissues need to be lysed to release the proteins of interest. This may be done either mechanically e.g. using a sonicator, or chemically, using a lysis buffer. In either event important consideration must be given to the solution that will contain the protein upon cell/tissue lysis. How the protein is solubilised will ultimately depend (i) on the nature of the protein – membrane proteins, typically hydrophobic in nature, can be more difficult to solubilise than proteins that are more hydrophilic in nature and (ii) how the protein will be analysed – analysis of the protein sample may either be ‘gel based’ i.e. utilise Polyacrylamide Gel Electrophoresis (PAGE) (discussed section 1.5.2.1) or ‘non-gel based’ i.e. utilise liquid chromatography/mass spectrometry (LC/MS) (discussed section 1.5.3). Where a non-gel based LC/MS approach is used, the preparation of a sample is sometimes referred to as ‘in solution’. It should be noted however, that the use of LC/MS is not restricted to non-gel based methods and can be used to identify proteins initially analysed by PAGE.

In order to aid protein solubilisation, the sample buffer may contain a combination of detergents and strong chaotropic reagents such as urea and thiourea. Although detergents are widely used in the disruption of tissue and cell and are often required to solubilise hydrophobic proteins such as membrane proteins (Arachea et al. 2012), the presence of a detergent may be unsuitable for certain applications. For example, sodium dodecyl sulphate (SDS) is commonly used as a reducing agent for the analysis of proteins by 1D polyacrylamide gel electrophoresis (PAGE), it is entirely unsuitable for 2D PAGE and (LC/MS) analysis. Although 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS) is suitable for use in both 1D and 2D PAGE methodologies, like SDS, CHAPS cannot be used for samples that are being prepared for 'in solution' analysis by LC/MS. It should be noted that in the last number of years a number of LC/MS compatible detergents have become available for enhanced protein solubilisation and digestion including RapiGest, PPS and ProteaseMAX (Chen et al. 2007; Lo et al. 2013). Cleavage of the detergent, usually under acidic conditions, prior to sample loading on the LC/MS system, ensures that the detergent does not interfere with LC/MS analysis. Compared to SDS and CHAPS, however, they are expensive and as such their application is likely to be limited to LC/MS analysis for the foreseeable future.

### **1.5.2 Sample separation**

It is estimated that a single cell contains 10,000 proteins, of which 4,000 – 6,000 can currently be routinely identified (Geiger et al. 2012). While the sheer number of proteins contained in a cell can itself hinder the identification of all proteins present in a single analysis, the fact that the expression of these proteins occurs over a large dynamic range adds another layer of complexity to any attempt to characterise a proteome (Picotti et al. 2009).

Given the complex nature of a proteome, some form of sample separation is typically required to facilitate its analysis. This section will provide a brief overview of two techniques that are commonly used to achieve this reduction in complexity; PAGE and liquid chromatography.

### **1.5.2.1 Protein separation by 2D-PAGE**

In two dimensional Polyacrylamide Gel Electrophoresis (2D PAGE), the proteins are electrophoretically separated along a pH gradient in the first dimension; this fractionates proteins on the basis of their isoelectric point. In the second dimension electrophoretic separation of proteins is generally based on the molecular weight of the protein, as this step usually occurs in the presence of SDS (Hanash and Taguchi 2010).

Although 2D PAGE based techniques are capable of resolving thousands of proteins (Li, Seillier-Moiseiwitsch, Korostyshevskiy 2011), it does suffer from a number of shortcomings including not being able to analyse proteins at the extremes of pH or molecular weight (Issaq and Veenstra 2008). Despite this, 2D PAGE has been successfully employed in the investigation of a wide variety of proteomes including human (Hardt et al. 2005), bacterial (Wickström et al. 2011), plant (Sarma, Oehrle, Emerich 2008) and CHO (Hayduk, Choe, Lee 2004)

Proteins separated on a gel may be visualised using a number of different stains, these can be fluorescent (e.g. SYPRO red) or colorimetric (e.g. Silver stain, Coomassie blue). The choice of stain will depend on a number of factors including desired level of sensitivity and ease of use. For example, Silver stain is two orders of magnitude more sensitive than Coomassie blue (Issaq and Veenstra 2008); however Coomassie is lower in cost and better suited to quantitative analysis (Dong et al. 2011). Although fluorescent stains such as SYPRO red are highly sensitive (as low as 2ng) and have a greater dynamic range than silver staining (Smejkal, Robinson, Lazarev 2004), specialist equipment (i.e. fluorescent scanner) is required to view the proteins on the gel (Cong et al. 2008).

Given that the colorimetric intensity of a spot is proportional to the quantity of the corresponding protein on the gel, the intensity of staining for each spot can be compared between gels containing samples from different conditions, thus relative quantitative information can be obtained using specialist software (Li, Seillier-Moiseiwitsch, Korostyshevskiy 2011).

An alternative approach to obtain relative quantitative information when comparing gels containing samples from two conditions is the use of a technique called Differential In-Gel Electrophoresis (DIGE). In this method the samples from the control, test, and a mixture of the control and test (to form an internal standard, providing both inter- and intra-gel matching) are pre-labelled with a cyanine dye (cy2, cy3 or cy5). The three labelled sample sets are then combined and run on a single gel. This means that matching proteins from the different samples will all migrate to the same location on the gel, yet each protein in that location will have its own unique fluorescent marker. The gel is then scanned at a suitable wavelength to facilitate detection of each labelled protein, making this approach more accurate for quantitation than the staining methods (Issaq and Veenstra 2008).

#### **1.5.2.2 Protein separation by Liquid chromatography**

The use of liquid chromatography (LC) is another means by which sample complexity can be significantly reduced. Separation is based on an interaction occurring between stationary phase (typically a resin contained in a column) and the analyte (protein/peptide) in a mobile phase (solution that is suitable for the analyte and that facilitates the interaction between the analyte and stationary phase). A number of different resins may be used in the separation of a sample, each of which takes advantage of different characteristics of the sample. These can be broadly categorised as affinity (biological interaction), ion-exchange (charge), size exclusion (size) and reversed phase (polarity). Referred to as multi-dimensional protein identification technology (Mud-PIT) when coupled to a mass spectrometer, the use of multiple chromatography columns enabling the further separation of complex samples is gaining in popularity (Fränzel and Wolters 2011). The use of liquid chromatography in conjunction with mass spectrometry is further discussed in section 1.5.3. It should be noted that, with the exception of phosphoprotein enrichment and its use for the reduction of the IgG molecule from conditioned media samples, the work discussed in this thesis primarily employs liquid chromatography for peptide separation.

### **1.5.3 Mass spectrometry for protein identification and characterisation**

Mass spectrometry has become the cornerstone upon which proteomics has come to rely. Mass spectrometry is used in a wide range of proteomic applications, from the qualitative identification of a single protein in a 2D-gel spot to large scale quantitative experiment seeking to decipher the signalling dynamics of a biological system (Gygi et al. 2000; Munoz et al. 2011).

The following section provides a brief overview of how a protein sample is prepared for, and analysed by a mass spectrometer culminating in the identification of the proteins contained in that sample.

#### **1.5.3.1 Sample preparation for Mass Spectrometry**

Mass spectral proteomic analysis may be referred to as ‘top down’, whereby proteins are analysed as intact molecules in the mass spectrometer, or more commonly, ‘bottom up’, where the proteins are enzymatically digested and the resulting peptides are analysed (Armirotti and Damonte 2010). Given the intrinsic structural complexity of proteins in their native form, analysis by mass spectrometry is challenging; sufficient separation of complex samples and obtaining sufficient information to identify the proteins can prove difficult (Yates and Kelleher 2013). However, the top down strategy has proven useful for the identification of post translational modifications (PTMs) and sample preparation time is shorter compared to the bottom up approach (Armirotti and Damonte 2010). For this reason top down analysis is typically used for the analysis of samples containing a homogeneous, or very small heterogeneous, population of proteins. Despite having a longer sample preparation time, the bottom up approach is more established and has been found to lend itself well to the analysis of complex protein mixtures (Zhang et al. 2013).

Assuming the protein sample has been obtained from a cell or tissue source using mechanical or chemical means as described in section 1.5.1, sample preparation of proteins for analysis at the peptide level follows a typical workflow involving protein denaturation, reduction, alkylation and digestion. The presence of LC/MS compatible detergents such as ProteaseMax (described in section 1.5.1) and/or the chaotropic agent urea in the sample buffer causes the protein to lose its quaternary structure to form a single stranded polypeptide.

In other words, the protein becomes denatured. It should be remembered however that some disulphide bonds may remain intact, causing the protein to retain some of its structure. Therefore a reducing agent such as dithiothreitol (DTT) is required for the reduction of disulfide bonds on the protein. The free sulfhydryls are then reacted with a reagent such as Iodoacetamide to prevent the disulfide bonds reforming; this step is commonly referred to as alkylation (Boja and Fales 2001). Lastly, the protein is enzymatically digested. Although there are a number of enzymes that can be employed, trypsin is perhaps one of the most commonly used enzymes for protein digestion (Swaney, Wenger, Coon 2010). This is because it is a protease with a specificity for cleaving at the C-terminus of lysine and arginine, which tends to generate peptides of 7-20 amino acids long, a length well suited to analysis by LC/MS (Tran et al. 2010). Other commonly used digestion enzymes include Lys-C which specifically cleaves at the C-terminus of lysine residues and Asp-N which cleaves at the amine side of aspartic and cysteic acid (Zhang et al. 2013).

### **1.5.3.2 Nanoflow Liquid Chromatography coupled to Mass Spectrometers**

The enzymatic digestion of a sample, possibly containing hundreds or even thousands of proteins, results in tens of thousands of peptides being generated. As such, some means of sample separation is often required in order to allow the mass spectrometer time to analyse such a large number of peptides. High pressure liquid chromatography (HPLC) using a reverse phase (RP) is compatible with mass spectrometry, enabling a RP-HPLC system to be directly coupled to a mass spectrometer, as a result RP-HPLC is the most common means of reducing this sample complexity (Moruz et al. 2013). With RP-HPLC, the peptide mixture dissolved in a polar liquid mobile phase, is passed over a non-polar stationary phase contained in the chromatography column resulting in the peptides binding to the stationary phase based on their hydrophobicity. By increasing the aqueous/organic ratio of the mobile phase to form a solvent gradient, peptides are removed (eluted) from the column with hydrophilic peptides eluting first and hydrophobic peptides eluting towards the end of the gradient (Pitt 2009). Proteomic peptide separations generally flow at very low flow rates 200-400nl/min through narrow chromatography columns usually 75µm wide and finally leave the HPLC through a spray needle/tip between 10-20 µm wide into a mass spectrometer.

### 1.5.3.3 The basic components of a mass spectrometer



**Figure 1.5.1** Schematic of the basic components of a mass spectrometer

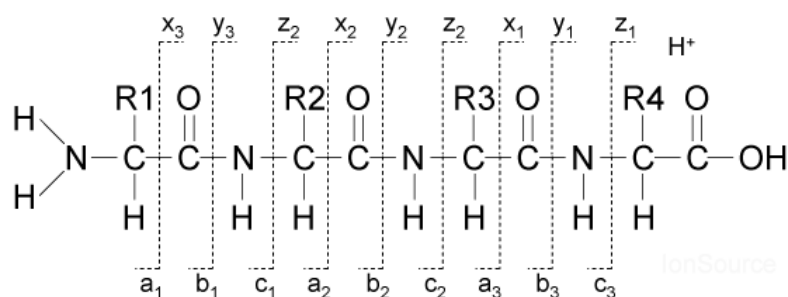
The first step in the analysis of a protein sample by mass spectrometry is the loading of the sample into the mass spectrometer and whereupon it is ionised by the ion source (Figure 1.5.1). Examples of ion sources include Matrix Assisted Laser Desorption Ionisation (MALDI) or Electrospray Ionisation (ESI). In the case of MALDI, samples are ‘spotted’ onto a plate which is then loaded into the mass spectrometer for analysis. It should be noted that samples may be pre-fractionated by liquid chromatography prior to spotting onto the plate in order to reduce sample complexity (Mirgorodskaya et al. 2005). ESI on the other hand, (referred to as nano spray ionisation (NSI) when being sprayed from a nanoHPLC system directly into a mass spectrometer), applies a voltage (1.5KV) to the liquid samples which causes an evaporating aerosol of single ions which are then drawn into the mass spectrometer by the vacuum system.

Once ionised, the charged molecules accelerate through the system and encounter magnetic and/or electrical fields from mass analysers, deflecting their path based on their mass-to-charge ratio ( $m/z$ ). Examples of mass analysers include Ion Traps, Quadrupoles and Time of Flight (ToF), each of which has their own specific characteristics. Mass analysers may either be used to separate all the ions in a sample for global analysis or act as a filter to deflect specific ions for analysis by the detector. Ion detectors are electron multipliers or microchannel plates which, on contact by an ion, emit a cascade of electrons. The resulting cascade amplifies the ion for improved detection sensitivity. This all occurs under a vacuum ( $10^{-8}$  torr) to remove any contaminating (i.e. non sample) ions which could collide with the sample ion. Software is then used to analyse the ion detector data generated and the information is output is typically visualised as a graph that arranges the detected ions by their individual  $m/z$  and relative abundance.



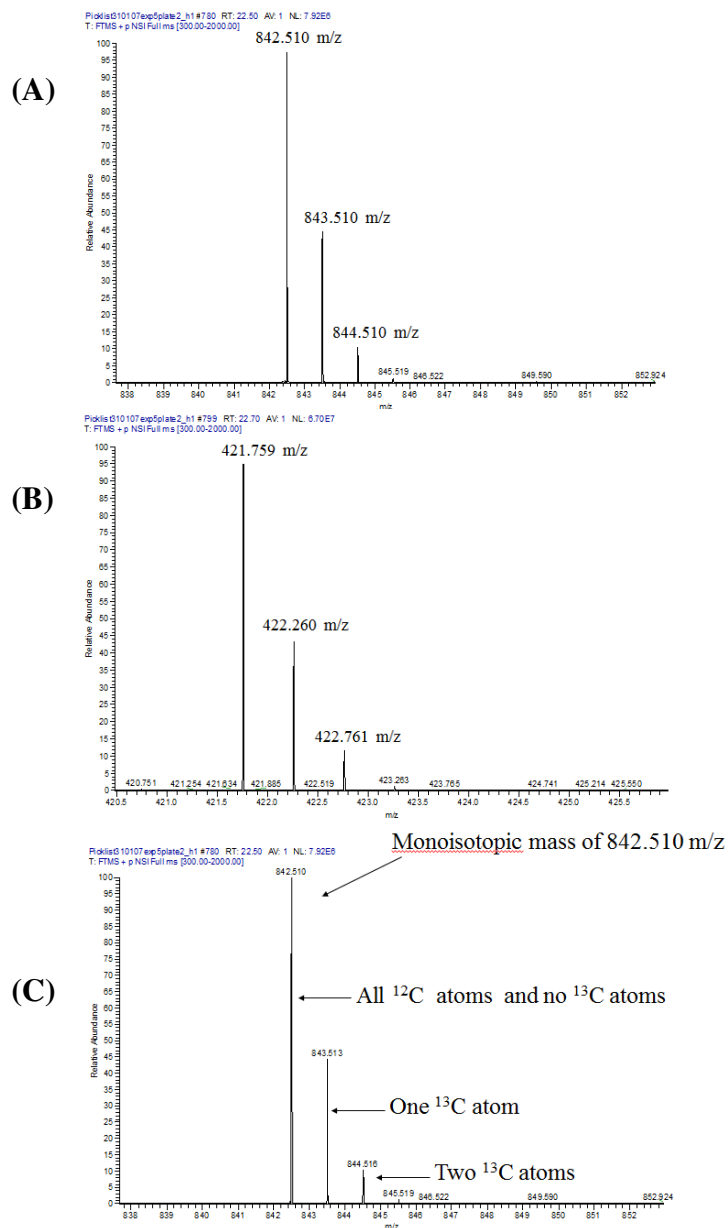
At this point, preliminary mass spectrum (MS1) data has now been acquired from the intact peptide (Angel et al. 2012). Further information about the detected ions can be obtained by using tandem mass spectrometry (MS/MS). Using this approach, a specific ion of interested is selected based on its  $m/z$  and is fragmented by colliding the ions with a chemically inert gas (Walther and Mann 2010). The fragments produced are separated on their individual  $m/z$  ratios and detected by the ion detector. It should be noted that some instruments are capable of generating fragmentation information on the already fragmented ions, this strategy is referred to as MS/MS/MS or MS<sub>n</sub> (further discussed in section 1.5.7).

There are a number of means by which ion fragmentation can be induced including Electron Transfer Dissociation (ETD), High Collision Dissociation (HCD) and Electron Capture Dissociation (ECD) but perhaps one of the most commonly used methods is Collision-induced dissociation (CID) (Frese et al. 2011). In the case of CID, the amide bond along the peptide backbone dissociates resulting in two fragments being produced; one containing the N-terminus and the other containing the C-terminus. The Roepstorff–Fohlmann nomenclature denotes N-terminal fragments with the letters a, b and c and C-terminal fragments with the letters x, y and z (Roepstorff and Fohlman 1984; Sadygov, Cociorva, Yates 2004). The use of CID in ion traps, triple quadrupoles and quadrupole time of flight instruments tend to result in b-ions, y-ions and neutral losses of water and ammonia dominating the mass spectrum (Figure 1.5.2)



**Figure 1.5.2** The proposal for a common nomenclature for sequence ions in mass spectra of peptides as put forward by Roepstorff and Fohlman (1984).

Another consideration is the charge state of the ion in the spectrometer as this is used to determine the  $m/z$  ratio of the specific ion. Although ToF instruments only produce molecules with a +1 charge, ion traps can produce multiple charged species. Therefore the charge state is required to determine the mass of the molecule, however it should be noted that it is the use of tandem MS/MS that allows for the molecule identification (further discussed in 1.5.3.4). The ion charge can be determined by distance between the peaks of the naturally occurring isotopes of different masses (1Da) on the  $m/z$  scale. If, for example, the distance between peaks is one unit on the  $m/z$  scale, then the charge of the peptide would also be one (as seen with isotopes 842.51 and 843.51 in Figure 1.5.3 A). If however, the distance was 0.5, then the charge would be two (as seen with isotopes 421.759 and 422.260 in Figure 1.5.3 B). This distance between the peaks results from the 1% probability that each carbon atom is a  $^{13}\text{C}$  isotope instead of the usual  $^{12}\text{C}$  atom (Figure 1.5.3 C).



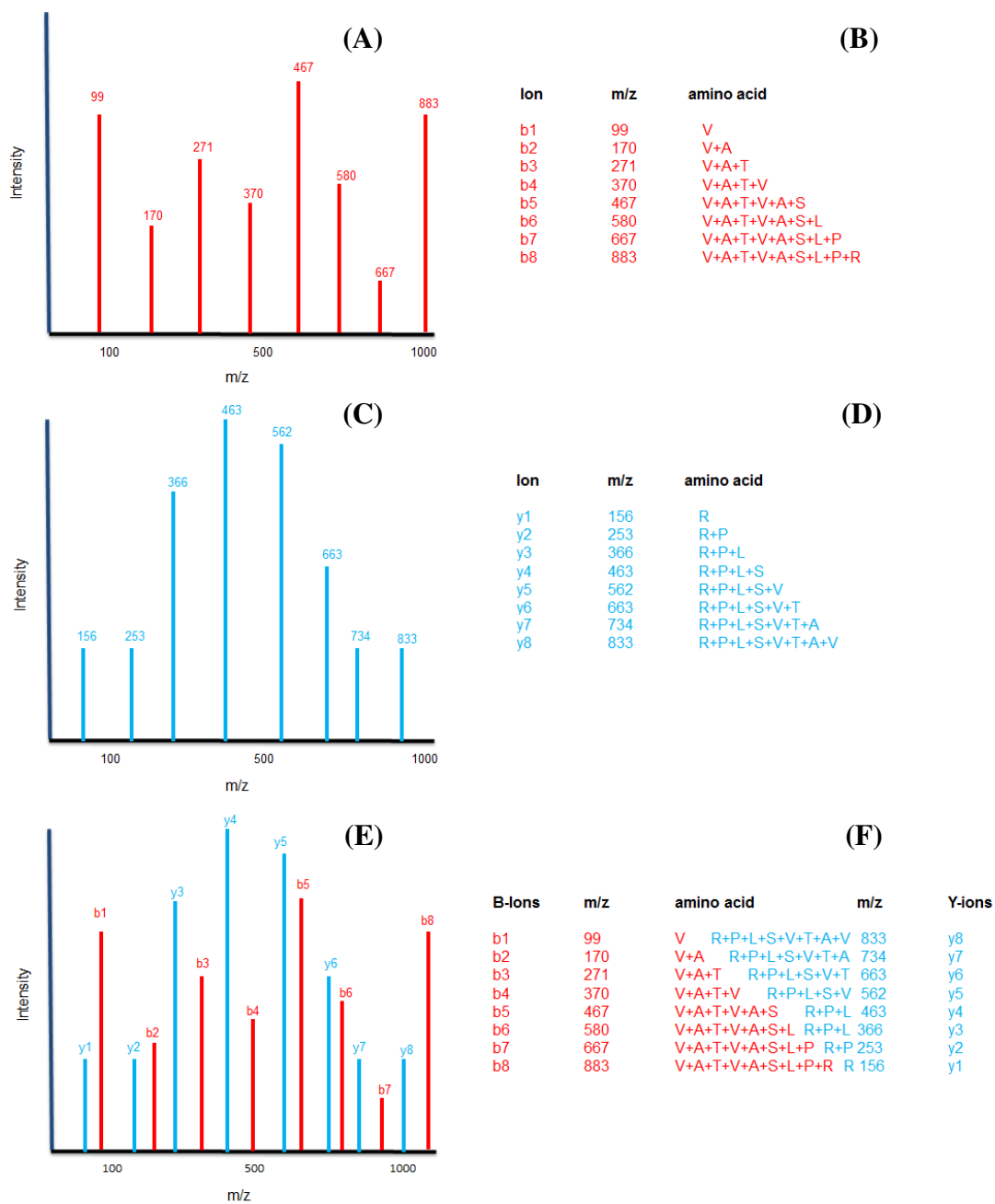
**Figure 1.5.3** (A) Full MS of singly charged peptide as denoted by difference of 1Da between the isotopic peaks 842.51 and 843.51 (B) Full MS of double charged peptide as denoted by difference of 0.5Da between the isotopic peaks 421.759 and 422.260 (C) Distribution of naturally occurring isotopes of different masses caused by the frequency with which  $^{13}\text{C}$  occurs in nature.

### 1.5.3.4 Determining the peptide sequence from mass spectrometry data

Using the mass of the ions detected following MS/MS, it is possible to deduce the amino acid sequence present in a given peptide by calculating the difference in the mass between the peptide fragments as they increase in size from the N terminus or C terminus. In its simplest form, this series of specific mass differences corresponds to a successive amino acid (See Table 1.5.1 for the list of amino acids and their associated masses). For example, when a peptide is fragmented in a mass spectrometer using CID, a unique pattern is generated creating an ion ladder made up of b-ions and y-ions. Figure 1.5.4 (A) shows an example of a series of b-ions from a fragmented peptide with an m/z of 883 and associated amino acids (Figure 1.5.4 B). Figure 1.5.4 (C) shows an example of a series of the y-ions from the same fragmented peptide (m/z of 883) and associated amino acids (Figure 1.5.4 D). The information obtained from the detection of b ions complements the data obtained from the detection of y ions (Figure 1.5.4 E and F). As such, the b and y series of ions observed can be used to determine the amino acid sequence of a given peptide.

Amino Acid	1-letter code	3-letter code	monoisotopic mass (Da)
glycine	G	Gly	57.021464
alanine	A	Ala	71.037114
serine	S	Ser	87.032028
proline	P	Pro	97.052764
valine	V	Val	99.068414
threonine	T	Thr	101.047678
cysteine	C	Cys	103.009184
isoleucine	I	Ile	113.084064
leucine	L	Leu	113.084064
asparagine	N	AsN	114.042927
aspartic acid	D	Asp	115.026943
glutamine	Q	GIN	128.058578
lysine	K	Lys	128.094963
glutamic acid	E	Glu	129.042593
methionine	M	Met	131.040485
histidine	H	His	137.058912
phenylalanine	F	Phe	147.068414
arginine	R	Arg	156.101111
tyrosine	Y	Tyr	163.063329
tryptophan	W	Trp	186.079313

**Table 1.5.1** List of amino acids, their 1 and 3 letter codes, and associated monoisotopic masses.



**Figure 1.5.4** shows an example of the peptide VATVSLPR and the ion ladder made up of b-ions and y-ions resulting from its fragmentation. (A) The series of b-ions from a fragmented peptide with a m/z of 883. (B) Corresponding amino acid to the MS/MS fragments generated as determined by the b-ions. (C) The series of y-ions from the same fragmented peptide shown in (A). (D) Corresponding amino acid to the MS/MS fragments generated as determined by the y-ions. (E) Displays the overlap between b and y-ions as a result of the peptide fragmentation and (F) how the ion series from both b and y-ions complement each other to provide peptide sequence information.

To obtain the mass of every amino acid in a fragmented peptide (i.e. de novo sequencing) in a complex protein sample is not currently a realistic goal. Although peptide fragmentation may only contain partial sequence information, when coupled with information on the mass of the peptide, this may still be sufficient to derive protein identification when searched against a well annotated database (Sadygov et al., 2004) (the use of proteomic databases is discussed further in section 1.5.8).

#### **1.5.4 Mass spectrometry data output**

During analysis of a sample by LC/MS, data is collected in a 'file' on a computer connected to the LC/MS system. A mass spectrometry file can contain peak list information such as peptide mass value and peak area as well as MS<sub>n</sub> peak list information including intensity values and beginning and end ion information. While there are a number of file formats specific to certain vendors e.g. RAW files (Thermo Scientific), wiff files (Ab Sciex) or as multiple files within a folder with a .raw extension (Waters and Agilent), open source file formats are also available such as mzML. Open source formats provide the user with greater freedom to share information and develop software for data analysis (Deutsch 2010), while vendor specific formats typically require commercially available software to facilitate their use (Deutsch 2012). The mass spectra contained in these files may be represented either as 'profile' data which contains all acquired data points or as 'centroided' data where only m/z and intensity pairs of the detectable peaks are extracted (Deutsch 2012). The file size can depend on the complexity of the sample and the nature of the data being collected (Patel et al. 2009).

#### **1.5.5 Bioinformatic analysis of mass spectral data**

Once a sample has been analysed on the mass spectrometer, the information gathered is then used to identify the proteins present in the analyte. Although manual interpretation can identify amino acid sequences from peptides, many proteomics experiments utilise the power of protein databases for automated spectrum interpretation.

Publically available databases such as uniprot-swissprot, NCBI etc, contain the amino acid sequence information for all the respective proteins identified in a particular species.

Databases are continuously changing as new information is obtained and added, or existing information modified. How well annotated a database is will depend on how much experimental work has been carried out in a particular species. For example, species such as *Homo sapiens*, *Mus musculus*, *Drosophila melanogaster*, *Caenorhabditis elegans*, *Arabidopsis thaliana*, and *Saccharomyces cerevisiae* are generally considered to be well annotated (Holt and Yandell 2011). Using a search algorithm such as MASCOT or SEQUEST (Cottrell and London 1999; Eng, McCormack, Yates 1994), the mass spectral data is searched against a database. The objective of a spectral database search is to identify the best sequence match to the actual spectrum generated by the mass spectrometer. There are four basic approaches that may be used for this comparison, this includes; descriptive, interpretative, stochastic and probability-based modelling (Sadygov, Cociorva, Yates 2004).

A theoretical digestion of the database takes place when the operator of the software inputs the enzyme used to digest the protein sample. Therefore, if trypsin was used as the digestion enzyme, all the protein sequences in the database will be theoretically cleaved at the amino acids lysine and arginine to provide every combination of amino acid sequence possible including miss cleaved events. Naturally searches become more complex as static or dynamic modifications are added to the search criteria (Savitski, Nielsen, Zubarev 2006). How the mass spectral data is then compared against this theoretical digest depends on the software used. For example, SEQUEST, described as a descriptive model, compares fragment ions against the MS/MS spectrum and generates a preliminary score for each amino acid sequence. A cross correlation analysis is then performed against the top 500 preliminary scoring peptides by correlating theoretical, reconstructed spectra against the experimental spectrum and a cross correlation score is given to any peptide match (Xcorr) (Eng, McCormack, Yates 1994).

MASCOT uses a probability-based approach for protein identification by first comparing the calculated peptide masses for each entry in the sequence database with those identified by the mass spectrometer. Matches using mass values (either peptide masses or MS/MS fragment ion masses) are always handled on a probabilistic basis. A score is then awarded based on the absolute probability that the observed match of either peptide masses or MS/MS fragment ion masses is a random event (Cottrell and London 1999).

The use of such programmes such as SEQUEST or MASCOT relies on the individual to use their discretion to select a cut-off score that will prevent the inclusion erroneous results while at the same time not excluding correctly assigned identifications (Cooper 2011).

#### **1.5.5.1 CHO specific proteomic databases**

Historically, due to the limited number of proteins available in the CHO database, studies relied on sequence homology between CHO and well annotated mammalian databases such as Human, mouse or rat (Baik et al. 2011; Carlage et al. 2009; Doolan et al. 2010; Kantardjieff et al. 2010; Lee et al. 2010). It is only more recently that two CHO specific databases have become available (Meleady et al. 2012; Xu et al. 2011). The first, released by Xu et al. in 2011, was generated following genomic sequencing of CHO-K1 cells (Xu et al. 2011).

A year later data from the expression of cDNA in CHO cells growing under different conditions was used to produce a second database (Meleady et al. 2012). Growing the cells under different conditions resulted in the expression of a greater variety of genes, as such it was found that the use of both CHO databases provided increased coverage of the CHO proteome of CHO (Meleady et al. 2012). Although CHO specific databases have been desired for some time, this recent advance in CHO proteomics is clearly a significant step forward as proteins will be identified with increased confidence and new, previously unidentified proteins are discovered. As more information about the CHO genome becomes available (Lewis et al. 2013), it is likely that even greater proteome coverage will be possible in the near future.



### 1.5.6 Post Translational Modifications (PTM) Analysis using Mass Spectrometry

Much work has been carried out on the development of different peptide fragmentation methodologies, understanding their basic mechanics and implementing their use. A significant development over the past few years has been the development of the “high-high” strategy (Mann and Kelleher 2008) - high resolution at both MS and MS/MS levels - which has been particularly useful for the analysis of PTMs (Olsen and Mann 2013). An example of one instrument capable of implementing this “high-high” strategy is the LTQ-OrbitrapXL from Thermo Fisher Scientific which has allowed for high mass accuracy (<2ppm) and high resolution analysis (100,000 FWHM). This high resolution capability enables the identification of specific amino acid modifications on a given peptide based on the observation of the mass difference associated with the particular modification (Table 1.5.2).

Modification	Amino Acid Involved	Context	Mass Difference (Da)
Oxidation	M	Sample preparation (Variable-not all methionines are modified)	15.995
Carbamidation	C	Sample preparation (Complete – all cystines become modified)	57.021
Phosphorylation	S, T, Y	Post translational	79.966
Water loss	S, T	Mass Spec	-18.01
Deamination	N,Q	Post translational	0.984

**Table 1.5.2** List of modifications for amino acids and the mass difference associated with the modification.

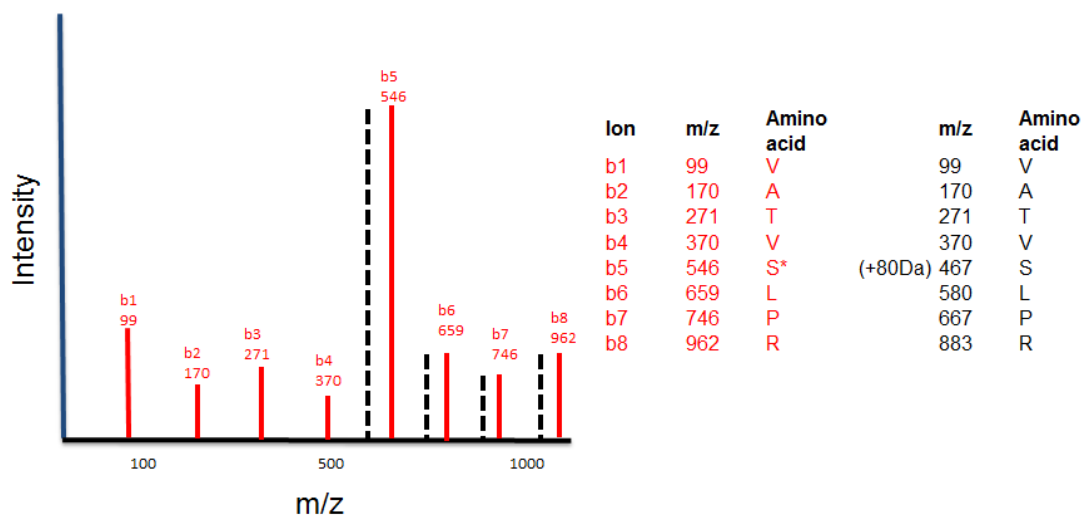
It should also be pointed out that in an instrument such as an LTQ-OrbitrapXL, fragmentation can take place either in the ion trap or the Orbitrap of the mass spectrometer. Where the fragmentation occurs will depend upon the end application as the ion trap uses faster rates of fragmentation but has low resolution while the Orbitrap is capable of high resolution, but has slower fragmentation speed compared to the ion trap. High resolution is crucial for applications such as a label-free experiment where the ability of the mass spectrometer to discern between separate peptide peaks is defined by its resolving power (Cox and Mann 2011) (Quantitative label-free LC/MS is further discussed in section 1.5.9).

In other words, in instances where peptides are co-eluting, a low resolution mass spectrometer may observe the co-eluting peptides as one peak, whereas a high resolution mass spectrometer will be able to differentiate between the different peptide profiles (Mann and Kelleher 2008). The slow scan speed of the Orbitrap may result in less MS/MS data being produced as less fragmentation occurs. MS and MS/MS tasks can, however, be divided between the Orbitrap and the ion trap enabling high resolution MS and multiple MS/MS fragmentation data to be gathered simultaneously (Yates, Ruse, Nakorchevsky 2009).

### **1.5.7 Identification of phosphopeptides by mass spectrometry**

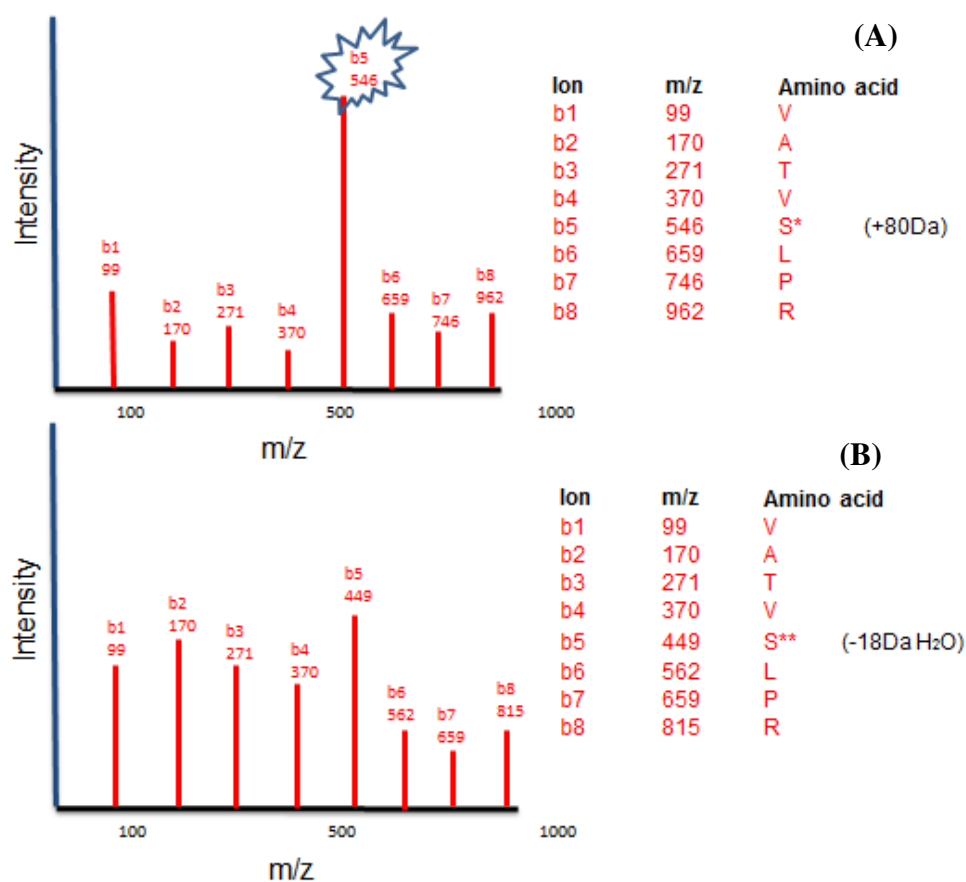
The identification of proteins through the analysis of their respective peptide was outlined in section 1.5.5. The following piece provides an overview of how sites of phosphorylation are determined during phosphopeptide analysis by LC/MS.

The fragmentation pattern of a phosphorylated peptide will appear shifted by a mass of +80Da relative to predicted mass of the unphosphorylated species due to the presence of the phosphoric acid group as shown in Figure 1.5.5. The use of collision-induced dissociation to fragment peptides containing phosphorylated threonine and serine results in the cleavage of the phosphoester bond and dissociation of Phosphoric acid (H<sub>3</sub>PO<sub>4</sub>) as a neutral loss (neutral species). This fragmentation, particularly with the use of high resolution mass spectrometers, is usually adequate to obtain sufficient sequence information to enable peptide identification (Villén, Beausoleil, Gygi 2008)



**Figure 1.5.5** Diagram of a phosphorylated MS/MS sequence in red overlapped with the unphosphorylated sequence in black. Phosphopeptides are indicated by peaks shifted by an increase of 80Da relative to the predicted unphosphorylated peptide mass due to the presence of a phosphate group. This is also shown in the accompanying table with the mass of the ions from the phosphorylated peptide indicated in red and the unphosphorylated sequence in black. (Phosphorylated serine denoted by \*)

However it should be pointed out that CID of phosphoric acid from a phosphopeptide is usually accompanied by only partial fragmentation of the peptide backbone with the fragmented ions being dominated by the neutral loss of phosphoric acid which can make peptide identification difficult. In other words, the presence of the phosphate group ‘shields’ the peptide from fragmentation resulting in the loss of peptide sequence information (Figure 1.5.6) (Villén, Beausoleil, Gygi 2008). In such instances, multiple fragmentation (MS<sub>n</sub>) strategies may be used to obtain peptide sequence data. This is a useful strategy for the study of PTMs such as phosphorylation. In instances where the MS/MS fragmentation results in the loss of a phosphate group from the peptide, the loss of this mass is detected by the mass spectrometer and triggers a third fragmentation event (MS/MS/MS). (This can result in the further loss of a loss of water (-18Da) from serine and threonine residues (Steen and Mann 2004)). Frequently this third fragmentation event aids in the gathering of further sequence information that would otherwise not have been observed due to the ‘shielding’ effect of the phosphate group (Villén, Beausoleil, Gygi 2008). Often the triggering of this MS<sub>3</sub> event is used as a ‘validation’ of the phosphorylation of the peptide of interest (Yu et al. 2007).



**Figure 1.5.6** (A) MS/MS showing a predominant peak corresponding to the neutral loss of H<sub>3</sub>PO<sub>4</sub> from a phosphorylated Serine. However the signal-to-noise ratio of the downstream ions are lower (Phosphorylated serine denoted by \*). (B) MS/MS/MS of neutral loss peak showing increased signal-to-noise ratio of all ions and the associated loss of water from serine as indicated in the accompanying table.

The identification of site-specific phosphorylation is not a trivial task, particularly as the size of the data sets obtained from large scale phosphoproteomic profiling campaigns continue to increase, manual validation becomes increasingly onerous (Beausoleil et al. 2006). Often the use of phosphorylation site localisation scoring algorithms to determine which amino acid residue is phosphorylated with confidence is required. Ascore, PTM score and PhosphoRS are examples of some of the site determining algorithms currently available, each of which uses a different approach to identifying the site of phosphorylation (Beausoleil et al. 2006; Olsen et al. 2006; Taus et al. 2011).

For example, the Ascore algorithm uses a probability based approach to score the likelihood of correct site assignment based on the presence or absence of site-determining fragment ions (Beausoleil et al. 2006).

Although PTM-Score also utilises a probability-based approach, unlike Ascore, PTM-Score groups potential phosphorylation sites into four categories (class I – class IV) based on their distinguishing fragment ions, with class I being the best quality (Beausoleil et al. 2004). In this way, rather than using a definitive threshold to claim a site has been identified with confidence, the quality of the fragmentation pattern is ranked. Despite this, those sites identified as class I are commonly deemed confident in phosphoproteomics studies and reported to public databases (Wiese et al. 2013). The leading commercially available tool PhosphoRS operates on a similar principal to PTM-Score (Collins et al. 2014; Wiese et al. 2013), however PhosphoRS distinguishes itself from PTM-Score through its ability to provide site localisation information based on data obtained from a number of dissociation methods including CID, HCD, and ETD on high resolution instruments (Taus et al. 2011), whereas PTM-Score was originally designed for use with CID data obtained from low resolution instruments (Chalkley and Clauser 2012).

### **1.5.8 Phosphoproteomic databases**

Given the important role that phosphorylation plays in many cellular processes such as growth and apoptosis, little wonder it has attracted much attention from those attempting to gain further insights into the working mechanisms of a cell. The presence or absence (or abundance) of phosphorylation on different proteins can have dramatic effects on cellular function. Presenting an understanding of all the phosphorylation sites and their precise function is well beyond the capabilities of any one research group. As such, phosphorylation specific data bases are seen as a useful resource to pool all current knowledge on the phosphoproteome of different species. Since the creation of the first phospho-specific database in 1998 (Blom, Kreegipuu, Brunak 1998), a large number of databases specific to specific species have been developed (some of these are further reviewed elsewhere (Xue et al. 2010)).

Some of the more well-known databases include PhosphoSitePlus, a manually curated database developed by Cell Signalling Technologies and largely contains information on PTMs associated with human and mouse proteins (Hornbeck et al. 2012). PhosphoSitePlus currently contains 155,545 non-redundant phosphorylation sites. Phospho.ELM stores in vivo and in vitro phosphorylation data on 42,500 non-redundant phosphorylation sites extracted from the scientific literature and phosphoproteomic analyses of human, mouse, *Drosophila* and *C.elegans* proteins (Dinkel et al. 2011). Another database, PHOSIDA originated from an experiment which identified and quantified phosphorylation events in HeLa cells as a result of EGF stimulation (Beausoleil et al. 2004), it now contains information on over 70,000 phosphorylation sites from nine different species (Gnad, Gunawardena, Mann 2011). Over the last number of years improvements in sample fractionation, LC/MS technologies and search algorithms has led to a dramatic increase in our ability to generate vast amounts of data on the phosphoproteome (Dulla et al. 2010; Zhou et al. 2012). While this represents a huge step forward in terms of profiling different biological systems, it also presents a number of problems for the scientific community.

Given that different databases have different acceptance criteria for data regarding peptide identification and phosphosite localisation any effort to reduce redundancy is going to be a challenge (Rogers and Foster 2009). While there are a number of comprehensive databases currently available, no one site as yet contains a full data set (Rogers and Foster 2009). This means that there is still a gap in our knowledge of the phosphoproteome of any given species. Coupled with the fact that the precise number of phosphorylation sites in an organism are based on estimates, just how big this gap in our knowledge is has yet to be determined. Even when all phosphorylation sites become known, identifying the role each one plays remains an enormous challenge (Newman et al. 2013). While phosphoproteomic databases are a very useful resource for pooling current knowledge, there is still a significant amount of work to do before any phosphoproteome can be said to be fully characterised.

### 1.5.9 Quantitative label-free LC/MS

Quantitative mass spectrometry based means of proteome quantitation has typically relied on the use of labelling in order to obtain quantitative information. Labelling strategies such as iTRAQ (Ross et al. 2004), iCAT (Gygi et al. 1999) (chemical labelling) or SILAC (metabolic labelling) (Ong et al. 2002) have been used with great success in a large number of proteomic studies. The primary advantage of these labelling methods is the reduction in both sample to sample variation and mass spectrometry analysis time due to the pooling of samples (America and Cordewener 2008). These methods are not without their limitations, however; aside from expense (Grossmann et al. 2010; Winefield, Williams, Himes 2009), poor labelling efficiency may mean that not all peptides in the sample are represented leading to a decrease in protein sequence coverage (Melanson, Avery, Pinto 2006; Sakai et al. 2005). Labelling strategies also often suffer from a limited dynamic range (Ow et al. 2009; Wu et al. 2003).

The number of comparators in a given experiment is also limited by the number of labels available. Although this has improved with the advent of increased number of labels e.g. iTRAQ 8-plex (Choe et al. 2007) and numerous SILAC isotopes, this may lead to other problems as the increase in co-eluting peptides may make data acquisition difficult (America and Cordewener 2008).

One means of overcoming these issues is to use a label-free approach (Patel et al. 2009). Label-free LC-MS can be conducted using one of two methods; (i) spectral counting, or (ii) ion peak intensity. Spectral counting protein quantification is calculated using the linear correlation that exists between MS spectra intensity and peptide abundance (Liu, Sadygov, Yates III 2004).

Peak ion intensity uses the chromatographic peak intensity of peptide precursor ions to calculate protein abundance in different samples. Since there is no labelling involved, each sample must be analysed individually and data from each run collated once all the samples have been analysed. So while time may be saved during sample preparation (since labelling is not used), time spent analysing all the samples on the mass spec may negate any time saved.

Quantitative label-free LC/MS has successfully been used in a wide a variety of proteomic investigations such as cancer, both whole cell (Zhu et al., 200 and

secretome (Lawlor et al. 2009), serum (Levin et al. 2007), cardiac membrane (Donoghue et al. 2008) and prokaryotic (Zhang et al. 2006). More recently it has been used to study post translational modifications (PTMs) such as phosphorylation (Froehlich et al. 2011) and glycosylation (Rebecchi et al. 2009).

While data normalization and statistical analysis are required for spectral counting strategies (Zybailov et al. 2006), their implementation is perhaps more important for ion peak intensity techniques which requires normalisation to correct for minor variations in sample preparation and sample injection (Chelius and Bondarenko 2002). High chromatographic reproducibility is a basic prerequisite for this method. Poor reproducibility results in changes in retention time and  $m/z$ , and can create problems when trying to accurately quantitate protein abundance. Many of these issues have now been addressed through the use of (i) sophisticated algorithms to normalise and allow for slight shifts in retention time (Sandin et al. 2011) and (ii) chromatography equipment with improved reproducibility (Qian et al. 2006).

An interesting potential source of error raised by Turtoi and co-workers is that only a limited number of points can be acquired over a peak over a given time (Turtoi, Mazzucchelli, De Pauw 2010). While this does not cause significant problems at the moment due to the already fast scan speed of modern mass spectrometers such as the LTQ-FTMS or LTQ-Orbitrap, undoubtedly future generations of mass spectrometers will offer improved performance. It is perhaps also worth mentioning that a small number of studies have used matrix assisted laser desorption ionization (MALDI) for label-free quantification (Getie-Kebtie et al. 2011; Hattan and Parker 2006).

Despite a small number of studies being done, it is difficult to determine which method is superior. A comparative study by Old and co-workers found spectral counting to be more sensitive for detecting proteins that undergo changes in abundance, while ion peak intensity provided more accurate estimates of protein ratios (Old et al. 2005).

However if peptide abundance is low (thus reducing the number of spectral counts as a result), spectral counting becomes a less sensitive method (Old et al., 2005). Another study found the two methods to be comparable in terms of response linearity and reproducibility (Stevenson et al. 2009).



## 1.6 Analysis of protein phosphorylation

The reversible phosphorylation of protein is frequently used within cells to control signalling events for a wide variety of functions such as growth, cell cycle, metabolism, transcription and apoptosis (Dix et al. 2012; Volarević and Thomas 2000). Generally speaking, phosphorylation of serine, threonine and tyrosine are the most widely reported phosphorylated amino acids in a ratio of 1,800:200:1 (Gronborg et al. 2002). The complexity of the cellular processes that rely upon this post translational modification for such signalling events becomes apparent when it is realised that roughly one third to one half of all eukaryotic proteins can become phosphorylated (Cohen 2001; Olsen et al. 2010). It should be pointed out however that although phosphorylation sites in eukaryotic cells are evolutionarily well conserved (Gnad et al. 2010), many of these phosphorylation events may be physiologically irrelevant and may not be required for cell signalling (Dephoure et al. 2013). Despite the frequency with which phosphorylation occurs in the proteome, enrichment strategies at the protein level or the peptide level or a combination of the two is often required for phosphoproteomic analysis (Beausoleil et al. 2004). This is primarily due to the low abundant nature of many phosphoproteins coupled with the fact that phosphorylation is often sub-stoichiometric.

From the perspective of analysis by mass spectrometry, the relatively high abundance of the non-phosphorylated peptides often suppresses detection of the phosphorylated species (Wang et al. 2007). Similarly, the presence of low abundant phosphoproteins on a 2D-gel may be masked by high abundant proteins (Chen, Southwick, Thulin 2004). For these reasons, enrichment strategies are often utilized in the study of phosphorylated proteins (Beltran and Cutillas 2012; Fíla and Honys 2012). The following sections will discuss some of the enrichment strategies commonly used and a table of their associated advantages and disadvantages is provided in Table 1.6.1.

Phosphorylation type	Selective enrichment/ monitoring strategy	Starting protein amounts	Advantages	Disadvantages
Tyrosine	Immunoaffinity (anti-pTyr antibodies)	20mg–40mg	Selective enrichment of hundreds of tyrosine phosphorylated peptides detectable by high resolution and high sensitivity mass spectrometers	Requirement of more starting material limits its utility in studying clinical samples like biopsies
Serine/Threonine/ Tyrosine	TiO <sub>2</sub> / IMAC Fe (III), Gallium (III)	2mg–15mg	Unbiased selective enrichment of thousands of phosphorylation sites detectable by high resolution and high sensitivity mass spectrometers	Under representation of tyrosine phospho sites.
Serine/Threonine/ Tyrosine	SCX Fractionation	2mg-20mg	Capable of enriching thousands of phosphopeptides detectable by high resolution and high sensitivity mass spectrometers	Technical expertise and experience required to carry out efficient enrichment. Starting material requirement limits its utility. Usually coupled with TiO <sub>2</sub> / IMAC enrichment
Serine/Threonine/ Tyrosine	Antibody microarrays	micro gram amounts	Low sample requirement	A prior knowledge of target proteins to be assayed. Specific antibodies and biochemical optimization required

**Table 1.6.1** Table summarising the advantages and disadvantages of phosphoprotein / phosphopeptide enrichment strategies used in the analysis of the phosphoproteome. Table adapted from (Harsha and Pandey 2010).

Aside from the low abundance of phosphoproteins and their associated sub-stoichiometric nature, there are other challenges that also exist in the analysis of phosphoproteomes. The rapid phosphorylation and/or dephosphorylation of a protein that may occur as a result of a signalling event can be difficult to capture (Blagoev et al. 2004; Nita-Lazar, Saito-Benz, White 2008).

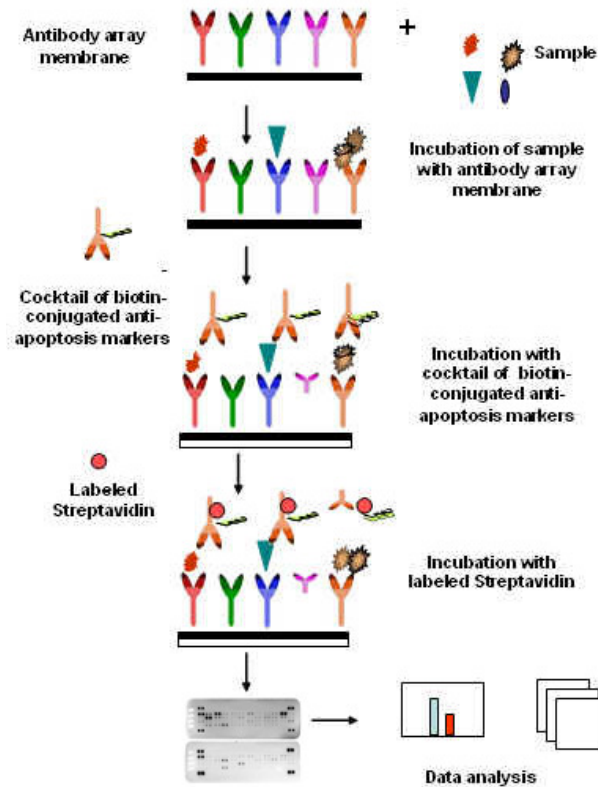
Another issue in 'large scale' studies involves the enzymatic digestion of the sample; although trypsin is widely regarded as the enzyme of choice for protein digestion prior to analysis by LC/MS, it does suffer from certain drawbacks. This includes having little or no activity toward arginine-proline and lysine-proline bonds and difficulty cleaving residues adjacent to amino acids containing bulky post-translational modifications or to highly acidic residues (Tran et al. 2010). While a number of strategies have been developed to improve the enzymatic digestion of a complex protein sample such as (trypsin-trypsin) double digestion (Tran et al. 2010), the use of multiple proteases (trypsin and lysine) (Brownridge and Beynon 2011) and trypsin enhancing reagents such as ProteaseMax (Orsburn, Stockwin, Newton 2011), obtaining complete digestion of a proteome is still difficult to achieve.

A difficulty that can also be encountered is the assignment of a phosphopeptide to the correct protein isoform (Engholm-Keller and Larsen 2013). As peptide sequences can be shared between isoforms, accurate identification of a particular isoform relies on the detection of a peptide unique to that protein (Stastna and Van Eyk 2012). Given that phosphoproteomic analysis often relies on the identification of single phosphopeptides, matching this to the correct isoform is fraught with difficulty and often requires a second approach (such as western blot) to confidently identify the phosphorylated isoform of interest.

### **1.6.1 Protein arrays to study phosphorylation**

Arrays offer a means of simultaneously detecting 10's to 100's of proteins in parallel in a single experiment on a miniaturised, and sometimes automated, platform (HS Lu et al. 2012). Arrays contain a capture protein immobilized onto a surface such as glass or nitrocellulose. This enables the detection of a protein of interest from a complex sample such as a whole cell lysate. The protein may then be detected either through the use of a secondary antibody designed to bind to the protein of interest as it is retained by the capture antibody. Alternatively, the starting material may be labelled, thus indicating the presence proteins retained by the capture protein after all non-binding starting material has been removed following a wash step. An overview of the process workflow is provided in Figure 1.6.1.

Protein arrays have been used for a wide variety of investigations such as proteomic profiling, protein-protein interaction, enzyme activity and antibody screening (Gordus and MacBeath 2006; Jones et al. 2005; Ptacek et al. 2005; Zhu et al. 2006). It has also been used in the in the study of post translational modification such as phosphorylation (Lin et al. 2010). For example, one investigation aimed at identifying all the potential protein substrates of each yeast kinase detected 4129 phosphorylation events in 1325 different proteins (Ptacek et al. 2005). The phosphokinase array utilized in this thesis has also been successfully employed in a number of studies including the identification of activated Erk1/2 and Akt as potential targets for the multi-targeted kinase inhibitors, sorafenib and sunitinib in the treatment of cholangiocarcinoma (Dokduang et al. 2013). Similarly, stimulation of MAPK, AKT and FAK signalling pathways following the activation of the tyrosine receptor kinases Mer and Axl as identified through the use of a phosphokinase array suggested that Mer and Axl should be investigated further as potential therapeutic targets in the treatment of non-small cell lung cancer (Linger et al. 2013). Protein arrays are not without their problems however; antibody cross-reactivity resulting in false positives or false negatives being one of the main ones. For this reason it is important that results are validated using a complementary method such as immunoblot or ELISA (Zhang and Pelech 2012). Despite this, the ability of kinase arrays to screen large elements of the proteome in a targeted manner means that they look set to remain a popular choice for the foreseeable future (Stoevesandt, Taussig, He 2009).



**Figure 1.6.1** Overview of the process workflow for a protein array (Image obtained from <http://www.baria.cz/apoptosis-array-0>).

### 1.6.2 Staining of phosphoproteins for gel based approaches

Of course phosphoproteins enriched by an IMAC or immunoprecipitation based enrichment do not have to be analysed by LC/MS, they may instead be visualized by 2D-gel electrophoresis. In such instance, changes in phosphorylation can be detected on a 2D-gel through the use of a phosphorylation specific stain such as Pro-Q Diamond (Schulenberg et al. 2003); however this will not in itself indicate the specific site of phosphorylation. A second methodology such as western blotting using site-specific antibodies or mass spectrometry is frequently employed to provide this information. There are of course limitations with both of these follow up approaches: the use of western blotting requires prior knowledge of the protein present in order to select a suitable antibody for its detection.

Reliance on the use of mass spectrometry to identify the protein could be complicated by the presence of other proteins that may migrate to the same point on the gel, thus leading to incorrect protein identification (D'Ambrosio et al. 2007). Nonetheless, Pro-Q diamond has been utilized in the study of the phosphoproteomes of a wide variety of species including *Bacillus subtilis* and *Arabidopsis* (Chitteti and Peng 2007; Eymann et al. 2007) as well as neuronal and breast cell lines (Kang et al. 2007; Zhang, Wong, Koay 2007).

### **1.6.3 Chromatography for the enrichment of phosphoproteins and phosphopeptides**

A popular means of enriching for phosphorylation at the protein and at the peptide level is through the use of chromatography columns that have an affinity for the phosphate group. This section provides a brief overview of some of these techniques.

Broadly speaking, metal affinity columns for phosphoprotein/peptide enrichment fall into one of two categories, immobilized metal affinity column (IMAC) or metal oxide affinity column (MOAC). The main difference between the two being that the oxidised form of a metal is used for MOAC, while the unoxidised form is used for IMAC. In both instances phosphopeptide enrichment works on the basis of electrostatic interactions occurring between the phosphate group on the peptide and the metal oxide sorbent (Engholm-Keller and Larsen 2013). It should be noted however that the precise chemistry for kit based phosphoprotein enrichment kits is generally not revealed by the manufacturer making it difficult to draw inferences from the column used and the results obtained. Nonetheless, the use of such columns has become more popular and has been reported for the enrichment of phosphoproteins in a number of studies. An example of one such study that successfully employed phosphoprotein enrichment investigated the tubular epithelial to mesenchymal cell transition upon induction by transforming growth factor beta-1 (Chen et al. 2010). Enriched phosphoproteins were enzymatically digested and iTRAQ labelled before undergoing 2D-nano-LC/MS resulting in the identification of 38 differentially-expressed phosphoproteins.

An alternative to analysis of the elution fraction by mass spectrometry is to use western blot in conjunction with the phosphoprotein enrichment column to detect changes in the abundance of protein phosphorylation. For example, phosphoprotein enrichment prior to western blot analysis has been used to identify changes in the phosphorylation and / or abundance of the protein under investigation.

This technique is based on the principle that the phosphorylated form of the protein of interest will have been preferentially retained by the enrichment column, while the unphosphorylated form will not bind. Following elution of the phosphorylated protein from the column, this sample is then analysed by immunoblot. If there has been an increase in the abundance of the phosphorylated form of a protein under a given condition, this change in abundance will be detected by an antibody specific to the protein of interest. One advantage of this methodology is that, in the absence of an antibody specific to the phosphorylated form of the protein of interest, antibodies raised against the unphosphorylated form of the protein may be utilized.

Similar to phosphoprotein enrichment, metal affinity chromatography can also be used for the enrichment of phosphorylated peptides. Unlike phosphoprotein enrichment however, phosphopeptide enrichment strategies are almost always referred to by the metal chemistry used in the column. The following sections will discuss two of the more common methodologies; IMAC and Titanium dioxide.

### **1.6.3.1 Titanium dioxide phosphopeptide enrichment**

A MOAC based methodology called titanium dioxide enrichment is perhaps one of the most widely used phosphopeptide enrichment methods due to its high capacity and high specificity for the phosphate group (Eyrich, Sickmann, Zahedi 2011; Larsen et al. 2005). Reports of the phosphopeptide enrichment of samples such as whole cell lysates without any prior enrichment strategies have yielded significant numbers of phosphopeptides.

For example, titanium dioxide was employed in the identification of 937 phosphopeptides in acute myeloid leukaemia cells treated with sodium pervanadate (Montoya et al. 2011). Of course, phosphoprotein and phosphopeptide enrichment techniques can be combined in an effort to obtain further phosphorylation information. One such instance of this is provided in a report by Nilsson et al. who used phosphoprotein enrichment and titanium dioxide phosphopeptide enrichment to measure responses of glioblastoma cancer stem cells (GSC11) to STAT3 phosphorylation inhibition by WP1193 treatment and IL-6 stimulation under normoxic and hypoxic conditions (Nilsson et al. 2009).

One drawback initially associated with the use of titanium dioxide columns is that in addition to the binding of phosphate, they also bind acidic groups with high affinity (Gates, Tomer, Deterding 2010). This was overcome however by the use of organic acids such as 2,5-DHB, as they bind metal oxides with greater affinity than carboxyl groups, but less affinity than phosphate (Aryal and Ross 2010). It is typically considered that titanium dioxide phosphopeptide enrichment columns preferentially bind monophosphorylated species, while IMAC preferentially binds multiply phosphorylated peptides (Aryal, Olson, Ross 2008; Zhao et al. 2013). It has been proposed however that titanium dioxide does not discriminate between multi and mono phosphorylated peptides, but multiple phosphorylated species are difficult to elute due to the high affinity with which they bind (Thingholm et al. 2008). Although the optimisation of experimental conditions for the enrichment of phosphopeptide using titanium dioxide columns is on-going (Aryal and Ross 2010), as yet there is no one set of conditions that can be used to enrich for the entire phosphoproteome (Thingholm et al. 2008). This can be said to be true of all phosphopeptide enrichment strategies. The greatest coverage of a phosphoproteome is typically provided when methods such as titanium dioxide are used in conjunction with other complementary methods such as IMAC (Bodenmiller et al. 2007; Carrascal et al. 2008).



### 1.6.3.2 IMAC for phosphopeptide enrichment

The specificity with which phosphopeptides are enriched using IMAC columns depends on (i) the resin used (ii) the metal ion used. Choosing the optimum enrichment resin / metal ion is not as straightforward as it might appear. For example, a resin of nitrilotriacetic acid (NTA) coupled with an iron ( $\text{Fe}^{3+}$ ) metal ion was found to be more specific for the enrichment of phosphopeptides than a resin containing iminodiacetic acid (IDA) and  $\text{Fe}^{3+}$  (Neville et al. 1997). Conversely, a combination of IDA and gallium ( $\text{Ga}^{3+}$ ) was found to outperform NTA combined with  $\text{Ga}^{3+}$  (Posewitz and Tempst 1999). Despite this, there are a number of other IMAC based metal affinity resins such as alumina, zirconium and gallium that have also been used for the enrichment of phosphopeptides with great success (Posewitz and Tempst 1999; Qiao et al. 2010).

For example, it has been demonstrated that such phosphopeptide enrichment methodologies are capable of obtaining large number of phosphopeptides from samples such as whole cell lysates without any prior enrichment strategies; one report identified 654 phosphopeptides in the identification of hormone dependent signalling pathways in renal medullary thick ascending limb (mTAL) cells using gallium phosphopeptide enrichment (Gunaratne et al. 2010). This is hardly surprising given that it has been shown that gallium is comparable with the likes of titanium, iron, aluminium and zirconium for the enrichment of phosphopeptides (Aryal and Ross 2010; Posewitz and Tempst 1999). Similar to titanium dioxide, IMAC also suffers from the non-specific binding of acidic rich glutamic and aspartic acid residues. One approach to overcome this involves the acidification of samples prior to loading on the IMAC column; this protonates the carboxyl groups on the highly acidic amino acid residues and reduces instances of nonspecific binding (Posewitz and Tempst 1999). A more recent advance in IMAC based phosphopeptide enrichment has been the development of a strategy to sequentially elute phosphopeptides from the IMAC column (Thingholm et al. 2008).

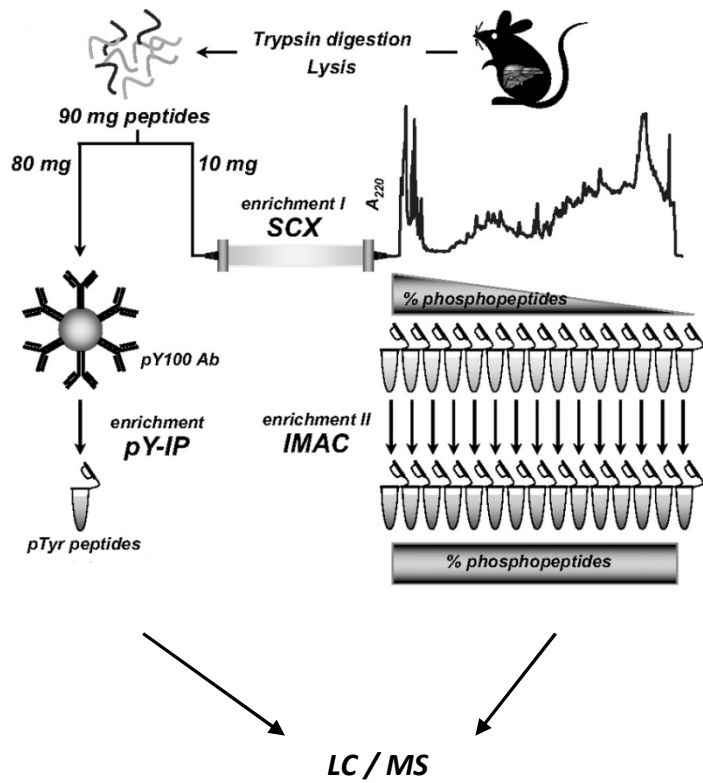
SIMAC (Sequential elution from IMAC) facilitates the elution of monophosphorylated and multi-phosphorylated peptides into distinct fractions enabling analysis of low abundant multi-phosphorylated peptides that otherwise may go undetected due to the relatively high abundance of the monophosphorylated species.

It has been through the continuous refinement of experimental conditions such as this that has made IMAC one of the most popular means of enriching for phosphopeptides (Niklew et al. 2010).

### **1.6.3.3 Strong cation exchange (SCX)**

Strong cation exchange (SCX) has also been used as an additional phosphopeptide enrichment step prior to the use of an IMAC or TiO<sub>2</sub> column (Beausoleil et al. 2004; Villen et al. 2007). SCX operates based on the principle that positively charged proteins or peptides will bind to the negatively charged functional groups on the SCX resin. Because phosphorylated peptides possess an additional negative charge due to the presence of the phosphate group, these will elute from the column earlier than the non-phosphorylated species.

This technique has been used as part of the enrichment protocol in a number of large scale phosphoproteome profiling studies including the identification of 2,002 phosphorylation sites from HeLa cell nuclear fractions, and 6,600 phosphorylation sites in EGF stimulated HeLa cells (Beausoleil et al. 2004; Olsen et al. 2006). SCX has been compiled with both immunoprecipitation and IMAC phosphopeptide enrichment methodologies to identify 5,635 phosphorylation sites in mouse liver (Figure 1.6.2) (Villen et al. 2007). Indeed, the coupling of multiple enrichment methodologies has enabled the identification of over 20,000 phosphorylation sites in some biological systems (Olsen et al. 2010). Although the use of this technique provides enhanced coverage of the phosphoproteome, its implementation requires a high level of technical ability on the part of the user (Harsha and Pandey 2010). In addition to this, large volumes of starting material are also required to ensure optimal recovery of phosphopeptide (Olsen and Macek 2009).



**Figure 1.6.2** Schematic of phosphopeptide enrichment strategies (SCX coupled with immunoprecipitation and IMAC) as used by Villen et al. (2007) in the identification of 5,635 phosphorylation sites in mouse liver.

## **Aims**

### **Phosphoproteomic characterisation of CHO cells subject to temperature-shift**

Our current knowledge of the cellular mechanisms involved in the cell response to reduced culture temperature is limited. However, understanding such mechanisms could aid in the identification of cell engineering targets to extend cellular viability, increase productivity or influence cellular growth rate. This project aims to:

- To assess the feasibility of phospho-array based technology to characterise the phosphoproteome of CHO cells, using antibodies spotted onto a nitrocellulose membrane to detect changes in the relative phosphorylation of kinases in temperature-shifted CHO cells;
- Profile the phosphoproteome of CHO cells subjected to temperature-shift and identify phosphoproteins that are potentially involved in the cellular response to reduced culture temperature.
- Establish and evaluate a protocol for the study of phosphopeptides in CHO cells;
- Utilise mass spectrometry to obtain site-specific information on the differential-phosphorylation of phosphoproteins occurring in direct response to reduced culture temperature conditions.

## **Proteomic profiling of Host Cell Proteins (HCPs) from CHO cells**

The secretome can also be a rich source of information, providing insights into cell-cell signalling events that effect cell growth and viability. Analysis of the host cell protein profile of bioprocess relevant cell lines such as hybridoma and CHO cells could further our understanding of how the cell culture environment can impact on product quality and cell viability. This project aims to:

- Characterise the extracellular proteome of CHO cells over time in culture to identify proteins that potentially impact cellular growth and viability. In addition, identify any proteins that may negatively impact product yield and / or quality;
- Determine whether there are significant differences in the secretome of CHO cells grown under temperature-shifted and non-temperature-shifted conditions;
- Determine whether there are significant differences in the secretome of a recombinant protein-producing and non-producing cell line;
- Catalogue proteins secreted from IgG-producing, phenotypically distinct and fast and slow growing CHO cell lines.

## Materials & Methods

## **2.0 Materials & Method**

### **2.1 Preparation for Cell Culture**

#### **2.1.1 Ultrapure water**

Ultrapure water was used in the preparation of all solutions used for cell culture. Pre-treatment, involving activated carbon, pre-filtration and anti-scaling was first carried out. This water was then purified by a reverse osmosis system (Millipore Milli-RO 10 Plus, Elgastat UHP), which is low in organic salts, organic matter, colloids and bacteria with a standard of 12 – 18 MΩ/cm resistance.

#### **2.1.2 Glassware**

All glassware and lids used for any cell related work were prepared as follows: all glassware and lids were soaked in a 2% (v/v) solution of RBS-25 (AGB scientific, cat. 83460) for at least 1 hour. Glassware was scrubbed and rinsed several times in tap water; the bottles were then washed by machine using Neodisher detergent. The bottles were then rinsed twice with distilled water, once with UHP water and sterilised by autoclaving.

#### **2.1.3 Sterilisation**

Water, glassware and all thermolabile solutions were sterilised by autoclaving at 121°C for 20 mins under 15 p.s.i. pressure. Thermolabile solutions were filtered through a 0.22µm sterile filter (Millipore, cat. SL6V033RB), with low protein binding filters used for all protein-containing solutions.

#### **2.1.4 Subculturing of cell lines suspension cells**

All routine cell culture work was carried out in a class II down flow re-circulating laminar flow cabinet (Nuair Biological Cabinet). Strict aseptic techniques were adhered to at all times (see NICB SOP No. 000-01). Laminar flow cabinets were swabbed with 70% industrial methylated spirits (IMS) prior to and following all work, as well as all equipment used during experiments. Only one cell line was worked with at a time in a cabinet, with 15 minutes clearance time given between work with individual cell lines. Each week cell culture cabinets and any incubators used were cleaned with industrial detergents (Virkon) and IMS.

Cell lines adapted to suspension growth were generally grown in either sterile disposable 250 ml flasks (Corning, cat. 431144) with a 50 ml working media volume or a disposable 50 ml spin-tube (Sartorius, cat. DF-050MB-SSH) with a 5 ml working media volume.

Prior to subculturing cells, a 100µl aliquot was removed from the flask or tube and counted using trypan blue (section 2.1.5). The cell suspension was then transferred to a 30 ml sterile container (Greiner, cat. 201151) and centrifuged at 170xg for 5 mins. The resulting pellet was resuspended in fresh warm growth medium, and an appropriate volume used to seed fresh flasks/tubes for further subculture or planned experiments.

Details of all the cell lines used in experiments are given in Table 2.1.1.

Cell line	Cell line origin	Recombinant protein product	Media	Selective reagents	Culture format	Source
4/2D	Mouse spleen cells fused with mouse myeloma (SP2 ) cells	Non-producing	PHFM-II (Gibco, 12040-077 )	N/A	Suspension (T-175)	In-house (Moran et al. 1996)
CHO K1	CHO K1	Non-producing	OptiCHO (Gibco, 12681-011)	N/A	Suspension (Shake flask)	ATCC
CHO-K1-SEAP	CHO K1	SEAP	SFM-II (Gibco, 12052-098)	G418	Suspension (Shake flask)	In-house (Dr. Niraj Kumar, Thesis)
DP12	CHO K1	IgG	OptiCHO (Gibco, 12681-011)	MTX	Suspension (Shake flask)	ATCC

**Table 2.1.1** Table of cell lines used in experiments presented in this thesis.



### 2.1.5 Assessment of cell number and viability

Cell counting and viability estimations were carried out using a trypan blue dye exclusion technique.

First a sample was taken from a cell suspension and mixed 1:1 with trypan blue (i.e. 100µl cell sample mixed with 100µl trypan blue). This was incubated for 3 mins at room temperature, after which a sample was applied to a haemocytometer over which a glass coverslip is then placed. Cells in the outer four corners of the chamber were counted microscopically using Nikon TS100 microscope and 10x objective lens. Non-viable cells were stained blue, while viable cells remained clear. Non-viable cells were stained blue, while viable cells remained clear. The following formula was used to calculate cell concentration:

$$\left( \frac{\text{Total number of cells counted}}{4} \right) \times (\text{dilution factor}) \times 10^4$$

### 2.1.6 Cryopreservation of cells

To allow for long term storage of cell stocks, cells were cryo-preserved in liquid nitrogen. Dimethyl sulphoxide (DMSO) (Sigma, cat. D8418) was used as a cryoprotectant. Cells to be frozen down were grown until the exponential phase, harvested and counted. Cells were then pelleted by centrifugation at 170xg for 5 mins and resuspended in media to give a concentration of at least 5 x 10<sup>6</sup> cells/ml. An equal volume of 2X concentrated freezing media (DMSO-7.5% and serum 92.5%) was then added dropwise due to the toxicity of the DMSO. The cell suspension was then aliquoted into cryovials (Greiner, cat. 12279) and quickly placed into a freezer at -80°C overnight before being placed into storage in liquid nitrogen.

### 2.1.7 Thawing of cryopreserved cells

Upon removal from liquid nitrogen, the cryovial was thawed in warm water. Once slightly thawed, the contents of the cryovial was transferred to the pre-warmed media (37°C). This suspension was centrifuged at 170xg for 5 mins. The supernatant was then removed and the pellet re-suspended in fresh warm growth medium. This thawed cell suspension was then placed in a 5ml spin tube for suspension growth.

After 24 hours, the medium was then replaced with fresh growth medium to remove any potential traces of DMSO and the cells subcultured if necessary.

### **2.1.8 Sterility checks**

Sterility checks were routinely carried out on all media, supplements and trypsin used for cell culture (see NICB SOP 002-01). A 1ml sample of the material being examined was inoculated into 9mls of both tryptone soya broth (TSB) and thioglycollate broth (Thio) and incubated at 37°C for 7 days. Both samples were then checked for turbidity and change in colour, indicating contamination.

### **2.1.9 Mycoplasma testing**

Cells were regularly checked for the presence of Mycoplasma using the Fluorescent Hoechst stain method. Mycoplasma-negative NRK (Normal rat kidney fibroblast) cells were used as indicator cells for this analysis. NRK cells were incubated with a sample of supernatant from the cell lines being tested for the presence of mycoplasma and then stained. Testing was carried out every four months for each cell line in use.

## **2.2 Protein Quantification**

### **2.2.1 Bradford Assay**

Protein levels were determined using the Bio-Rad Quick Start™ Bradford Dye Reagent (Bio-Rad, 500-0205) as follows. A 2 mg/ml bovine serum albumin (BSA) solution (Sigma, A9543) was prepared freshly in lysis buffer. A protein standard curve (0, 0.125, 0.25, 0.5, 0.75 and 1.0 mg/ml) was prepared from the BSA stock with dilutions made in lysis buffer. An aliquot of the protein samples were diluted as appropriate (1:2, 1:5 or 1:10) with lysis buffer. 5µl of standards and samples were added to a minimum of three separate wells in a 96-well plate. 250µl of the Bio-Rad solution was added to each well. After 5 minutes incubation, absorbance was measured at 595 nm and the concentration of the protein samples was determined from the plot of the absorbance at 595 nm versus concentration of the protein standard.

### **2.2.2 ELISA Assay**

An ELISA (Enzyme-linked immunosorbent assay) was used to quantify the recombinant Human IgG1 (rHuIgG1) antibodies secreted by the DP12 cell line. Conditioned media samples were centrifuged at 170xg for 5min and supernatant was aliquoted into an eppendorf tube. Supernatant was then centrifuged a second time at 19,000xg for 5 min to remove any residual cells and aliquoted into a fresh eppendorf tube. 1µl of coating antibody was diluted in 100µl coating buffer (0.05M Carbonate Bicarbonate, pH 9.60) for well required (number of samples plus number of standards). Coated wells were incubated at 4°C overnight. After incubation, the coating buffer was removed and the plate was washed 5 times with wash buffer (50mM Tris, 0.14M NaCl, 0.05% Tween20, pH 8.0). 200µl of blocking solution (50mM Tris, 0.14M NaCl, 1% BSA, pH 8.0) was then added to each well and incubated for 1 hour at RT. During this incubation period samples and standards were prepared in sample diluent (50mM Tris, 0.14M NaCl, 1% BSA, 0.05% Tween20). Blocking solution was then removed and the plate was washed 5 times with wash buffer. 100µl sample/standard was then added to their assigned wells and incubated for 1 hour at RT.

After incubation, the samples were removed and the plate washed 5 times with wash buffer. The enzyme (HRP) linked detection antibody was diluted in sample diluent at a dilution of 1:100,000 and 100µl added to each well and incubated for 1 hour at RT. After an hour the coating buffer was removed and the plate was washed 5 times with wash buffer. 100µl of HRP substrate solution (TMB - 3,3',5,5'-Tetramethylbenzidine) was then added to each well and incubated in the dark for 15 mins. 100µl of stop solution (0.18M H<sub>2</sub>SO<sub>4</sub>) was then added to each well. Absorbance of each well was read at a wavelength of 450nm.

### **2.2.3 SEAP Assay**

Secreted Alkaline Phosphatase (SEAP) is a secreted enzyme that was used as a model recombinant glycoprotein (Lipscomb et al. 2005). CHO SEAP cells were generated in house by Dr. Niraj Kumar (Ph.D Thesis 2008). The enzymatic assay for quantification of SEAP protein was adapted from the method reported previously by Berger et al. (Berger et al. 1988). The following protocol was used to detect SEAP levels in the supernatant of CHO cells transfected with a SEAP containing plasmid.

Conditioned medium was collected from CHO SEAP cells and centrifuged at 19,000xg for 5min. 500µl aliquots of conditioned media were collected and stored at -20°C until such time as the assay was ready to be performed. 50µl of cell-free conditioned medium was transferred to individual wells of a 96-well flat bottom plate. 50µl of 2x SEAP reaction buffer (solution stock contains 10.50 g diethanolamine (100%), 50 µl of 1 M MgCl<sub>2</sub>, and 226 mg of L-homoarginine in a total volume of 50 mL) was added to each sample. Plates were incubated for 10 min at 37°C and then 10 µl of substrate solution (158 mg of p-nitrophenolphosphate from Sigma (P4744) in 5 mL of 1x SEAP reaction buffer, made fresh for each use) was then added to each well. The change in absorbance per minute (OD 405/min) of each well was monitored by a microplate reader. The change in absorbance per minute was considered as indicator of the amount of SEAP present in the sample.

## **2.3 Proteomic Sample Collection and Preparation**

### **2.3.1 Cell Lysate Sample Preparation**

Cells were pelleted by centrifugation at 170xg for 5 minutes and the media was removed and cells were washed with ice cold 50mM HEPES. The cell suspension was then transferred to an eppendorf and centrifuged at 14000xg for one minute at 4°C. HEPES was then decanted, the eppendorf containing the cell pellet was submersed in liquid nitrogen for 30 seconds before being stored at -80 C° until required. All procedures from this point forward were performed on ice. Cells were lysed using an appropriate volume of lysis buffer (7M Urea, 2M Thiourea, 4% CHAPS, 30mM Tris, pH 8.5) and incubated on ice for 20 minutes with occasional vortexing. Protease inhibitor (PI-78415 Thermo Scientific), halt phosphatase inhibitor (78428 Thermo Scientific) and Nuclease mix (80-6501-42 GE) were also added to the lysis buffer to form a final concentration of 1x for each inhibitor. Cells were lysed by pipetting up and down and by vortexing. Sample lysates were centrifuged at 19,000xg for 15 minutes at 4 C°. Supernatant containing extracted protein was transferred to a fresh chilled eppendorf tube and stored in aliquots at -80 °C if they were not going to be used immediately.

### **2.3.2 Conditioned Media Collection**

Media containing cells were placed into a 50ml sterile universal container and centrifuged at 2000xg for 6 minutes. Upon completion the media was decanted into a fresh sterilin and centrifuged at 2000xg for a further 6 minutes. The media was then decanted into a fresh sterilin and filtered using a 0.45µm filter (VWR, 514-0075), one filter per 25ml of media. Upon completion, the media underwent further filtration through a 0.22µm filter (VWR, 514-0073), one filter per 25ml of media. Filtered conditioned media was then either stored at -80°C for long term storage or kept on ice for immediate further processing.

### **2.3.3 Conditioned Media Sample Preparation**

Conditioned media sample was concentrated by centrifugation at 3,200xg at 16°C using Vivaspin 20 3,000 MWCO centrifugal concentrators (Sartorius VS2092) until samples had decreased in volume ~100-fold. 50-100µg of protein from the sample of interest was placed in an eppendorf and underwent a 'clean up' using ReadyPrep™ 2-D Cleanup Kit (Bio-rad, 163-2130) as detailed 2.3.4.

The protein pellet was then resuspended in 6M Urea, 2M Thiourea, 10mM Tris, pH 8 (Label-free lysis buffer). The sample was sonicated and vortexed until the precipitate was no longer visible and quantified by Bradford assay.

### **2.3.4 Sample clean-up for Mass Spectral Analysis**

80µl of conditioned media (≈25µg protein) or 100µl of whole cell lysate (≈500µg protein) was cleaned up in order to remove any agents that could interfere with Mass Spectral Analysis. Sample clean-up was carried out using ReadyPrep 2D Cleanup Kit (BioRad, 163-2130) according to manufacturer's instructions with the exception of the final incubation period where the sample was incubated at -20°C overnight as opposed to 30 minutes. Following the final centrifugation step, the supernatant was removed and the pellet was allowed to air dry for 1 minute before being resuspended in the appropriate buffer depending on whether the sample was going to be analysed by LC/MS or western blot.

### **2.3.5 In solution digestion for preparation of conditioned media samples**

10µg of protein was placed into a new eppendorf. All samples were then brought up to an equal volume using label-free lysis buffer. 1µl of reduction buffer was added to

each sample (DTT in 50mM ammonium bicarbonate such that the final concentration of DTT will be 5mM). This was incubated at 56°C for 20 minutes on a shaker platform. Once the reduced samples had cooled, 1µl of alkylation buffer was added to the sample (Iodoacetamide in 50mM Ammonium Bicarbonate such that the final concentration of Iodoacetamide is 55mM), and incubated for 20 minutes at room temperature in the dark. 0.2µg of Lys-C (Promega, V1071) was then added to each sample and left on a shaking platform for 4 hours at 37°C. After this incubation period, each sample was diluted four volumes of 50mM Ammonium Bicarbonate.

0.4µg of sequence grade Trypsin (Promega, V5280) was then added and all samples were incubated overnight at 37°C on a shaker platform. After overnight digestion PepClean™ C-18 Spin Columns (Pierce, 89870) were used to purify and concentrate the resulting peptides as detailed in 2.3.7.

### **2.3.6 In solution digestion for preparation of whole cell lysate samples**

100µl of whole cell lysate containing 500µg of protein was cleaned up in order to remove any agents that could interfere with mass spectrometry analysis. Sample clean-up was carried out using ReadyPrep 2D Cleanup Kit (BioRad) as detailed 2.3.4. To prepare protein samples for digestion, samples were resuspended in label-free lysis buffer (6M Urea, 2M Thiourea, 10mM Tris, pH 8) and all volumes were brought up to 70µl and 20µl of 50mM ammonium bicarbonate added. The samples were first incubated with 10µl 50mM dithiothreitol for 30 minutes at 37°C. 11µl of 150mM iodoacetamide was then added to each sample and incubated in the dark for 20 minutes at room temperature. 314µl of 50mM ammonium bicarbonate was then added to each sample. Digestion of protein samples was performed with trypsin (sequence grade, Pierce) at a ratio of 31.25/1, protein/enzyme w/w, i.e. 500µg of protein was digested with 16µg of trypsin. 4.5µl of ProteaseMax (Promega V2071). Samples were incubated overnight at 37°C to allow the digestion to occur.

### **2.3.7 Peptide sample clean-up using C18 spin column**

With the exception of two additional wash steps (four washes total) and two additional elutions (three elutions total) C18 sample clean-up was carried out as per the manufacturer's instruction using PepClean™ C-18 Spin Columns (Pierce, 89870).

Samples were then placed in a vacuum evaporator until they had dried fully and re-suspended in an appropriate volume of 0.1% TFA in 2% ACN with vortexing and sonication.

### **2.3.8 Peptide sample clean-up using Graphite spin column**

Graphite spin columns (Pierce, 88302) were used to clean up samples containing >30µg of peptide. As well as concentrating the samples, this was done to remove any agent that may interfere with Mass Spectral analysis. All samples were processed as per the manufacturer's instructions such that the final elution volume was 400µl.

Samples were then placed in a vacuum evaporator until they had dried fully and re-suspended in an appropriate solution/volume for the next application e.g. titanium phosphopeptide enrichment (2.4.4), iron phosphopeptide enrichment (2.4.5), LC-MS analysis (2.4.6).

## **2.4 Proteomic Enrichment and Depletion Methods**

### **2.4.1 Phosphoprotein enrichment Pierce**

Enrichment for Phosphoproteins took place using the Pierce Phosphoprotein Enrichment Kit (Pierce, 90003). Cell pellets were lysed using the Lysis/Binding/Wash buffer provided which was supplemented with CHAPS (0.25%) (Sigma, C9426), 1X Halt Protease Inhibitor (EDTA-free) (Thermo Scientific, 78425), 1X Halt Phosphatase Inhibitor Cocktail (Thermo Scientific, 78420) and 1X Nuclease Mix (GE Healthcare, 80-6501-42). Upon addition of cell lysis buffer to the cell pellet, cells were passed through an 18-gauge needle 5 five times and then vortexed.

Samples were placed on a rotator (Stuart, SB2) and left spinning at 15rpm for 45 minutes at 4°C with periodic vortexing during this period. The lysate was then centrifuged at 14,000xg for 20 minutes at 4°C and the supernatant collected. After protein quantification by Bradford assay, phosphoprotein enrichment and concentration was carried out as per the manufacturer's instructions such that the final volume of samples was 100-150µl. At this point samples were either stored at -80°C or kept on ice for further processing.

#### **2.4.2 Phosphoprotein-enriched Label-free sample preparation**

Phosphoprotein-enriched sample fractions from the Pierce Phosphoprotein enrichment kit (Pierce, 90003) were quantified by Bradford assay and 120µg of protein was placed into a new eppendorf. All samples were then brought up to 50µl using the elution buffer supplied in the Pierce Phosphoprotein enrichment kit. This volume was then brought up to 93.5µl using 50mM of Ammonium Bicarbonate. 1µl of reduction buffer was added to each sample (0.5M DTT in 50mM ammonium bicarbonate), vortexed, briefly centrifuged and incubated at 56°C for 20 minutes on a shaker platform. Once the reduced samples had cooled, 2.7µl of alkylation buffer was added to the sample (0.55M Iodoacetamide in 50mM Ammonium Bicarbonate). The sample was then briefly vortexed and centrifuged before being incubated for 20 minutes at room temperature in the dark. After this incubation period, 1µl of 1% ProteaseMAX™ surfactant trypsin enhancer (Promega, V2071) and 2.2µl of 2µg/µl Trypsin was then added and all samples were incubated overnight at 37°C on a shaker platform. A volume of each sample equal to 20µg of peptide was then taken and stored at -20°C for quantitative label-free LC/MS analysis. The remaining 100µg of peptide was processed using Graphite spin columns (Pierce, 88302) and Phosphopeptide enriched using Swellgel® Gallium spin columns (Pierce, 89853).

#### **2.4.3 Gallium Phosphopeptide enrichment**

Phosphopeptide enrichment using SwellGel® Gallium chelated discs was carried out as per the manufacturer's instructions (Pierce, 89853).

#### **2.4.4 TiO<sub>2</sub> Phosphopeptide enrichment**

Following clean up using graphite spin columns (detailed 2.3.8) peptides obtained from CHO whole cell lysates for TiO<sub>2</sub> phosphopeptide enrichment were enriched using TiO<sub>2</sub> phosphopeptide enrichment columns (Pierce 88303) as per the manufacturer's instructions. Each sample was then cleaned up using graphite spin column as detailed 2.3.8.

#### **2.4.5 Iron Phosphopeptide enrichment**

Following clean up using graphite spin columns (detailed 2.3.8) peptides obtained from CHO whole cell lysates for iron phosphopeptide enrichment were enriched using iron phosphopeptide enrichment columns (Pierce 88300) as per the



manufacturer's instructions such that the elution fractions were pooled. Samples were then cleaned up using graphite spin columns as per 2.3.8.

#### **2.4.6 LC/MS analysis**

Samples were resolubilised in 25 $\mu$ l of LC-MS grade water with 0.1% TFA and 2% ACN. 12.5 $\mu$ l was then transferred to glass vials and the remaining reconstituted sample stored at -80°C. Mass spectrometry was performed by Mr. Michael Henry. Nano LC-MS/MS analysis was carried out using an Ultimate 3000 nanoLC system (Dionex) coupled to a hybrid linear ion trap/Orbitrap mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific). Note, SilicaTip™ Standard Coating Tubing OD/ID 360/20  $\mu$ m Tip ID 10  $\mu$ m Length 5cm, (NewObjective) were used as emitter tips for nano electrospray.

A 2.5 $\mu$ l injection of digested sample were picked up using an Ultimate 3000 nanoLC system (Dionex) autosampler using direct injection pickup onto a 20 microlitre injection loop. The sample was loaded onto a C18 trap column (C18 PepMap, 300  $\mu$ m ID  $\times$  5 mm, 5  $\mu$ m particle size, 100 Å pore size; Dionex) and desalted for 10 min using a flow rate of 25 $\mu$ l/min in 0.1% TFA containing 2% acetonitrile (Loading Buffer). The trap column was then switched online with the analytical column (PepMap C18, 75  $\mu$ m ID  $\times$  250 mm, 3  $\mu$ m particle and 100 Å pore size; (Dionex)) using a column oven at 35 °C and peptides were eluted with the following binary gradients of:

Mobile Phase Buffer A and Mobile phase buffer B: 0–25% solvent B in 120 min and 25–50% solvent B in a further 60 min, where solvent A consisted of 2% acetonitrile (ACN) and 0.1% formic acid in water and solvent B consisted of 80% ACN and 0.08% formic acid in water. Column flow rate was set to 350 nL/min. Data were acquired with Xcalibur software, version 2.0.7 (Thermo Fisher Scientific).

The Hybrid linear ion trap/Orbitrap mass spectrometer (LTQ Orbitrap XL; Thermo Fisher Scientific) was operated in data-dependent mode and externally calibrated. Survey MS scans were acquired in the Orbitrap in the 400–1800 m/z range with the resolution set to a value of 30,000 at m/z 400. Up to three of the most intense ions (1+, 2+ and 3+) per scan were CID fragmented in the linear ion trap followed by an MS3 event.

MS3 was only triggered if a neutral loss of phosphoric acid was detected indicated by a -98, -49 or -32.7 Da loss from the parent precursor. Dynamic exclusion was enabled with a repeat count of 1, repeat duration of 30 seconds, exclusion list size of 500 and exclusion duration of 40 seconds. The minimum signal was set to 500. All tandem mass spectra were collected using a normalised collision energy of 35%, an isolation window of 2 m/z, activation Q was set to 0.250 with an activation time of 30.

#### **2.4.7 Mass Spectrometry Protein Identifications**

Mass spectrometry data is generated as \*.RAW files from LTQ XL instruments. Data was analysed using the search algorithms TurboSequest (Thermo Fisher Scientific) and MASCOT (v2.3.01, Matrix Science, London, UK) through Proteome Discover 1.3 (Thermo Fisher Scientific) against Bielefeld-BOKU-CHO database (Meleady et al 2011) and CHO-K1 genomic data as published by Xu et al. (2011) for all LC/MS analysis of samples originating from CHO cultures. For samples originating from Hybridoma cultures, Swissprot-uniprot Mouse databases (fast file Jan 2014) were used.

The search parameters that were used allowed two missed cleavages, fixed modification of cysteine (carbamidomethyl-cysteine), variable modifications of methionine (oxidised), serine, threonine and tyrosine (phosphorylated) and serine and threonine (dehydrated). The Hybrid linear ion trap/Orbitrap mass spectrometer used a peptide tolerance of 20ppm and the MS2 and MS3 tolerances were set at 0.5Da.

On completion of the database search the peptide results were filtered as follows for host cell protein analysis: a MASCOT criteria of 95% confidence interval (C.I.) threshold ( $p < 0.05$ ), with a minimum MASCOT score of  $\geq 40$  while the following TurboSequest filters were applied: for charge state 1, XCorr  $> 1.9$ ; for charge state 2, XCorr  $> 2.2$ ; for charge state 3, XCorr  $> 3.75$  and a delta CN of 0.1.

For phosphoproteomic analysis the following filters were applied: a MASCOT criteria of 95% confidence interval (C.I.) threshold ( $p < 0.05$ ), with a minimum MASCOT score of  $\geq 40$  while the following TurboSequest filters were applied: for charge state 1, XCorr  $> 1.9$ ; for charge state 2, XCorr  $> 2.2$ ; for charge state 3, XCorr  $> 3.75$ . In addition, PhosphoRS filters were set at  $> 99\%$ .

#### **2.4.8 Label-free LC-MS data analysis**

Data analysis was performed using Progenesis LC-MS software available from Non-Linear Dynamics. Raw MS data files were imported to Progenesis LC-MS software package.

Where specified reference run alignment was carried out twice, once using the automatic reference run selection feature on the software and once by means of manual reference run selection, otherwise, reference run selection was carried out only once by manual reference run selection. A reference run is selected as the sample which is most representative of the data, all additional sample runs are then aligned to the reference sample run.

Alignment of sample runs allows correction of variable peptide retention times during chromatographic separation thus allowing comparison of the different sample runs.

(In instances where reference run alignment was carried out twice, lists of differentially-expressed proteins or phosphopeptides was generated by overlapping the list of differentially-expressed proteins obtained from the automatic reference run selection with the list of differentially-expressed proteins obtained from the manual reference run selection. Proteins / phosphopeptides common to both lists were accepted as being differentially-expressed.) Upon alignment of all runs, detected features were then filtered based on an anova p-value of less than 0.05 between experimental groups. From this list of filtered features a principal component analysis (PCA) plot is generated. The MS/MS data from this list of filtered features is then exported into the external search engine MASCOT to match these identified features to known peptides as detailed in section 2.4.7.

XML files containing MS/MS identifications were then imported back into Progenesis. Conflicting peptides i.e. peptides common to two or more proteins, were then sorted; the conflicting peptide with the highest ion score was retained while all other peptides for that conflict were removed. In instances where two or more conflicting peptides shared a common 'highest' ion score, these peptides were retained for identification purposes but not used for quantitation. Statistical criteria of anova p-value less than 0.05 and fold-change greater than 1.5-fold between experimental groups was then applied to the protein lists.

## **2.5 Other Proteomic Techniques**

### **2.5.1 Gel electrophoresis**

Proteins for analysis by Western blotting were resolved by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using precast 4-12% Bis-Tris gels (Life Technologies, NP0322BOX (12 well), NP0321BOX (10 well)). Prior to samples being loaded in the relevant sample wells, 8-20 µg of protein was diluted in 2x laemmli loading buffer (Sigma, S3401-1VL). Pageruler™ Plus prestained protein ladder (Pierce, 26619) were loaded alongside samples. The gels were run at constant voltage (200V) until the Bromophenol blue dye front reached the end of the gel, at which time sufficient resolution of the molecular weight markers was achieved.

### **2.5.2 Western blotting**

Once electrophoresis was complete, the SDS-PAGE gel was equilibrated in transfer buffer (10% 10X Tris/Glycine (Biorad, 161-0734), 20% Methanol (Romil, H409), 70% diH<sub>2</sub>O) for approximately 10 minutes. Six sheets of 3 mm filter paper (Whatman, 1001-824) were soaked in freshly prepared transfer buffer and placed on the cathode plate of a semi-dry blotting apparatus (Bio-Rad, TransBlotR) after excess buffer was removed. Hybond™-P PVDF membrane (GE Healthcare, RPN 303F), which had been equilibrated in the same transfer buffer, was placed over the filter paper on the cathode plate. The gels were then aligned on to the membrane. Six additional sheets of transfer buffer soaked filter paper were placed on top of the gel and all air pockets removed. The anode was carefully laid on top of the stack and the proteins were transferred from the gel to the membrane at a current of 0.24A at 15 V for 35 minutes.

The membranes were then blocked for 2 hours at room temperature using either 5% BSA in 0.1% Tris Buffered Saline with Tween 20 (TBSt) for phosphorylated proteins or 5% skimmed milk powder (Marvel) in 0.1% TBSt for non-phosphorylated proteins. Membranes were incubated with primary antibody overnight at 4°C . All antibodies were prepared in 0.1% TBST. Primary antibody was removed after overnight incubation and the membranes rinsed 3 times with 0.5% TBSt for a total of 15-30 minutes. Secondary antibody was added for 30-90 minutes at room temperature. The membranes were washed thoroughly in 0.5% TBSt for 15-30 minutes.

Antibody	Phospho Site	Dilution	Supplier	Product Number	Secondary
Akt	NA	1:3000	Cell Signaling	9272	Rabbit
Akt1 Phos	Ser473	1:1000	Cell Signaling	9271	Rabbit
Erk	NA	1:1000	R&D Systems	AF1575	Rabbit
Erk Phos	Y204+Y187	1:1000	Abcam	Ab76299	Rabbit
Amp Kinase $\alpha$ 1	NA	1:1000	Cell Signaling	2795	Rabbit
Amp Kinase $\alpha$ 2	NA	1:1000	Cell Signaling	2757	Rabbit
Amp Kinase $\beta$ 1	NA	1:1000	Cell Signaling	4178	Rabbit
Amp Kinase $\beta$ 2	NA	1:1000	Cell Signaling	4148	Rabbit
Amp Kinase $\gamma$ 1	NA	1:1000	Cell Signaling	4187	Rabbit
Amp Kinase $\gamma$ 2	NA	1:1000	Cell Signaling	2536	Rabbit
Amp Kinase $\gamma$ 3	NA	1:1000	Cell Signaling	2550	Rabbit
Amp Kinase $\alpha$ 1 and $\alpha$ 2 Phos	Thr172	1:1000	Cell Signaling	2531	Rabbit
JNK1	NA	1:1000	R&D Systems	MAB17761	Mouse
JNK Phos	Thr183/Tyr185	1:1000	Cell Signaling	4668	Rabbit
SAP/JNK	NA	1:1000	Cell Signaling	9258	Rabbit
Mek 6 Phos	Ser207	1:1000	Abcam	Ab36722	Rabbit
Mek 6	NA	1:2000	R&D Systems	AF1604	Rabbit
ATF-2	NA	1:1000	Abcam	Ab31483	Rabbit
ATF-2 Phos	Thr51/Thr69	1:1000	Abcam	Ab131106	Rabbit
NDRG1	NA	1:1000	Cell Signaling	5196	Rabbit
NDRG2 Phos	Ser330	1:1000	Cell Signaling	3506	Rabbit
$\alpha$ -Tubulin	NA	1:2000	Abcam	Ab55611	Mouse
HRP Conjugated Goat anti-Human IgG-Fc	NA	1:50000	Bethly Laboratories	A80-104P	N/A
MMP19	NA	1:1000	Abcam	Ab53146	Rabbit
BMP1	NA	1:1000	Abcam	Ab118520	Rabbit
PSAP	NA	1:1000	Abcam	Ab68466	Rabbit
GRP87/Bip	NA	1:1000	Sigma	G8918	Rabbit

**Table 2.5.1** List of primary antibodies used in the experiments used in this thesis including those used to detect phosphorylated and non-phosphorylated versions of the proteins studied.

### 2.5.3 Enhanced chemiluminescence (ECL) detection

Immunoblots were developed using ECL2 Western Blotting Substrate (Pierce, 80196) which facilitated the detection of bound peroxidase-conjugated secondary antibody. Following the final washing membranes were incubated with 2ml of ECL2 reagent, 5:1 (Buffer A : Buffer B), for 1-5 minutes. The membrane was then placed between two sheets of clear acetate and then exposed to autoradiographic film (Kodak, X-OMATS) for various times (from 10 seconds to 30 49 minutes depending on the signal). The exposed autoradiographic film was developed for 1-5 minutes in developer (Kodak, LX-24). The film was then washed in water for 15 seconds and transferred to a fixative (Kodak, FX-40) for 5 minutes. The film was then washed with water for 15 seconds minutes and left to dry at room temperature.

## 2.5.4 MAP Kinase array

Samples were analysed on the Human Phospho-MAPK Array Kit (R&D systems, ARY002B) as per the manufacturer's instructions. Table 2.5.2 shows the kinases detected by the array. Antibody detection was by ECL as detailed section 2.5.3. Exposure time to autoradiographic film was 1 hour.

Target/Control	Alternate Nomenclature	Phosphorylation Site Detected
Positive Control	Control (+)	—
Positive Control	Control (+)	—
Akt1	PKB $\alpha$ , RAC $\alpha$	S473
Akt2	PKB $\beta$ , RAC $\beta$	S474
Akt3	PKB $\gamma$ , RAC $\gamma$	S472
Akt pan	—	S473, S474, S472
CREB	—	S133
ERK1	MAPK3, p44 MAPK	T202/Y204
ERK2	MAPK1, p42 MAPK	T185/Y187
GSK-3 $\alpha$ / $\beta$	GSK3A/GSK3B	S21/S9
GSK-3 $\beta$	GSK3B	S9
HSP27	HSPB1, SRP27	S78/S82
JNK1	MAPK8, SAPK1 $\gamma$	T183/Y185
JNK2	MAPK9, SAPK1 $\alpha$	T183/Y185
JNK3	MAPK10, SAPK1 $\beta$	T221/Y223
JNK pan	—	T183/Y185, T221/Y223
MKK3	MEK3, MAP2K3	S218/T222
MKK6	MEK6, MAP2K6	S207/T211
MSK2	RSK $\beta$ , RPS6KA4	S360
p38 $\alpha$	MAPK14, SAPK2A, CSBP1	T180/Y182
p38 $\beta$	MAPK11, SAPK2B, p38-2	T180/Y182
p38 $\delta$	MAPK13, SAPK4	T180/Y182
p38 $\gamma$	MAPK12, SAPK3, ERK6	T183/Y185
p53	—	S46
p70 S6 Kinase	S6K1, p70 $\alpha$ , RPS6KB1	T421/S424
RSK1	MAPKAPK1 $\alpha$ , RPS6KA1	S380
RSK2	ISPK-1, RPS6KA3	S386
TOR	—	S2448
PBS	Control (-)	—
Positive Control	Control (+)	—

**Table 2.5.2** Shows list of kinases and the specific site of phosphorylation detected by the kinase array.

### **2.5.5 Depletion of IgG from conditioned media collected from DP12 cell line**

Protein A affinity chromatography work was carried out on an Äktaprime plus using a 5ml HiTrap MabSelect SuRe column (GE Healthcare, 11-0034-94). Each sample was processed manually on the Äktaprime plus. It should be noted that all runs used a flow rate of 5ml/min and the pressure alarm was set to 1MPa. Equilibration / wash buffer consisted of 20mM sodium phosphate, 0.15M sodium chloride pH adjusted to 7.2 with HCl solution. Elution buffer was 60M sodium citrate pH adjusted to 3.3 using 1M citric acid. All buffers were made up using LC/MS grade water.

The column was equilibrated using 5 column volumes (CVs) of equilibration/wash buffer. Conditioned media sample was then run through the column and collected in a clean tube. The collected conditioned media sample was then returned to the original sample tube and run through the Protein A column for a second time. This was repeated until the sample had been passed through the column four times. Conditioned media samples were then either stored at -80°C or prepared immediately for subsequent analysis (section 2.3.3). The column was then equilibrated for further depletion by running an 80% solution of elution buffer through the column for 10 CV's followed by a 100% equilibration / wash buffer for 10 CV's.

# Results



### **3.0 Host Cell Protein Analysis**

#### **Chapter overview**

For some time now, engineering of recombinant protein producing cell lines has been the primary focus of the bioprocessing community in an effort to create faster growing, higher producing cell lines. It is only in more recent times that the host cell protein profile has been looked at in greater detail to identify how the cell culture process can impact on downstream purification and product quality. In addition to this however, the secretome can also be a rich source of information, providing insights into cell-cell signalling events that effect cell growth and viability.

The first study conducted in this thesis looked at changes in the relative abundance of proteins between early exponential and early death phase in a hybridoma cell culture. Given the widespread use of temperature-shift in industry, the effect of the reduction in culture temperature over time (and compared to early and late phases at 37°C) were also examined.

In industry, the primary cell line of choice is the CHO cell. It was therefore decided to extend the investigation to include a study of the HCP profile of non-producing CHO cells over time in culture under temperature-shifted and non-temperature-shifted conditions.

To complement the analysis of the non-producing CHO cells, an IgG-secreting CHO-K1 cell line, CHO-DP12, was also selected for analysis. Analysed in a similar manner to the CHO-K1 cells, this allowed direct comparisons to be made between the two CHO cell lines enabling the identification of proteins with the potential to impact on product quality and cell culture viability as well as the confirmation of changes in the HCP profile as a direct result of temperature-shift.

To the best of my knowledge, the work presented in this thesis is the first time that the HCP profiling of a hybridoma and CHO cell lines over time in culture under temperature-shifted and non-temperature-shifted conditions has been reported.

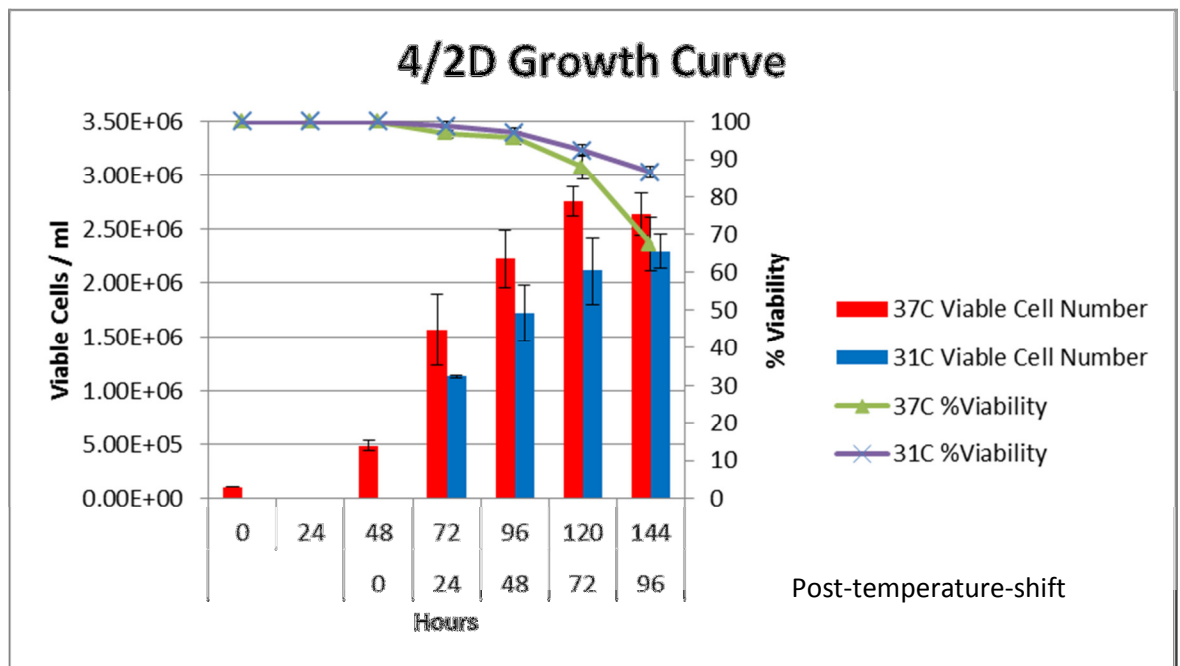
### **3.1 Characterisation of the host cell protein profile of a non-producing hybridoma cell line**

Although there have been a number of publications describing some of the proteins secreted by hybridoma cell lines, these have almost exclusively focused on specific proteases or growth factors. To the best of my knowledge, there have been no reports in the literature detailing the proteomic profiling of host cell proteins produced by hybridoma cells. In order to facilitate the use of a quantitative, label-free, LC/MS method a non-producing hybridoma cell line was selected to generate conditioned media samples, as the presence of an actively secreted IgG molecule, the quantity of which would change both over time and under different culture conditions, would interfere with the accurate quantitation of HCPs. The availability of a well annotated mouse protein database would also facilitate identification of any proteins detected in the cell culture supernatant of the hybridoma cell line. This chapter reports the proteomic analysis of proteins secreted into the cell culture supernatant of a non-producing hybridoma cell line under both temperature-shifted and non-temperature-shifted conditions by quantitative label-free mass spectrometry. In addition, this chapter will also look at the changes in the HCP profile over the life time of the culture from early exponential phase to late stationary phase.

#### **3.1.1 Growth profiling of non-producing hybridoma cell line**

A hybridoma cell line called 4/2D (generated in house by Dr. Annemarie Larkin NICB) was grown in 20ml volume of chemically defined protein free PHFM-II™ media in T-175 monolayer culture. Cells were seeded at  $2 \times 10^5$  cell/ml and after 48 hours, flasks selected for temperature-shift were placed in a 31°C incubator for the remainder of the culture period. Cell culture supernatants were collected 24, 72 and 96 hours post-temperature-shift from both temperature-shift and non-temperature-shifted cultures see Figure 3.1.1 (as per materials and methods 2.3.2). Due to the low volumes of media used, conditioned media from two flasks were pooled to provide a sufficient sample size for analysis. Media from two pooled flasks was considered one replicate. Enough flasks were set up so that three ‘replicates’ from each time point at each temperature could be collected.

Although a lower cell density is placed under temperature-shift by reducing culture temperature at 48hrs, which may be sub-optimal for collecting maximum titre of a recombinant therapeutic, it enabled biological inferences to be drawn from proteomic analysis of two distinct cultures, thus providing insights into changes occurring in temperature-shifted hybridoma cells.



**Figure 3.1.1** Growth curve of a non-producing hybridoma cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions. Temperature-shifted culture was placed at 31°C after 48hrs in culture. Cell culture supernatants were collected 24, 72 and 96 hours post-temperature-shift and analysed using mass spectrometry.

### **3.1.2 Quantitative label-free LC/MS analysis of conditioned media from a non-producing hybridoma cell line**

Conditioned media from three biological replicates taken at 24, 72 and 96 hours post temperature from cultures grown at 37°C and 31°C were prepared for quantitative LC/MS analysis as described in materials and methods 2.3.2 and 2.3.3.

### **3.1.3 Data analysis of differentially-regulated proteins identified by label-free LC/MS profiling of conditioned media from a non-producing hybridoma cell line**

Data acquired following LC/MS analysis was analysed using Progenesis label-free software. A protein was deemed to be differentially-regulated if it has a statistically significant fold-change (anova >0.05) greater than 1.5-fold between experimental groups, and was identified with an ion protein score >40 (this score is a minimum score per peptide and is not based on a cumulative aggregate score between peptides).

Due to the use of multiple time points between two conditions in this experiment, it was decided that the data would be analysed as follows:

- 24hrs vs 96hrs at 37°C
- 24hrs vs 96hrs at 31°C
- 37°C vs 31°C at 24hrs
- 37°C vs 31°C at 96hrs

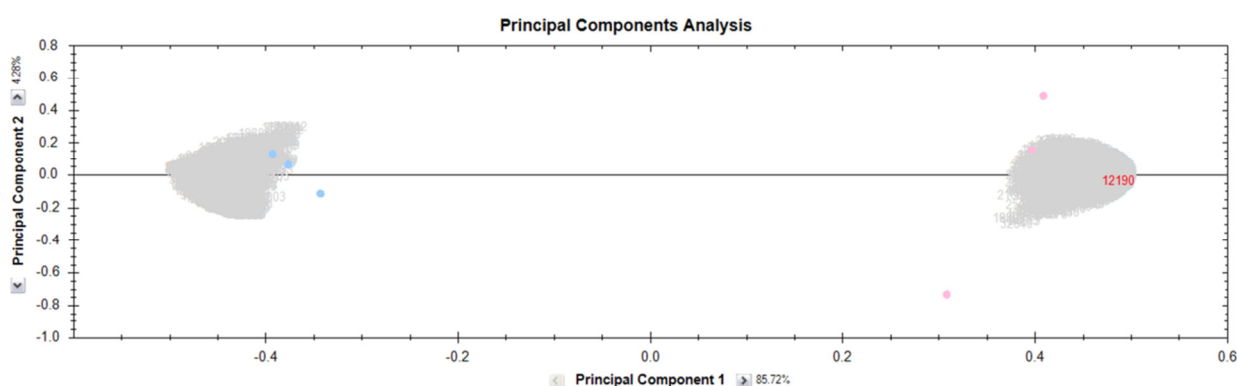
Note: The time in hours above is with respect to temperature-shift. i.e. 24hrs is 24hrs post-temperature-shift.

Analysing the data in this way allows (i) changes in the abundance of proteins between early stage and late stage culture under temperature-shifted and non-temperature-shifted conditions and (ii) changes in the abundance of proteins between temperature-shifted and non-temperature-shifted cultures at early stage and late stage culture, to be captured.

In order to identify what role the differentially-regulated proteins might play in the cell culture supernatant, the gene symbol for the mouse version of the protein identified was obtained and searched in the Uniprot database (<http://www.uniprot.org/>). Proteins classified by uniprot as being a protease, glycosidase, apoptotic protein or miscellaneous protein were deemed to be of particular interest and, as such, were labelled accordingly.

### 3.1.3.1 Differential regulation of HCPs 24hrs vs 96hrs at 37°C

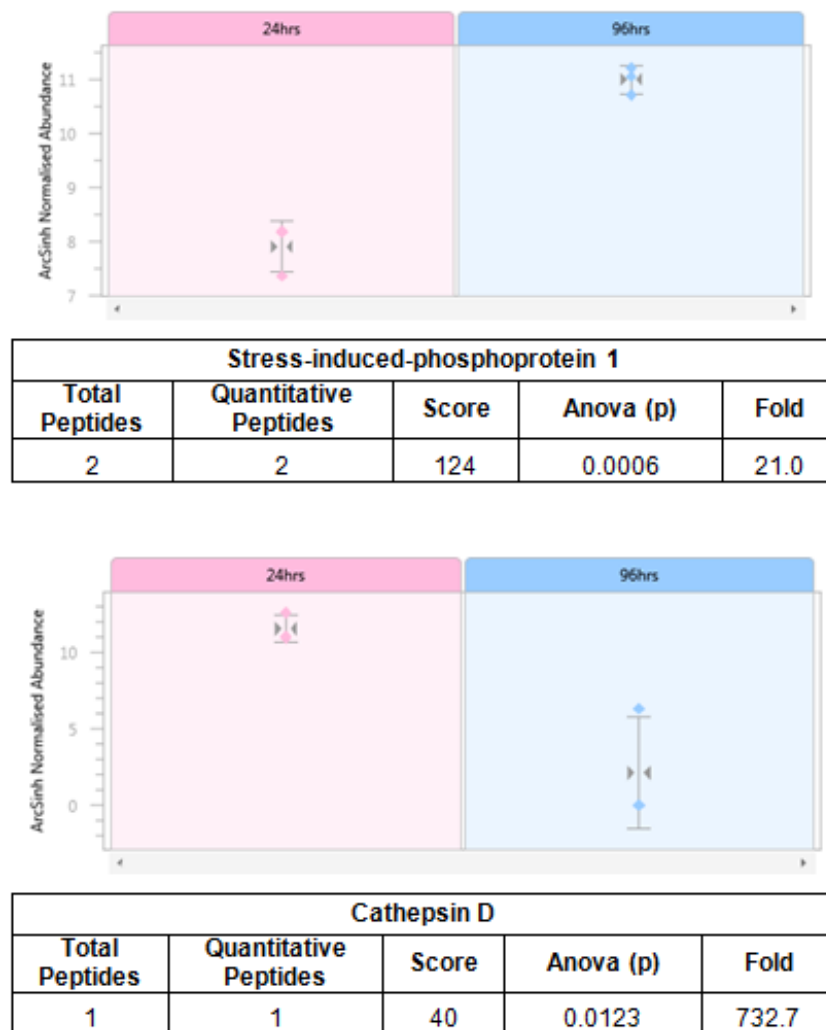
When using the Progenesis label-free software to analyse the LC/MS data, a principal component analysis (PCA) plot was generated showing how the samples from 24hr and 96hr cultures at 37°C clustered. This PCA plot is displayed in Figure 3.1.2.



**Figure 3.1.2** PCA plot for the non-producing hybridoma cell line sample groups 24hrs (pink) and 96hrs (blue) at 37°C generated during quantitative label-free LC/MS data analysis using Progenesis software.

A total of 79 proteins with 1 or more peptides were identified as being differentially-expressed at 37°C between early (24hrs) and late (96hrs). Of these, 25 (with 1 or more peptides) had increased expression at 96hr (Appendix A, Table 3.1.1), while 54 were found to be down-regulated (Appendix A, Table 3.1.2).

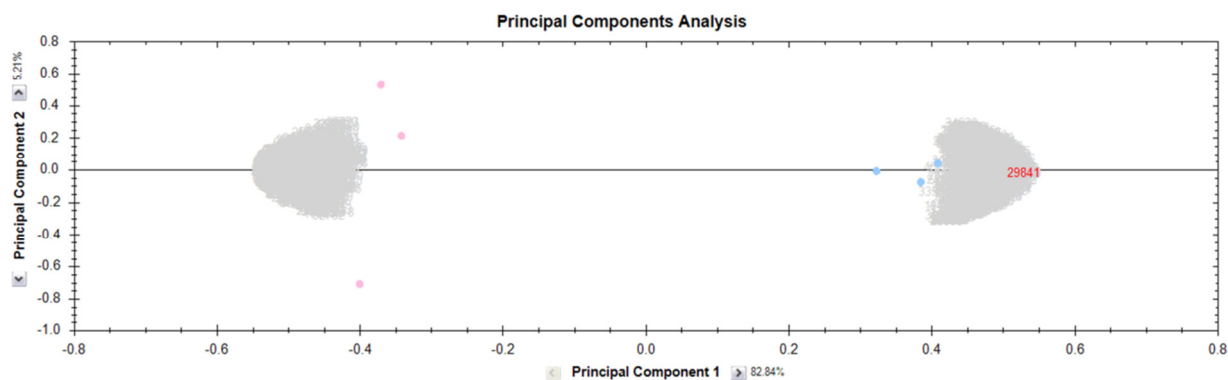
A number of proteins identified in this analysis are of potential interest to the bioprocessing community due to the role they may play in affecting culture viability or product quality. This includes the up-regulation of the stress protein STIP1 by almost 21.0-fold and the down-regulation of the secreted protease Cathepsin D (732.7-fold) at 96hrs compared to 24hrs at 37°C (Figure 3.1.3). The large fold-change observed in Cathepsin D is likely to be indicative of presence / absence scenario, particularly when coupled with the fact that the proteins were identified by one peptide.



**Figure 3.1.3** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the upregulation of Stress-induced-phosphoprotein 1 and the down-regulation of the secreted protease Cathepsin D. Both proteins were identified as being differentially-expressed when 96hrs was compared to 24hrs at 37°C in the conditioned media of a non-producing hybridoma cell line.

### 3.1.3.2 Differential regulation of HCPs 24hrs vs 96hrs at 31°C

A PCA plot showing how the samples from 24hr and 96hr cultures at 31°C cluster was generated by Progenesis label-free software during analysis of LC/MS data and is displayed in Figure 3.1.4.



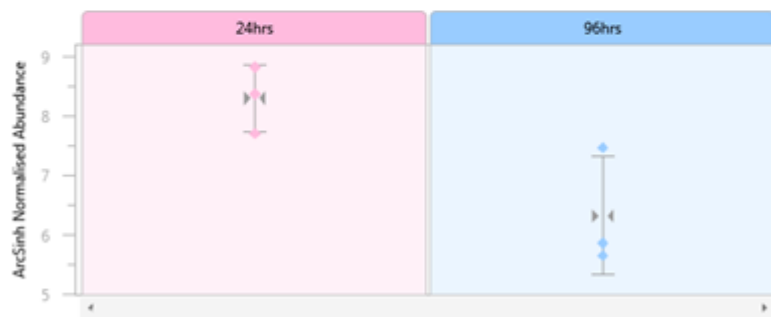
**Figure 3.1.4** PCA plot for non-producing hybridoma cell line sample groups compared at 24hrs (pink) and 96hrs (blue) at 31°C generated during label-free LC/MS data analysis using Progenesis software.

A comparison between early (24hrs) and late (96hrs) phases of the temperature-shifted (31°C) culture found 65 proteins (with 1 or more peptides) to be differentially-expressed. Of these, 31 were up-regulated at 96hrs (Appendix A, Table 3.1.3) and the remaining 34 were down-regulated (Appendix A, Table 3.1.4).

Among the differentially-expressed proteins identified in this study were the secreted proteins Olfactomedin-like protein 3 (2.4-fold up-regulated) and Calumenin (5.6-fold down-regulated) (Figure 3.1.5). Secreted proteins such as these are possibly involved in controlling culture growth and viability through cell-cell signalling events. Despite the relatively high viability of the cells at 31°C, the static nature of the culture i.e., the cells were grown in monolayer suspension format, would suggest that the increase in intracellular proteins is as a result of apoptosis and the subsequent release of proteins from dead cells into the culture media.



Transitional endoplasmic reticulum ATPase				
Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
2	2	116	0.0022	2.0



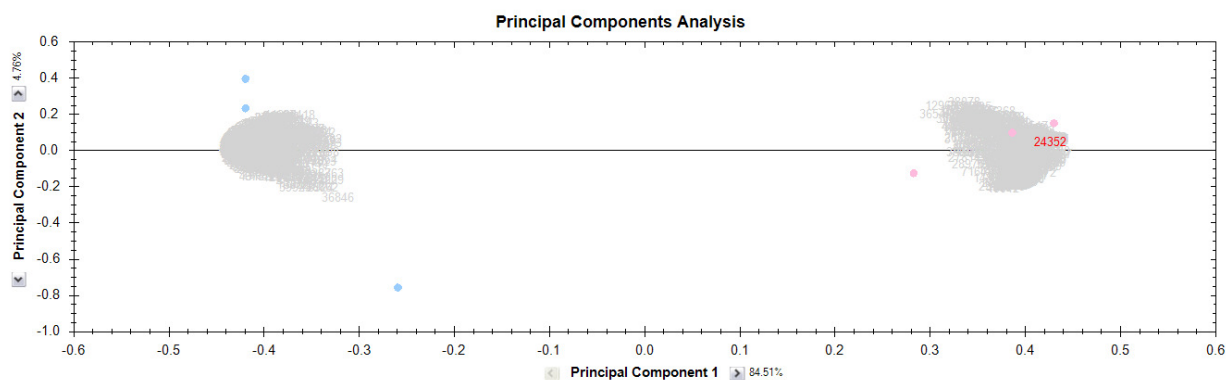
Calumenin				
Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
1	1	43	0.0402	5.6

**Figure 3.1.5** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the upregulation of the intracellular protein Transitional endoplasmic reticulum ATPase (VCP) and down-regulation of the secreted protein Calumenin. Both proteins were identified as being differentially-expressed when 96hrs was compared to 24hrs at 31°C in the conditioned media of the non-producing hybridoma cell line 4/2D.



### 3.1.3.3 Differential regulation of HCPs at 24hrs 37°C vs 31°C

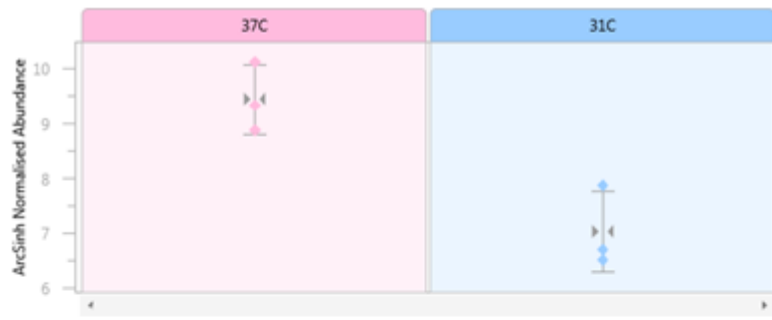
A PCA plot generated by Progenesis during data analysis indicates how the samples from 37°C and 31°C cultures at 24hrs cluster is displayed in Figure 3.1.6.



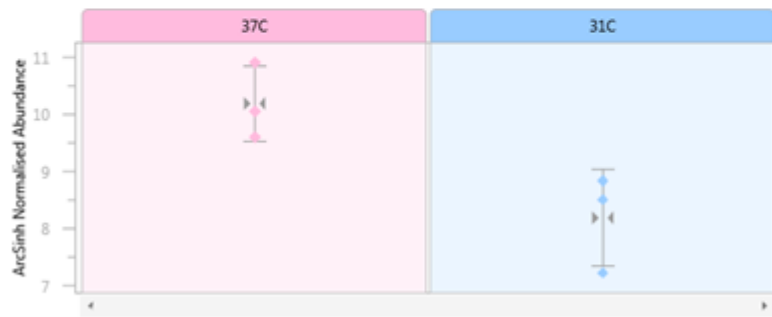
**Figure 3.1.6** PCA plot for non-producing hybridoma cell line sample groups 37°C (pink) and 31°C (blue) after 24hrs in culture generated during label-free LC/MS data analysis using Progenesis software.

In order to identify changes in the HCP profile of a non-producing hybridoma cell line as a result of temperature-shift, samples taken 24hrs post-temperature-shift from 37°C and 31°C cultures were compared. Of the 27 differentially-expressed proteins with 1 or more peptides identified, 4 were deemed to increase in abundance at 31°C (Appendix A, Table 3.1.5) while 23 were down-regulated (Appendix A, Table 3.1.6).

The Progenesis output from some of the differentially-expressed proteins identified, including Laminin subunit gamma-1 (10.5-fold) and Nidogen-2 (7-fold), are shown in Figure 3.1.7. Interestingly, a number of secreted proteins are down-regulated at 31°C, these are of particular interest because of the role they play in cell-cell signalling events. This raises the question; do secreted proteins play a role in the cellular response to temperature-shift? These proteins could warrant further investigation to determine whether extracellular-signalling events influence the growth of temperature-shifted cells.



Laminin subunit gamma-1				
Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
1	1	52	0.0125	10.5

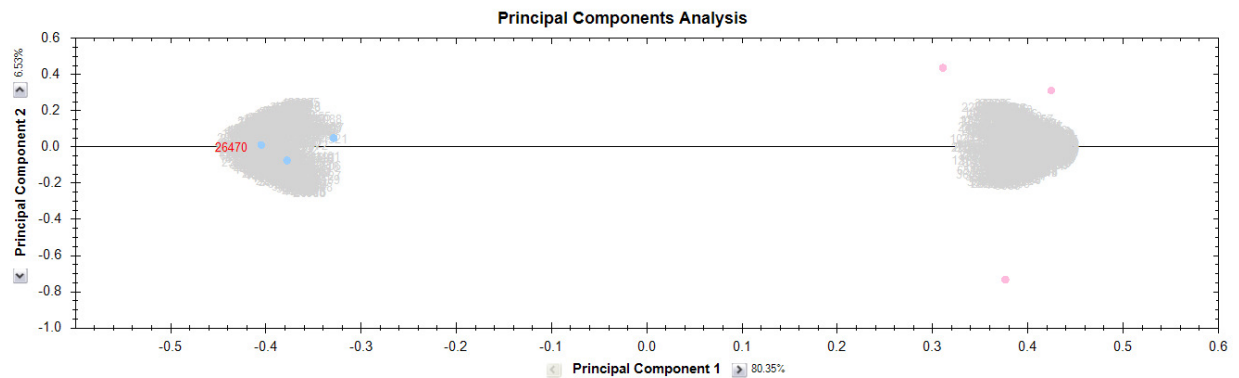


Nidogen-2				
Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
1	1	49	0.0324	7.0

**Figure 3.1.7** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the secreted proteins Laminin subunit gamma-1 and Nidogen 2, both of which were identified as being down-regulated at 31°C compared to 37°C 24hrs post-temperature-shift in the conditioned media of the non-producing hybridoma cell line 4/2D.

### 3.1.3.4 Differential regulation of HCPs at 96hrs 37°C vs 31°C

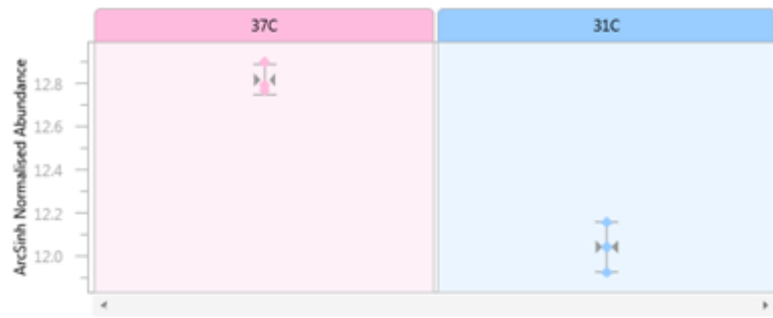
A PCA plot generated by Progenesis label-free software during analysis of LC/MS data is shown in Figure 3.1.8. The PCA plot displays how the samples from 37°C and 31°C cultures at 96hrs post-temperature-shift cluster.



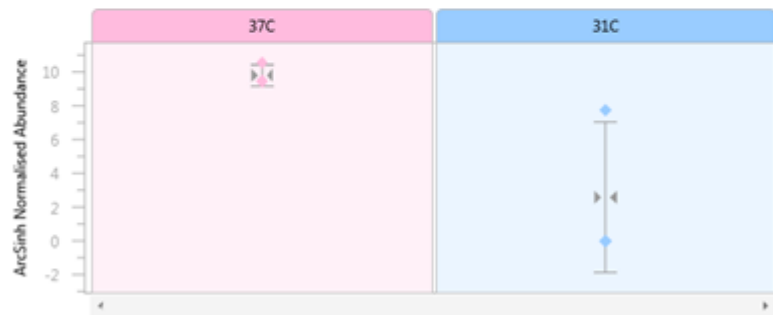
**Figure 3.1.8** PCA plot for non-producing hybridoma cell line sample groups 37°C (pink) and 31°C (blue) after 120hrs in culture generated during label-free LC/MS data analysis using Progenesis software.

A direct comparison of 37°C and 31°C cultures 96hrs post culture temperature reduction revealed the differential expression of 44 proteins with 1 or more peptides. The 31 HCPs identified as being down-regulated in the temperature-shifted culture are shown in Appendix A, Table 3.1.7. The 13 proteins that showed an increase in abundance are listed in Appendix A, Table 3.1.8.

Given the reduction of culture viability in the 37°C culture compared to the 31°C culture, the down-regulation of HCPs in the temperature-shifted culture makes sense given the likely increase in the release of intracellular proteins from dead cells at 37°C. Among the proteins identified as down-regulated at 31°C were a number of proteins that could negatively impact on product quality such as the protease Endoplasmic reticulum aminopeptidase 1 (2.2-fold) or the secreted protease Cathepsin B (27.9-fold), both of which were down-regulated at 31°C (Figure 3.1.9).



Endoplasmic reticulum aminopeptidase 1				
Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
4	4	178	0.0006	2.2



Cathepsin B				
Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
1	1	48	0.0497	27.9

**Figure 3.1.9** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the protease Endoplasmic reticulum aminopeptidase 1 and the secreted protease Cathepsin B, both of which were identified as being down-regulated at 31°C compared to 37°C, 96hrs post-temperature-shift in the conditioned media of the non-producing hybridoma cell line 4/2D.

### **3.1.3.5 Comparison of differentially-regulated HCPs over time in culture at 37°C and 31°C in a non-producing hybridoma cell line**

Quantitative analysis of conditioned media samples from 24hrs compared to 96hrs at 37°C identified 80 differentially-expressed proteins, while a comparison between 24hrs and 96hrs at 31°C revealed changes in the relative abundance of 66 proteins. In order to compare the effect of temperature-shift on proteins over time in culture with those in a non-temperature-shifted culture, lists of differentially-expressed proteins from sections 3.1.3.1 and 3.1.3.2 were combined and proteins common to both lists compiled in Table 3.1.9. The low number of proteins that overlap between the two lists may be as a result of the differential expression of proteins under one culture temperature and no change in relative abundance at the other temperature, as opposed to a complete absence of the protein from the culture media. As can be seen in Table 3.1.9, many of the proteins trend in the same direction over time in culture, such as the secreted protein Laminin subunit gamma-1 and the pro-apoptotic protein Guanine nucleotide-binding protein subunit beta-2-like 1, both of which were down-regulated after 96hrs in culture. There are a number of proteins however that show opposite trends, such as the intracellular proteins Transitional endoplasmic reticulum ATPase (VCP), 60S acidic ribosomal protein P0-like and Elongation factor 1-gamma. Opposite trends in the differential expression of intracellular proteins in the cell culture media may be as a result of differences in the viabilities of the two cultures at 96hrs. Functionally, these proteins do not group into specific cellular roles. However, information such as this could be useful to the bioprocessing community, for example; if a recombinant therapeutic is known to be negatively affected by the presence of a specific protease secreted by the cells into the culture media, comparing the data as provided in Table 3.1.9, would enable identification of the culture conditions that generates the lowest amount of this protease.

**Table 3.1.9** List of differentially-expressed proteins common to 24hrs vs 96hrs 37°C and 24hrs vs 96hrs 31°C from a non-producing hybridoma cell line.

Gene symbol	Description 37C	Total Peptide	Quantitative Peptides	Score	Anova (p)	Fold	Direction of fold change	Comparison
Hspd1	60 kDa heat shock protein, mitochondrial	2	2	104	0.018	28.9	Down at 96hrs	24hrs vs 96hrs 37C
		3	3	147	0.0108	3.8	Down at 96hrs	24hrs vs 96hrs 31C
Rplp0	60S acidic ribosomal protein P0	1	1	62	0.0217	102.1	Down at 96hrs	24hrs vs 96hrs 37C
		2	2	88	0.0035	3	Up at 96hrs	24hrs vs 96hrs 31C
Alpl	Alkaline phosphatase, tissue-nonspecific isozyme	1	1	54	3.57E-05	Infinity	Up at 96hrs	24hrs vs 96hrs 37C
		1	1	55	0.0438	1.7	Up at 96hrs	24hrs vs 96hrs 31C
Got2	Aspartate aminotransferase, mitochondrial	1	1	52	0.0066	Infinity	Down at 96hrs	24hrs vs 96hrs 37C
		2	2	113	0.0288	2.4	Down at 96hrs	24hrs vs 96hrs 31C
Eef1g	Elongation factor 1-gamma	1	1	74	0.0318	23	Down at 96hrs	24hrs vs 96hrs 37C
		1	1	47	0.0196	6.6	Up at 96hrs	24hrs vs 96hrs 31C
Erp29	Endoplasmic reticulum resident protein 29	1	1	43	0.0126	484.6	Up at 96hrs	24hrs vs 96hrs 37C
		1	1	42	0.0023	2.8	Down at 96hrs	24hrs vs 96hrs 31C
G6pdx	Glucose-6-phosphate 1-dehydrogenase X	2	2	92	0.0027	8.7	Down at 96hrs	24hrs vs 96hrs 37C
		1	1	62	0.0438	21492.2	Down at 96hrs	24hrs vs 96hrs 31C
Gnb2l1	Guanine nucleotide-binding protein subunit beta-2-like 1	1	1	44	0.0002	17.2	Down at 96hrs	24hrs vs 96hrs 37C
		1	1	52	0.0125	2.4	Down at 96hrs	24hrs vs 96hrs 31C
Hnrnpab	Heterogeneous nuclear ribonucleoprotein A/B	1	1	43	0.03	1545.6	Down at 96hrs	24hrs vs 96hrs 37C
		2	2	90	0.0204	4.2	Up at 96hrs	24hrs vs 96hrs 31C
Impdh2	Inosine-5'-monophosphate dehydrogenase 2	1	1	40	0.0407	12.6	Down at 96hrs	24hrs vs 96hrs 37C
		2	2	93	0.009	2.5	Down at 96hrs	24hrs vs 96hrs 31C

Lmb1	Lamin-B1	2	2	98	0.0284	214.2	Up at 96hrs	24hrs vs 96hrs 37C
		1	1	45	0.0489	7.2	Up at 96hrs	24hrs vs 96hrs 31C
Lamc1	Laminin subunit gamma-1	1	1	64	0.0096	16935.2	Down at 96hrs	24hrs vs 96hrs 37C
		1	1	41	0.0384	3.8	Down at 96hrs	24hrs vs 96hrs 31C
Nme2	Nucleoside diphosphate kinase B	1	1	42	0.0017	30	Down at 96hrs	24hrs vs 96hrs 37C
		1	1	53	0.0307	8.9	Up at 96hrs	24hrs vs 96hrs 31C
Olfm3	Olfactomedin-like protein 3	1	1	40	0.0295	1336.7	Down at 96hrs	24hrs vs 96hrs 37C
		1	1	53	0.0201	2.4	Up at 96hrs	24hrs vs 96hrs 31C
Pafah1b3	Platelet-activating factor acetylhydrolase IB subunit gamma	1	1	56	3.63E-06	Infinity	Up at 96hrs	24hrs vs 96hrs 37C
		1	1	60	0.0441	3.8	Up at 96hrs	24hrs vs 96hrs 31C
P4hb	Protein disulfide-isomerase	1	1	47	0.0044	Infinity	Down at 96hrs	24hrs vs 96hrs 37C
		1	1	49	0.0003	3.3	Down at 96hrs	24hrs vs 96hrs 31C
Pdia4	Protein disulfide-isomerase A4	1	1	65	0.0017	95.8	Down at 96hrs	24hrs vs 96hrs 37C
		1	1	50	0.0055	1.9	Down at 96hrs	24hrs vs 96hrs 31C
Vcp	Transitional endoplasmic reticulum ATPase	3	3	153	0.0201	323	Down at 96hrs	24hrs vs 96hrs 37C
		2	2	116	0.0022	2	Up at 96hrs	24hrs vs 96hrs 31C

### **3.1.4 Qualitative analysis of proteins identified by label-free LC/MS profiling of conditioned media from a non-producing hybridoma cell line**

Progenesis label-free software provides quantitative information on changes in the relative abundance of proteins. This however, ignores the fact that there are also proteins present in the sample, the expression of which may not change. Knowledge of these proteins may be useful in providing insights into whether significant differences exist between temperature-shifted and non-temperature-shifted hybridoma cultures.

The RAW files obtained from LC/MS analysis of the conditioned media samples collected in biological triplicate and taken at 24, 72, and 96 hours post-temperature-shift from 37°C and 31°C cultures as per Figure 3.1.1, were searched against the mouse UniproKB database using Mascot and SEQUEST search algorithms. These searches were conducted in parallel using Proteome Discoverer 1.4. Proteins were “identified” if they met the criteria as set out in section 2.4.7. In order for an ‘identified’ protein to be considered “present” either at 37°C or 31°C, it had to be observed in the mass spectral data from a minimum of two biological samples at the given temperature.

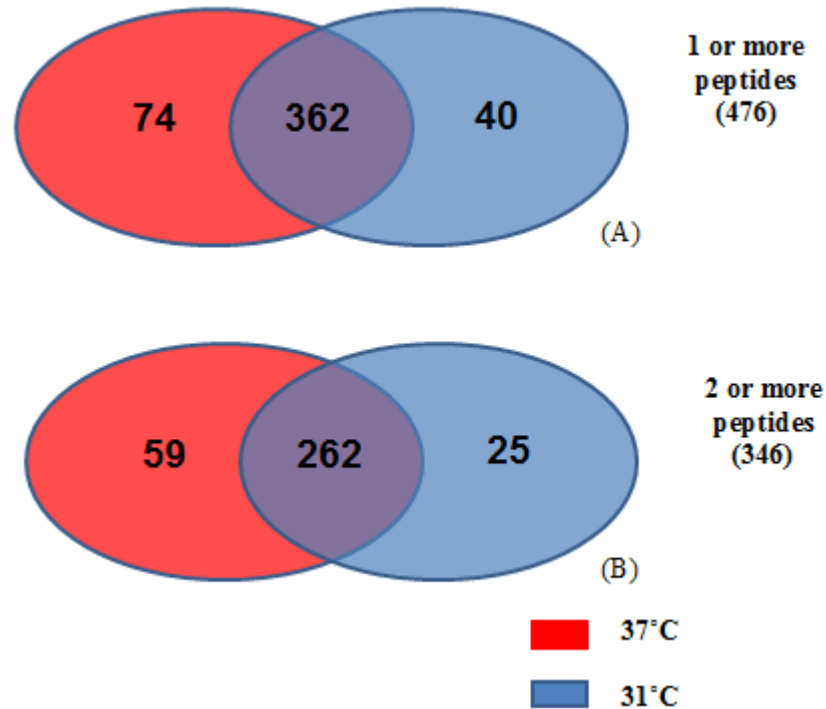
The following comparisons were made by overlapping the lists of identified proteins:

- Hybridoma cell line 4/2D - 37°C vs 31°C Cultures
  - Proteins identified with 1 or more peptides
  - Proteins identified with 2 or more peptides

As can be seen in Figure 3.1.10 A, 476 proteins with one or more peptides were identified in both temperature-shifted and non-temperature-shifted cultures. Of these, 362 were common to both cultures, 74 unique to the 37°C culture and 40 unique to the 31°C culture. By using a minimum of two peptides to identify a protein, the confidence with which that protein is identified increases. Figure 3.1.10 B shows that 262 proteins with two or more peptides are common to 37°C and 31°C cultures, while 59 are unique to the non-temperature-shifted culture and 25 are unique to the temperature-shifted culture.



This shows that despite increasing the criteria for protein identification, that differences exist between the proteins expressed at 37°C and 31°C. Lists of the proteins identified in Figure 3.1.10 A and Figure 3.1.10 B are available in Appendix A Tables 3.1.10 - Table 3.1.14.



**Figure 3.1.10** Overlap of the list of proteins identified in the conditioned media of a non-producing hybridoma cell line grown at 37°C and 31°C over 144 hours in stationary suspension culture. Samples were collected in biological triplicate and taken at 24, 72 and 96 hours post-temperature-shift. Venn diagram (A) represents those proteins identified using a minimum of 1 or more peptides. Venn diagram (B) is compiled from those proteins identified using a minimum of 2 or more peptides.

While some of the differences seen between temperature-shifted and non-temperature-shifted cultures could be attributed to a reduction in the viability of the 37°C culture and the subsequent release of intracellular proteins from dead cells, another possible explanation could be the induced expression of a cohort of proteins in the temperature-shift culture as a result of the cells undergoing temperature-shift. In either event, the identification of a set of proteins unique to temperature-shift and non-temperature-shift culture conditioned has important implications (i) for the ability of a generic anti-HCP ELISA to detect contaminating HCPs in the final drug product.

### **3.1.5 Summary**

Attempts to group proteins into cellular functions e.g. protease, growth factor, apoptotic etc., and identify trends in the differential-expression of these groups either over time in culture or between temperature-shift and non-temperature-shift conditions proved inconclusive (data not shown). Despite this, analysis of the HCP profile identified changes in the expression of proteases such as Cathepsin B, Cathepsin D and Endoplasmic reticulum aminopeptidase 1 as well as the apoptotic protein Guanine nucleotide-binding protein subunit beta-2-like 1 under both temperature-shifted and non-temperature-shifted conditions over time in culture. This provides useful information for the bioprocessing community for the selection of culture conditions for the production of a recombinant therapeutic sensitive to enzymatic activity. An interesting finding of this investigation was the identification of a number of secreted proteins that change in abundance when 37°C and 31°C cultures are compared at both early and later phases of the growth cycle. This possibly indicates an active cell-cell signalling response to reduction in culture temperature. Additionally, the identification of a cohort of proteins unique to both temperature-shift and non-temperature-shift cultures is a significant finding of this work. During the time that this analysis was conducted the CHO genome was sequenced and access to a CHO protein database became available. The availability of such a database would provide greater coverage of the CHO proteome and increased confidence in the assignment of protein identifications.

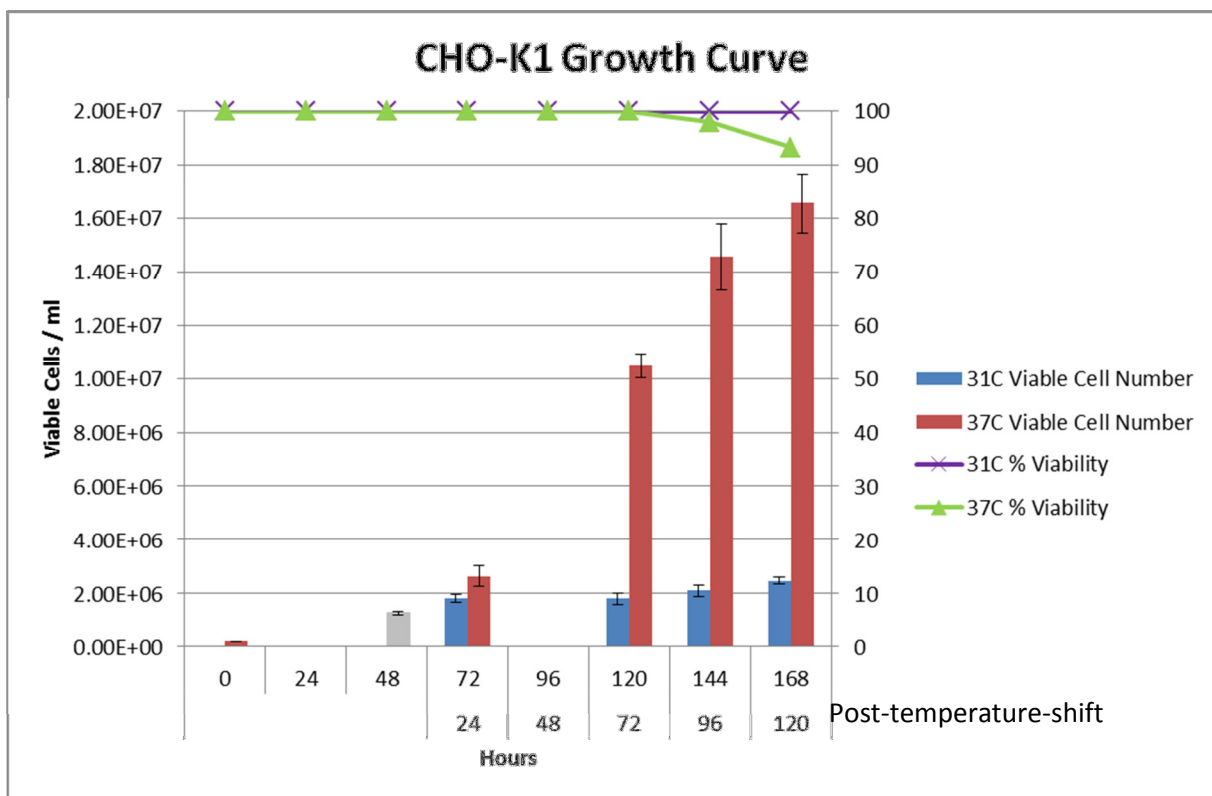
Given that this analysis was carried out in a hybridoma cell line, and we were able to successfully identify HCP proteins, it was decided that this investigation should be extended to include one of the most commonly used cell line in the production of biopharmaceuticals; the CHO cell.

## **3.2 Characterisation of the host cell protein profile of a non-producing CHO-K1 cell line**

While there have recently been a number of publications discussing the host cell protein (HCP) profile of CHO cell lines, they have primarily focused on late stationary and death phases of culture. This profiling has also almost exclusively been based around the use of (2D) gel based technology. This chapter details the use of a quantitative, label-free, mass spectrometry based approach to identify changes in the HCP profile of a non-producing cell line over time in culture, from early exponential to late stationary phase. In addition to this, the HCP profile of a culture under reduced culture temperature conditions is also examined. This allows for direct comparisons to be made between temperature-shifted and non-temperature-shifted conditions over the cell culture cycle.

### **3.2.1 Growth profiling of non-producing CHO-K1 cell line**

CHO-K1 cells were grown in triplicate in a 50ml volume of chemically-defined protein free OptiCHO™ media in shake flask batch culture. Cells were seeded at  $2 \times 10^5$  cell/ml and after 48 hours in culture flasks selected for temperature-shift were placed in a 31°C incubator for the remainder of the culture period. Cell culture supernatants were collected 24, 72, 96 and 120 hours post-temperature-shift from both temperature-shifted and non-temperature-shifted cultures Figure 3.2.1 (as per materials and methods 2.3.2). While it is recognised that 48 hours may not have been an optimum time to place cells under reduced culture conditions from the point of obtaining maximum yield of a therapeutic product from the process, it provides clarity from a biological perspective in terms of identifying a functional response to a change in the culture conditions.



**Figure 3.2.1** Growth curve of CHO-K1 non-producing cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions. Temperature-shifted culture was placed at 31°C after 48hrs in culture. Cell culture supernatants were collected 24, 72, 96 and 120 hours post-temperature-shift and analysed using a quantitative, label-free, mass spectrometry based approach.

### **3.2.2 Quantitative Label-free LC/MS analysis of conditioned media from a CHO-K1 non-producing cell line**

Cell culture supernatants from three biological replicates at time points under both temperature-shifted and non-temperature-shifted conditions were concentrated ~100 fold using centrifugal concentrators with a 3kDa molecular weight cut off. Protein from an aliquot of each sample was then precipitated using a ReadyPrep™ 2D clean up kits. 10µg of sample was then prepared for LC/MS analysis as detailed in materials and methods 2.3.3. It was established during preliminary studies that a 3hr reverse phase gradient was sufficient for LC/MS analysis of conditioned media as these samples did not have the same level of complexity as a whole cell lysate, which requires a 5hr gradient, however the use of 1hr gradients resulted in loss of peptides. After mass spectral analysis of all the samples was complete, inclusion lists were generated and analysis for some of the samples was repeated. This was done to obtain information on additional peptides to further strengthen protein identifications and identify low abundance proteins that were not previously identified in the initial analysis.

### **3.2.3 Data analysis of differentially-regulated proteins identified by quantitative label-free LC/MS profiling of conditioned media from a CHO-K1 non-producing cell line**

Progenesis label-free software was used to analyse the data acquired following LC/MS analysis. A protein was considered to be differentially-regulated if it met the criteria as set out in 3.1.3.

Due to the complex sampling nature of the experiment i.e. multiple time points between two conditions, it was decided that the data would be analysed as follows:

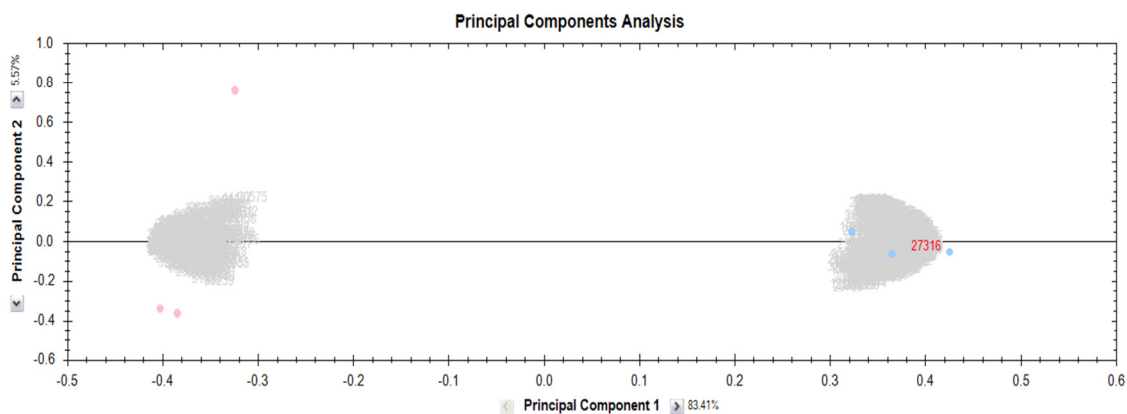
- 24hrs vs 120hrs at 37°C
- 24hrs vs 120hrs at 31°C
- 37°C vs 31°C at 24hrs
- 37°C vs 31°C at 120hrs
- Analysis of differential regulation over 24, 72, 96 and 120 hours at 37°C using Pearson correlation analysis
- Analysis of differential regulation over 24, 72, 96 and 120 hours at 31°C using Pearson correlation analysis

Note: The time in hours above is with respect to temperature-shift. i.e. 24hrs is 24hrs post-temperature-shift.

In order to infer a biological function for the differentially-regulated proteins identified, the gene symbol for the human version of the protein was obtained and searched in the Uniprot database (<http://www.uniprot.org/>). Proteins involved in in roles deemed to be of potential importance were classified as protease, glycosidase, apoptotic or miscellaneous.

### 3.2.3.1 Differential regulation of HCPs 24hrs vs 120hrs at 37°C

A principal component analysis (PCA) plot generated by Progenesis during data analysis showing how the samples from 24hr and 120hr cultures at 37°C clustered is displayed in Figure 3.2.2.



**Figure 3.2.2** PCA plot for CHO-K1 non-producing cell line sample groups 24hrs (pink) and 120hrs (blue) at 37°C generated during quantitative label-free LC/MS data analysis using Progenesis software.

A total of 322 proteins with 1 or more peptides were identified as being differentially-expressed at 37°C between early (24hrs) and late (120hrs). Of these, 274 (with 1 or more peptides) had increased expression at 120hr, while 48 were found to be down-regulated (Table 3.2.1) (see Appendix B, Table 3.2.2 and Table 3.2.3 for full lists).

As can be seen in Table 3.2.1 proteins involved in apoptosis increased in abundance at 120hrs compared to 24hrs, this most likely reflects the decrease in cell viability the later time point. An example of the up-regulation of one such pro-apoptotic protein, Guanine nucleotide-binding protein subunit beta-2-like 1 (3.2-fold) is shown in Figure 3.2.3. The predominant up-regulation of proteases at 120hrs would suggest that recombinant products would be more susceptible to enzymatic activity at the late phase of culture. Conversely, another group of proteins likely to negatively impact on product quality, glycosidases, are down-regulated at 120hrs compared to 24hrs at 37°C. Figure 3.2.3 shows the 1.8-fold down-regulation of the glycosidase Beta-galactosidase.

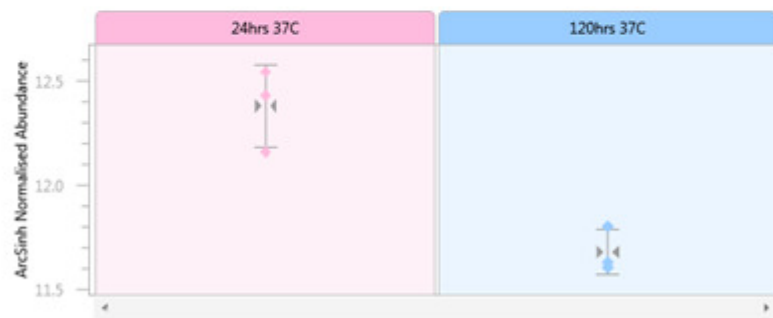


**Table 3.2.1** Number of differentially-expressed proteins at 24hrs compared to 120hrs at 37°C in a CHO-K1 non-producing cell culture supernatant. Using Uniprot protein database, proteins of biological relevance have been identified and the number of differentially-expressed proteins under the relevant classification included in the table. The protein count indicates the number of proteins identified by  $\geq 1$  peptides or  $\geq 2$  peptides.

	24hrs vs 120hrs at 37°C					
	Total		Up at 120hrs		Down at 120hrs	
	$\geq 1$ Peptide	$\geq 2$ Peptides	$\geq 1$ Peptide	$\geq 2$ Peptides	$\geq 1$ Peptide	$\geq 2$ Peptides
<b>Protein Count</b>	322	112	274	96	48	16
<b>Apoptosis</b>	5	2	5	2	0	0
	GNB2L1, CLU, API5, PDCD6IP, PPM1F	GNB2L1, CLU	GNB2L1, CLU, API5, PDCD6IP, PPM1F	GNB2L1, CLU		
<b>Protease</b>	20	5	14	2	6	3
	DPP7, THOP1, C1ra, Otub1, CTSZ, BMP1, ADAM10, ADAM15, ADAM17, EIF3F, PSMA3, PSMA4, PSMB1, PSMB4, PSMB6, UCHL3, XPNPEP1, ADAMTS1, MMP9, CTSA	DPP7, THOP1, C1ra, Otub1, CTSZ	THOP1, Otub1, BMP1, ADAM10, ADAM15, ADAM17, EIF3F, PSMA3, PSMA4, PSMB1, PSMB4, PSMB6, UCHL3, XPNPEP1	THOP1, Otub1	DPP7, C1ra, CTSZ, ADAMTS1, MMP9, CTSA	DPP7, C1ra, CTSZ
<b>Glycosidase</b>	6	3	2	0	4	3
	MAN2B2, NAGA, GLB1, NEU1, FUCA1, SMPD1	MAN2B2, NAGA, GLB1	NEU1, FUCA1		MAN2B2, NAGA, GLB1, SMPD1	MAN2B2, NAGA, GLB1



Guanine nucleotide-binding protein subunit beta-2-like 1				
Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
4	4	247	0.0027	3.2

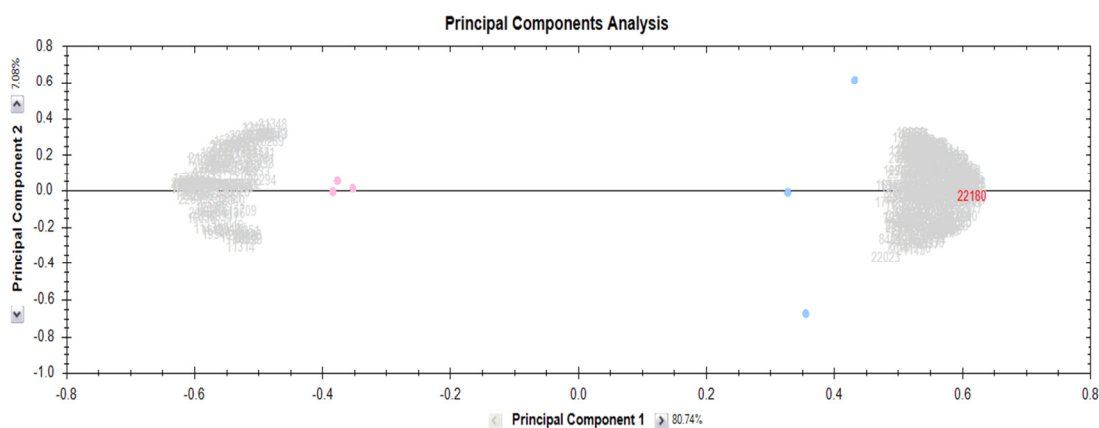


Beta-galactosidase				
Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
2	2	132	0.0181	1.8

**Figure 3.2.3** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the apoptotic protein Guanine nucleotide-binding protein subunit beta-2-like 1 and the glycosidase Beta-galactosidase, identified as being differentially-expressed between 24hrs and 120hrs at 37°C in the conditioned media of a CHO-K1 cell line.

### 3.2.3.2 Differential regulation of HCPs 24hrs vs 120hrs at 31°C

A PCA plot generated by Progenesis during data analysis showing how the samples from 24hr and 120hr cultures at 31°C cluster is displayed in Figure 3.2.4.



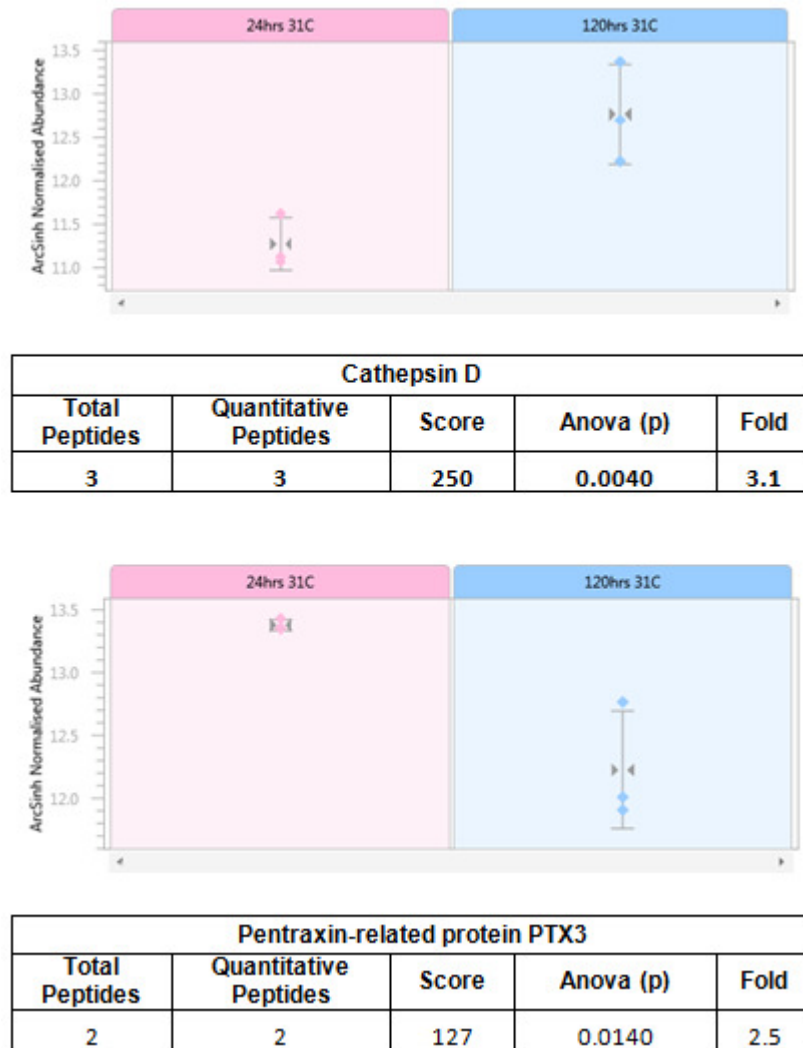
**Figure 3.2.4** PCA plot for non-producing CHO-K1 cell line sample groups 24hrs (pink) and 120hrs (blue) at 31°C generated during quantitative label-free LC/MS data analysis using Progenesis software.

When early (24hrs) and late (120hrs) phases of the temperature-shifted (31°C) were compared 146 proteins identified with 1 or more peptides were found to change in relative abundance (96 up-regulated at 120hrs and 50 down-regulated) (Table 3.2.4) (For full list see Appendix B, Table 3.2.5 and Table 3.2.6). Although there is still an overall increase in the expression of proteases at the later time point, there are not as many as was observed in the 37°C (Table 3.2.1). Interestingly, there is an increase the number of up-regulated glycosidases at the later time point at 31°C compared to the 37°C culture. Taken together, this would indicate that while a recombinant therapeutic would be less susceptible to protease activity at 31°C, glycans on the molecule would be more prone to cleavage due to the up-regulation of glycosidases at 120hrs in the temperature-shifted culture.

**Table 3.2.4** Number of differentially-expressed proteins at 24hrs compared to 120hrs at 31°C in a CHO-K1 non-producing cell culture supernatant. Using Uniprot protein database, proteins of biological relevance have been identified and the number of differentially-expressed proteins under the relevant classification included in the table. The protein count indicates the number of proteins identified by  $\geq 1$  peptides or  $\geq 2$  peptides.

	24hrs vs 120hrs at 31°C					
	Total		Up at 120hrs		Down at 120hrs	
	$\geq 1$ Peptide	$\geq 2$ Peptides	$\geq 1$ Peptide	$\geq 2$ Peptides	$\geq 1$ Peptide	$\geq 2$ Peptides
<b>Protein Count</b>	146	38	96	24	50	14
<b>Apoptosis</b>	2	0	0	0	2	0
	APP, Cycs				APP, Cycs	
<b>Protease</b>	8	2	6	1	2	1
	CTSD, C1ra, CTSL1, CTSB, PSMB1, PSMB6, Otub1, ADAMTS7	CTSD, C1ra	CTSD, CTSL1, CTSB, PSMB1, PSMB6, Otub1	CTSD	ADAMTS7, C1ra	C1ra
<b>Glycosidase</b>	6	3	5	2	1	1
	NAGLU, GANAB, NAGA, MAN2B2, GAA, IDUA	NAGLU, GANAB, NAGA	NAGLU, NAGA, MAN2B2, GAA, IDUA	NAGLU, NAGA	GANAB	GANAB

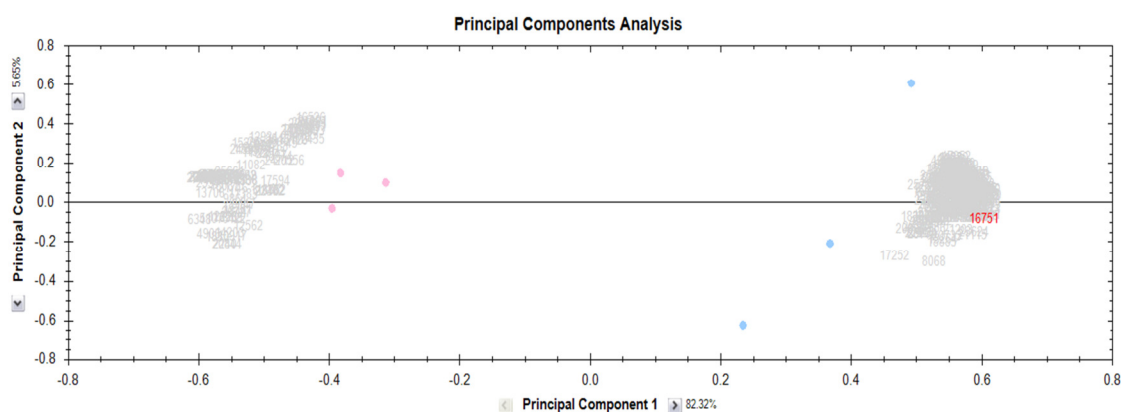
Among the differentially-expressed proteins identified in this study at 120hrs compared to 24hrs at 31°C was the protease Cathepsin D (3.1-fold up-regulated) and the secreted protein Pentraxin-related protein PTX3 (2.5-fold down-regulated) as shown in Figure 3.2.5.



**Figure 3.2.5** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the protease Cathepsin D (up-regulated at 24hrs and 120hrs at 31°C) and the secreted protein Pentraxin-related protein PTX3 (down-regulated at 24hrs and 120hrs at 31°C), both of which were identified in the conditioned media of a CHO-K1 cell line.

### 3.2.3.3 Differential regulation of HCPs at 24hrs 37°C vs 31°C

A PCA plot generated by Progenesis during data analysis indicates how the samples from 37°C and 31°C cultures at 24hrs cluster is displayed in Figure 3.2.6.



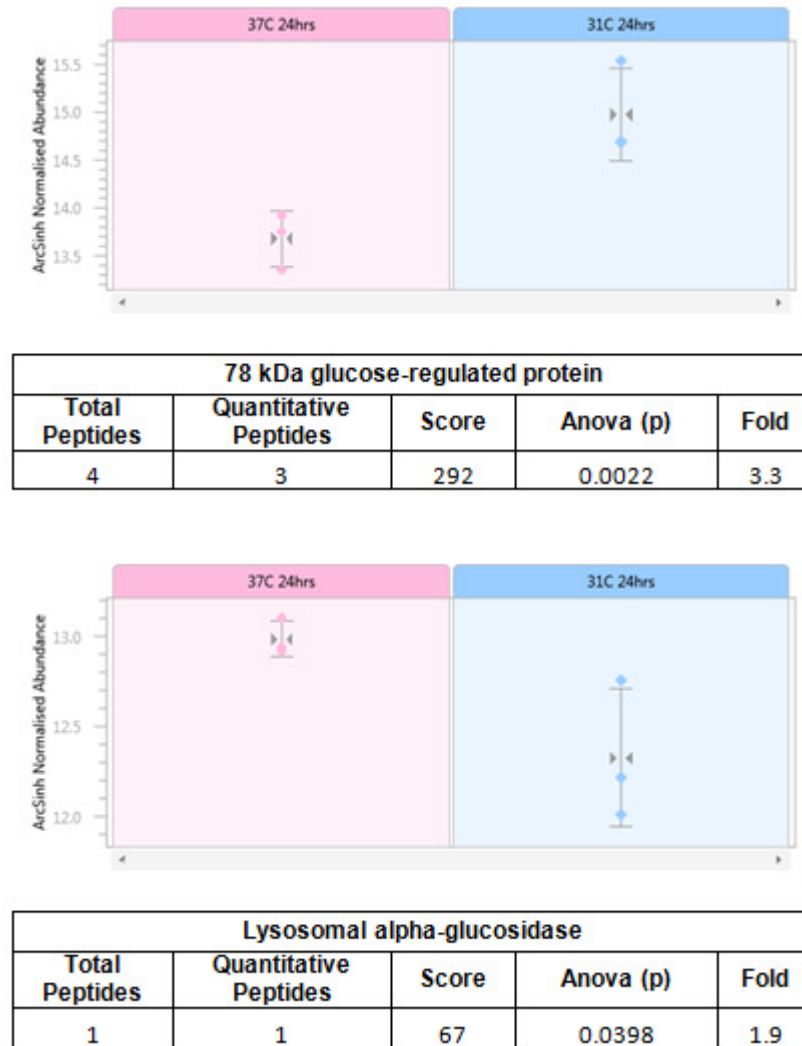
**Figure 3.2.6** PCA plot for CHO-K1 non-producing cell line sample groups 37°C (pink) and 31°C (blue) after 24hrs in culture generated during quantitative label-free LC/MS data analysis using Progenesis software.

A direct comparison of 37°C and 31°C cultures 24hrs post culture temperature reduction revealed a significant increase in the relative expression of HCPs in the 31°C culture. Of the 211 differentially-expressed proteins with 1 or more peptides identified, only 14 were down-regulated at 31°C while 197 were deemed to increase in overall expression. This included up-regulation of 4 apoptotic proteins and 7 proteases and a decrease in the abundance of 2 glycosidases (Table 3.2.7). (A full list of differentially-expressed HCPs is provided in Appendix B Table 3.2.8 and Table 3.2.9).

**Table 3.2.7** Number of differentially-expressed proteins between 31°C and 37°C at 24hrs post-temperature-shift in a CHO-K1 non-producing cell culture supernatant. Using Uniprot protein database, proteins of biological relevance have been identified and the number of differentially-expressed proteins under the relevant classification included in the table. The protein count indicates the number of proteins identified by  $\geq 1$  peptides or  $\geq 2$  peptides.

	37°C vs 31°C 24hrs					
	Total		Up at 31°C		Down at 31°C	
	$\geq 1$ Peptide	$\geq 2$ Peptides	$\geq 1$ Peptide	$\geq 2$ Peptides	$\geq 1$ Peptide	$\geq 2$ Peptides
<b>Protein Count</b>	211	55	197	52	14	3
<b>Apoptosis</b>	5	0	4	0	1	0
	CLU, API5, GNB2L1, KIAA1967, PPM1F		API5, GNB2L1, KIAA1967, PPM1F		CLU	
<b>Protease</b>	11	5	7	3	4	2
	CTSH, DPP7, THOP1, UCHL3, Otub1, MMP19, TPP1, CNDP2, DPP3, LTA4H, PREP	CTSH, DPP7, THOP1, UCHL3, Otub1	THOP1, UCHL3, Otub1, CNDP2, DPP3, LTA4H, PREP	THOP1, UCHL3, Otub1	CTSH, DPP7, MMP19, TPP1	CTSH, DPP7
<b>Glycosidase</b>	3	1	1	1	2	0
	GANAB, HEXA, GAA	GANAB	GANAB	GANAB	HEXA, GAA	

The Progenesis output from some of the differentially-expressed proteins identified are shown in Figure 3.2.7. This includes the up-regulation of the intracellular protein GRP78 (3.3-fold) and the 1.9-fold down-regulation of the glycosidase Lysosomal alpha-glucosidase at 31°C when compared to 37°C after 24hrs.

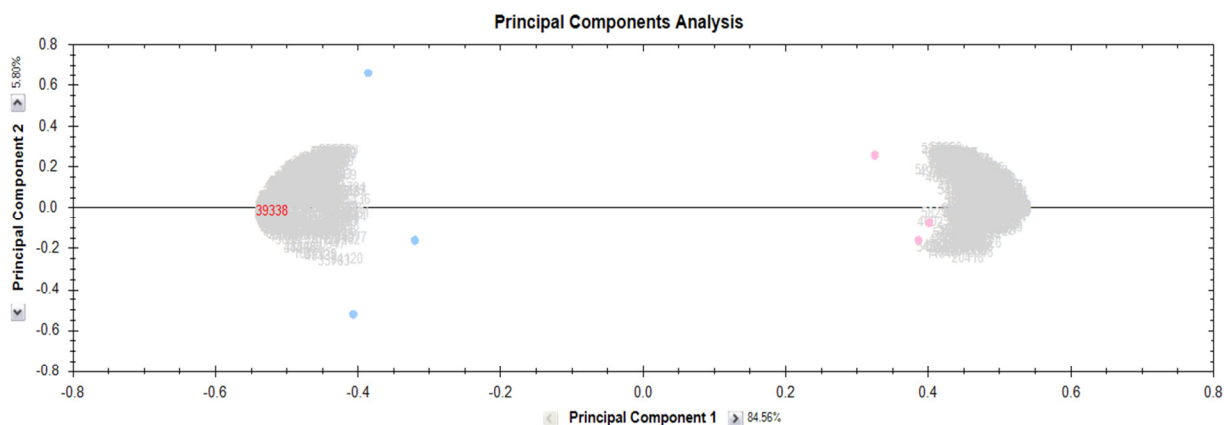


**Figure 3.2.7** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as GRP78, up-regulated at 31°C and the glycosidase Lysosomal alpha-glucosidase, down-regulated at 31°C, 24hrs post-temperature-shift in the conditioned media of a CHO-K1 cell line.



### 3.2.3.4 Differential regulation of HCPs at 120hrs 37°C vs 31°C

A PCA plot generated by Progenesis during data analysis showing how the samples from 37°C and 31°C cultures at 120hrs cluster is displayed in Figure 3.2.8.



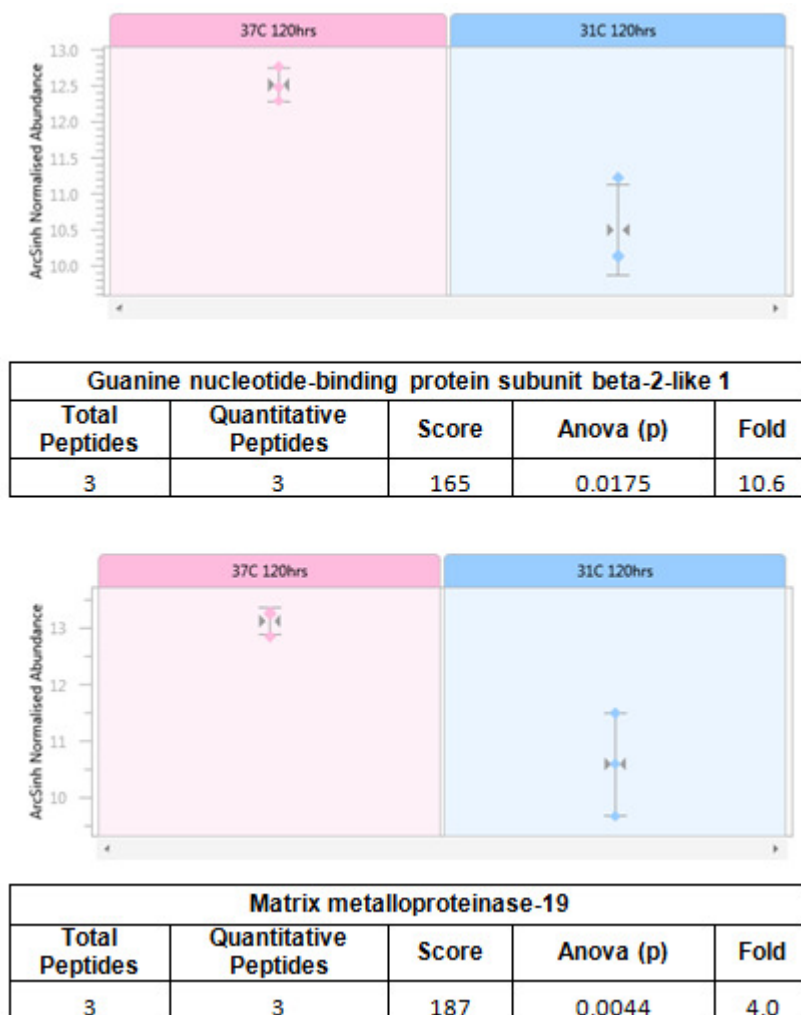
**Figure 3.2.8** PCA plot for CHO-K1 non-producing cell line sample groups 37°C (pink) and 31°C (blue) after 120hrs in culture generated during quantitative label-free LC/MS data analysis using Progenesis software

When 37°C and 31°C cultures are compared at the last time point, 120hrs post-temperature shift, 228 proteins with 1 or more peptides identified are seen to alter their expression, the majority of which are down-regulated in the 31°C culture (153 down-regulated, 75 up-regulated at 31°C compared to 37°C, 120hrs post-temperature-shift. See summary Table 3.3.10.) This is a reversal of the trend seen when both culture temperatures were compared at the 24hr time point (Table 3.2.7). The significant number of down-regulated proteases and glycosidases at 31°C compared to 37°C suggests that a recombinant product would be less susceptible to enzymatic activity in the temperature-shifted culture (Table 3.2.10). As seen in Table 3.2.10, six of the eight differentially-expressed apoptotic proteins identified are down-regulated at 31°C compared to 37°C. This could be attributed to the overall increased viability of the temperature-shifted culture compared to the non-temperature-shifted culture. (See Appendix B Table 3.2.11 and Table 3.2.12 for full list of differentially-expressed HCPs).

**Table 3.2.10** Number of differentially-expressed proteins between 31°C and 37°C at 120hrs post-temperature-shift in a CHO-K1 non-producing cell culture supernatant. Using Uniprot protein database, proteins of biological relevance have been identified and the number of differentially-expressed proteins under the relevant classification included in the table. The protein count indicates the number of proteins identified by  $\geq 1$  peptides or  $\geq 2$  peptides.

	37°C vs 31°C 120hrs					
	Total		Up at 31°C		Down at 31°C	
	$\geq 1$ Peptide	$\geq 2$ Peptides	$\geq 1$ Peptide	$\geq 2$ Peptides	$\geq 1$ Peptide	$\geq 2$ Peptides
<b>Protein Count</b>	228	67	75	18	153	49
<b>Apoptosis</b>	6	2	2	0	4	2
	GNB2L1, APP, GAPDH, API5, DDX3X, Cyps	GNB2L1, APP	DDX3X, Cyps		GNB2L1, APP, GAPDH, API5	GNB2L1, APP
<b>Protease</b>	17	5	2	1	15	4
	ADAMTS7, MMP19, PRSS22, BMP1, DPP7, HTRA1, P15156, PLAT, ADAM19, MMP9, THOP1, CTSZ, ADAM10, C1ra, ADAM17, ADAM15, CTSD	ADAMTS7, MMP1,9 PRSS22, BMP1, DPP7	DPP7 CTSD	DPP7	ADAMTS7, MMP19, PRSS22, BMP1, HTRA1, P15156, PLAT, ADAM19, MMP9, THOP1, CTSZ, ADAM10, C1ra, ADAM17, ADAM15	ADAMTS7, MMP19, PRSS22, BMP1
<b>Glycosidase</b>	7	1	3	1	4	0
	MAN2B2, CTBS, NEU1, FUCA1, GBA, GANAB, NAGA	MAN2B2	MAN2B2, GANAB, NAGA	MAN2B2	CTBS, NEU1, FUCA1, GBA	

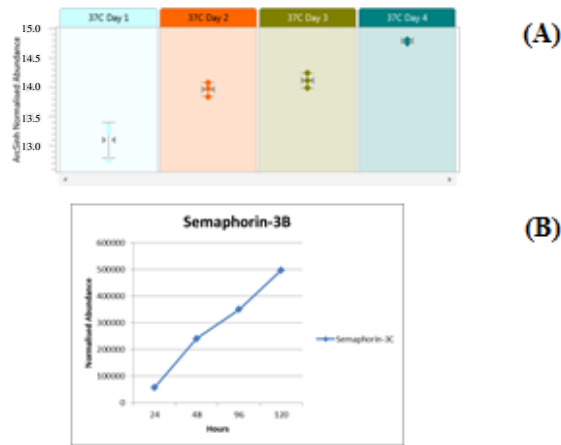
The Progenesis output indicating the down-regulation of the pro-apoptotic protein Guanine nucleotide-binding protein subunit beta-2-like 1 (10.6-fold) and the protease Matrix metalloproteinase-19 (4.0-fold) are shown in Figure 3.2.9.



**Figure 3.2.9** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the apoptotic protein Guanine nucleotide-binding protein subunit beta-2-like 1 and the protease Matrix metalloproteinase-19, identified as being down-regulated at 31°C 24hrs post-temperature-shift in the conditioned media of a CHO-K1 cell line.

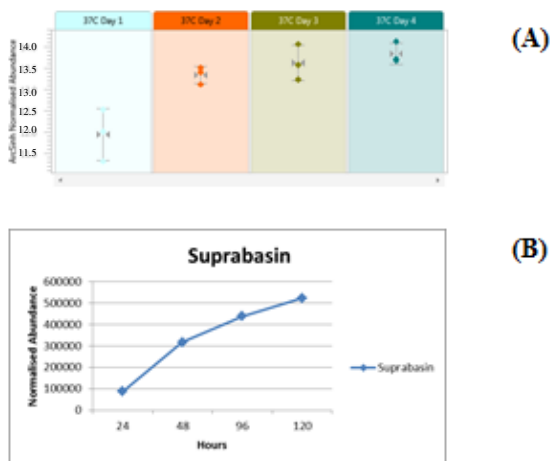
### **3.2.3.5 Pearson correlation analysis of differentially-regulated HCPs over 24, 72, 96 and 120 hours at 37°C**

Pearson's correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 37°C culture. An Anova cut-off of 0.05 was applied to add an additional layer of statistical rigour. Using this method of analysis 84 proteins were deemed to change over time in culture. See Appendix B, Table 3.2.13 for the proteins that decrease over time and Table 3.2.14 for proteins that increase over time in culture. An example of the linear increase in the expression of a protein, semaphorin-3B, is shown in Figure 3.2.10. With a Pearson correlation score of 0.995 the tumour suppressor protein Semaphorin-3B increases in expression 8.9-fold between 24 and 120hrs post-temperature-shift at 37°C. Expression of the secreted protein Suprabasin is also shown to increase linearly between 24 and 120hrs post-temperature-shift at 37°C by 6.0-fold in Figure 3.2.11 (Pearson correlation 0.974, p-value 0.026).



Semaphorin-3B								
Pearson's correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest	Lowest
0.995	0.005	4	4	275	0.0114	8.9	120hrs	24hrs

**Figure 3.2.10** Progenesis label-free output of the tumour suppressor semaphorin-3B showing its increase in expression from 24hrs (Day1) to 120hrs (Day4) at 37°C (A). Output from excel on which Pearson correlation calculation is based (Pearson correlation 0.995, p-value 0.005). (B)

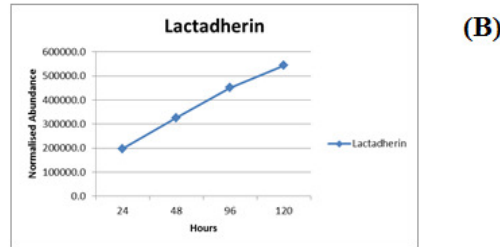


Suprabasin								
Pearson's correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest	Lowest
0.974	0.026	3	3	243	0.0017	6.0	120hrs	24hrs

**Figure 3.2.11** Progenesis label-free output of the secreted protein Suprabasin showing its increase in expression from 24hrs (Day1) to 120hrs (Day4) at 37°C (A). Output from excel on which Pearson correlation calculation is based (Pearson correlation 0.974, p-value 0.026). (B)

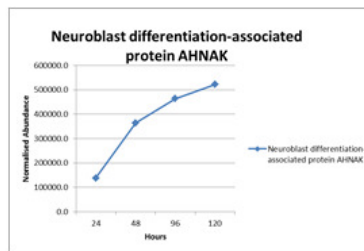
### **3.2.3.6 Pearson correlation analysis of differentially-regulated HCPs over 24, 72, 96 and 120 hours at 31°C**

Pearson's correlation analysis was also used as a means of measuring the linearity of protein expression over time in the temperature-shifted culture. This analysis identified that 28 proteins increase expression in a linear manner over time in culture (Appendix B, Table 3.2.15). Figure 3.2.12 and Figure 3.2.13 shows the increase in the secreted proteins Lactadherin (Pearson correlation 0.995, p-value 0.0048.) and Neuroblast differentiation-associated protein AHNAK (Pearson correlation 0.956, p-value 0.0442) over time in culture at 31°C.



Lactadherin								
Pearson's correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest	Lowest
0.997	0.0026	2	2	95	0.0024	2.8	120hrs	24hrs

**Figure 3.2.12** Progenesis label-free output of the secreted protein Lactadherin showing its increase in expression from 24hrs (Day1) to 120hrs (Day4) at 31°C (A). Output from excel on which Pearson correlation calculation is based (Pearson correlation 0.997, p-value 0.0026). (B)



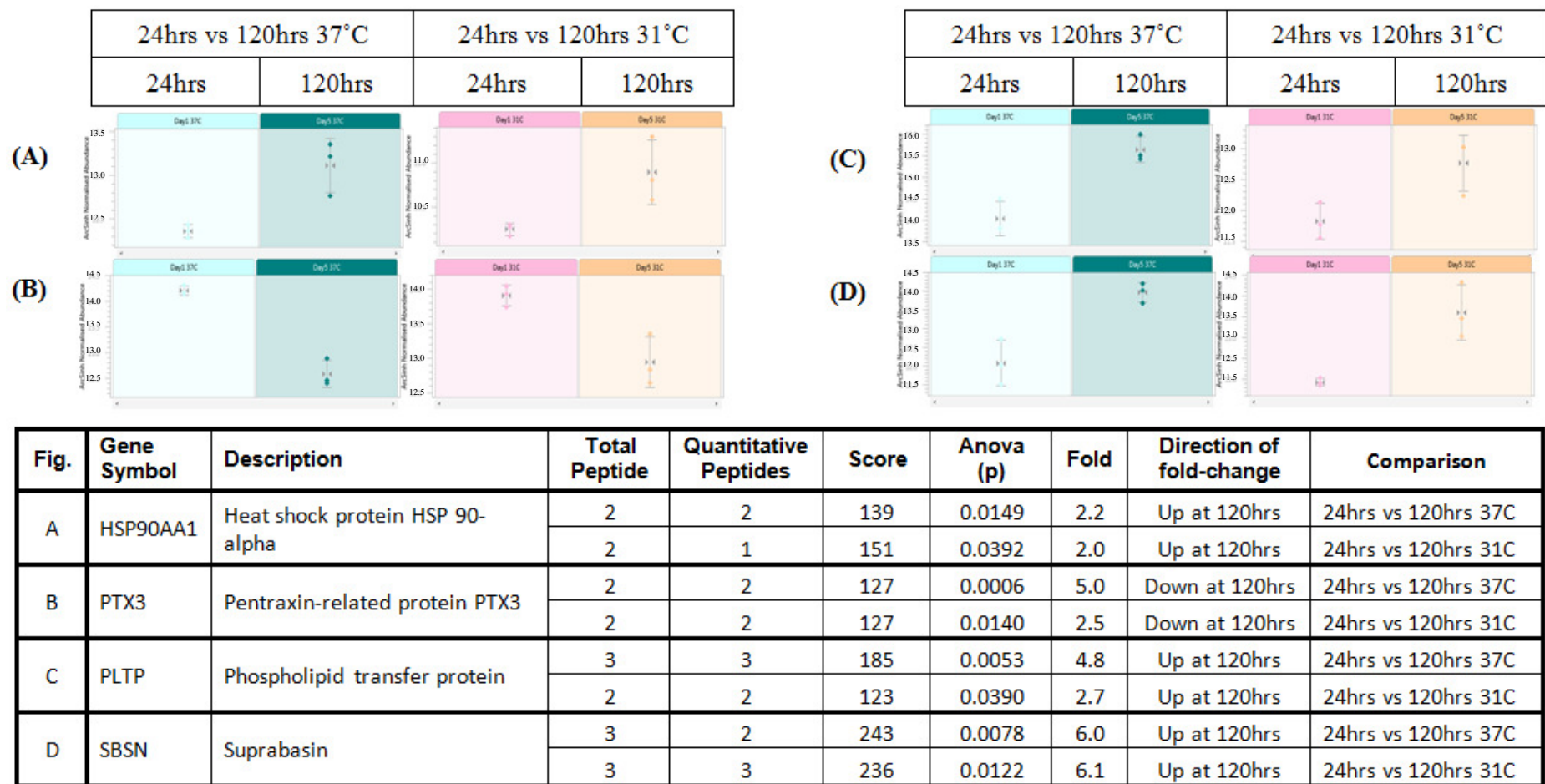
Neuroblast differentiation-associated protein AHNAK								
Pearson's correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest	Lowest
0.956	0.0442	9	9	486.0	1.78E-04	3.8	120hrs	24hrs

**Figure 3.2.13** Progenesis label-free output of the secreted protein Neuroblast differentiation-associated protein AHNAK showing its increase in expression from 24hrs (Day1) to 120hrs (Day4) at 31°C (A). Output from excel on which Pearson correlation calculation is based (Pearson correlation 0.956, p-value 0.0442) (B).

### **3.2.3.7 Comparison of differentially-regulated HCPs over time in culture at 37°C and 31°C in a non-producing CHO-K1 cell line**

This study has identified a significant number of differentially-expressed HCPs over time in culture under both temperature-shifted and non-temperature-shifted conditions (as detailed in sections 3.1.3.1 and 3.1.3.2). In addition, comparisons between 37°C and 31°C cultures at both early (24hrs) and late (120hrs) stage has provided further information on the differential expression of HCPs as a result of the change in culture temperature. In order to obtain a clearer picture of the differences in the HCP profile of both cultures data from 24hrs vs 120hrs at 37°C and 24hrs vs 120hrs at 31°C experiments were combined. In this way, the differential expression of a given HCP under temperature-shifted and non-temperature conditions could be directly compared (Figure 3.2.14). These proteins were not identified as differentially-expressed in the 37°C vs 31°C at 120hrs protein list, therefore it is most likely that the relative abundance of these proteins did not significantly change between temperature-shifted and non-temperature-shifted cultures at a late stage in culture.





**Figure 3.2.14** Shows Progenesis label-free software outputs and a table of proteins common to 24hrs vs 120hrs 37°C and 24hrs vs 120hrs 31°C differentially-expressed protein lists from non-producing CHO-K1 cell line.

As can be seen in Figure 3.2.14, in many instances the differentially-expressed protein trends in the same direction irrespective of the culture conditions. It may still be useful to know whether the relative expression of the HCP is greatest at 37°C or 31°C. For this reason, the list of proteins common to 24hrs vs 120hrs 37°C and 24hrs vs 120hrs 31°C lists were then compared to 37°C vs 31°C at 120hrs and the data from all common proteins combined. This would then provide an indication as to which culture had the greatest relative abundance of a given protein after 144hrs (Appendix B, Table 3.2.16). Information such as this could be useful for identifying culture conditions that accumulate the greatest concentration of a given protease or glycosidase and as such, increase the likelihood of product degradation through enzymatic activity.

To show how the abundance of some of HCPs identified change over time, a number of proteins were selected from Table 3.2.16 and, using the output from Progenesis label-free software as a visual representation, are shown in Figure 3.2.15 and Figure 3.2.16. Figure 3.2.15 shows the differential expression of the anti-apoptotic protein Endoplasmic reticulum chaperone protein (ERCP) and the protease Complement C1r-A subcomponent, while Figure 3.2.16 shows the secreted proteins semaphorin-3B and Basement membrane-specific heparan sulfate proteoglycan core protein (HSPG2). These proteins were selected as examples as they may impact on cell viability (Endoplasmic reticulum chaperone protein) or product quality (Complement C1r-A subcomponent). Knowledge of proteins actively secreted into the cell culture supernatant by CHO cells may also be useful as they are possibly involved in cell – cell signalling and so, are also included.

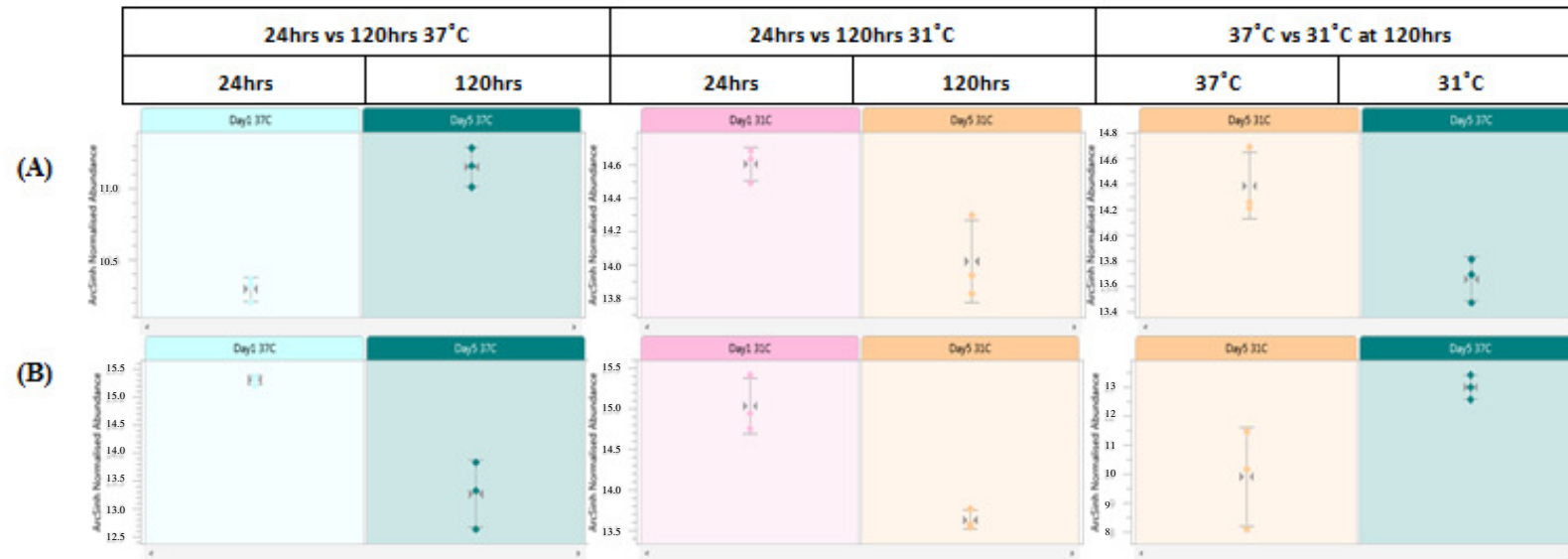
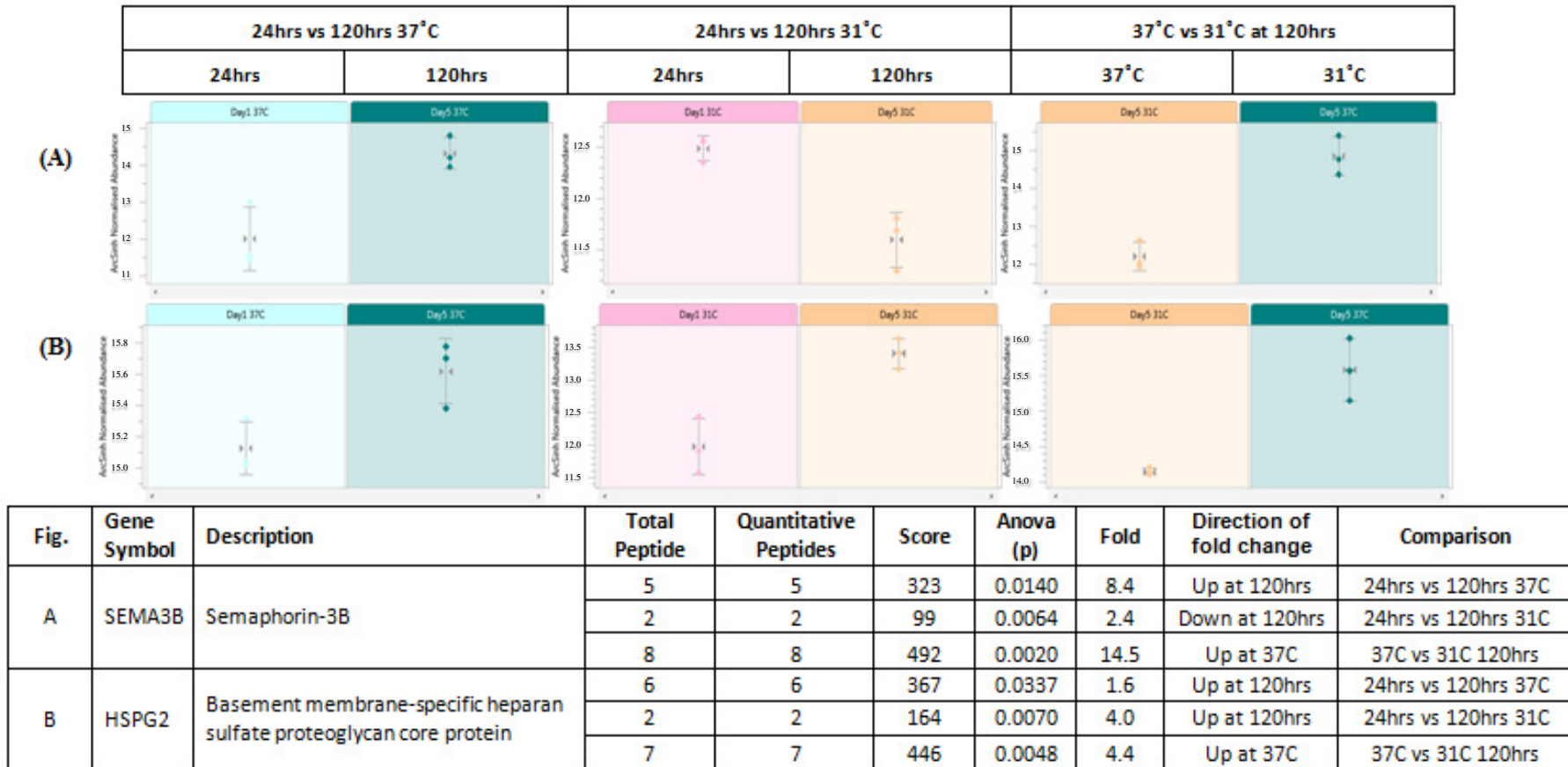


Fig.	Gene Symbol	Description	Total Peptide	Quantitative Peptides	Score	Anova (p)	Fold	Direction of fold change	Comparison
A	HSP90B1	Endoplasmin	2	2	104	0.0008	2.34	Up at 120hrs	24hrs vs 120hrs 37C
			2	2	131	0.0190	1.8	Down at 120hrs	24hrs vs 120hrs 31C
			3	3	188	0.0157	2.1	Up at 31C	37C vs 31C 120hrs
B	C1ra	Complement C1r-A subcomponent	3	3	179	0.0044	6.9	Down at 120hrs	24hrs vs 120hrs 37C
			3	3	179	0.0025	4.2	Down at 120hrs	24hrs vs 120hrs 31C
			1	1	47	0.0379	11.2	Down at 31C	37C vs 31C 120hrs

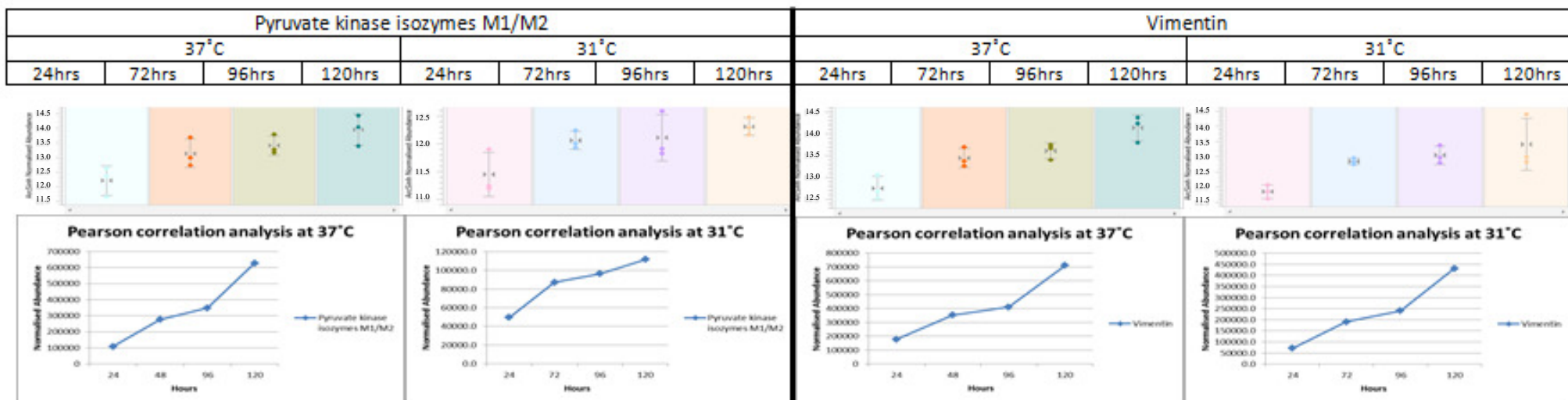
**Figure 3.2.15** Shows the differential expression of the HCPs (A) Endoplasmin and (B) Complement C1r-A subcomponent between early (24hrs) and late (120hrs) stages in culture at 37°C and 31°C. The differential expression between 37°C and 31°C at 120hrs is also shown.



**Figure 3.2.16** Shows the differential expression of the HCPs (A) Semaphorin-3B and (B) HSPG2 between early (24hrs) and late (120hrs) stages in culture at 37°C and 31°C. The differential expression between 37°C and 31°C at 120hrs is also shown.

Using Pearson correlations as described in sections 3.1.3.5 and 3.1.3.6, information was gathered on the linearity of the expression of proteins over time at 37°C and at 31°C. The lists generated were compared and any proteins common to both lists are shown in Appendix B, Table 3.2.17. To identify which temperature condition has the greatest overall abundance of a given protein the list of common proteins was compared to 37°C vs 31°C at 120hrs (Table 3.2.11 and Table 3.2.12). Based on this comparison, no proteins identified as being common to both 37°C and 31°C Pearson correlation lists were found to be differentially-expressed between 37°C vs 31°C at 120hrs apart from PKM which was found to be 4.6-fold up-regulated at 37°C compared to 31°C.

Pearson correlation plots and the outputs from Progenesis label-free software are used in Figure 3.2.21 to show the linear increase of the intracellular proteins (A) Pyruvate kinase isozymes M1/M2 and (B) Vimentin in the cell culture supernatant of CHO-K1 cells. This increase occurs over time in both temperature-shifted and non-temperature-shifted cultures. While it is possible that these proteins may be secreted via some non-classical secretory pathway, it is perhaps more likely that their presence in the media is as a result of cell lysis occurring over time. Information on the release of intracellular proteins is of interest as not only might they impact on culture viability or product quality, but their presence is also an additional burden to the purification of the final protein product.



Gene Symbol	Description	Pearson's Correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest at	Comparison
PMK	Pyruvate kinase isozymes M1/M2	0.9719	0.0281	3	3	213.6	0.0115	5.8	120hrs	24, 72, 96 and 120 hours at 37°C
		0.9559	0.0441	2	1	125.04	0.0427	2.3	120hrs	24, 72, 96 and 120 hours at 31°C
VIM	Vimentin	0.9650	0.0350	4	4	215.02	0.0010	4.0	120hrs	24, 72, 96 and 120 hours at 37°C
		0.9735	0.0265	2	2	94.03	0.0189	6.0	120hrs	24, 72, 96 and 120 hours at 31°C

**Figure 3.2.17** Pearson correlation plots and the outputs from Progenesis label-free software of the intracellular proteins Pyruvate kinase isozymes M1/M2 (A) and Vimentin (B) as they increase over time in culture under both temperature-shifted and non-temperature-shifted conditions.

### **3.2.3.8 Comparison of quantitative label-free LC/MS CHO-K1 HCP data normalised to cell number compared to normalisation by protein concentration**

Quantitative information on the differential expression in the cell culture supernatant of CHO-K1 cells (as described in sections 3.2.3.1-3.2.3.4) was obtained through the LC/MS analysis of equal concentrations of protein from the respective samples and applying a series of statistical cut-offs to the data obtained (as described in sections 3.2.3). A protein was considered to be differentially-regulated if its change in relative abundance was >1.5 fold between experimental groups. Another approach could consider normalising to cell numbers between the comparisons to take into account changes in cell concentration (both viable and non-viable) over time or under temperature-shifted and non-temperature-shifted conditions. It was decided that a correction factor would be applied to the lists of differentially-expressed proteins identified to facilitate normalisation of the data against cell numbers. This was achieved by taking the total number of cells per volume of media collected at each time point and calculating what this amounted to in HCP concentration in the conditioned media samples. The fold-change in protein concentration between experimental groups was then calculated and added to the 1.5-fold cut-off to establish a new threshold to determine differential expression. As can be seen in Table 3.2.18, applying the correction factor resulted in the loss of a small number of proteins from two comparisons (24hrs vs 120hrs at 37°C and 37°C vs 31°C at 24hrs). In the remaining two comparisons no correction factor was applied as there was no change in protein concentration between experimental groups. Due to the similar number of differentially-expressed proteins obtained with and without the correction factor applied (no change in two instances), it was decided that the 1.5-fold change was sufficiently robust as a cut-off for determining whether or not a protein was significantly differentially-regulated.

Comparison	Normalisation correction factor	Normalisation correction factor + 1.5-fold change cut-off	Number of differentially expressed proteins identified using correction factor	Number of differentially expressed proteins identified using 1.5-fold cut-off
24hrs vs 120hrs at 37°C	0.33	1.83	309	322
24hrs vs 120hrs at 31°C	0	1.50	146	146
37°C vs 31°C at 24hrs	0.27	1.77	190	211
37°C vs 31°C at 120hrs	0	1.50	228	228

**Table 3.2.18** The use of a correction factor was investigated to facilitate normalisation of the data by cell number (as opposed to normalisation by protein concentration). This table show a comparison of the number of differentially-expressed HCPs identified in the conditioned media of a CHO-K1 cell line with and without the correction factor applied.

### 3.2.4 Bioinformatic analysis of differentially-regulated protein lists

In order to determine what cellular compartment differentially-expressed proteins originated from, the gene symbol from each protein was searched using DAVID functional annotation software (<http://david.abcc.ncifcrf.gov/>).

DAVID uses gene ontology to classify the differentially-regulated proteins and identify enrichment of cellular locations where the proteins would typically be found. Protein lists generated from comparing 24hrs to 120hrs at both 37°C and 31°C were searched against DAVID using an adjusted p-value of  $\leq 0.05$  (Benjamini) as a statistical cut-off. Benjamini is a conservative correction technique to correct enrichment P-values to control gene family-wide false discovery rate under certain rate (e.g.  $\leq 0.05$ ). Based on the analysis of 24hrs versus 120hrs at 37°C using DAVID it is immediately obvious that proteins representing a wide variety of cellular components are released into the media of CHO-K1 cells (Table 3.2.20 and Figure 3.2.18). While there is some enrichment of extracellular proteins, it would appear that the majority are expected to be intracellularly located, their presence in the media possibly as a result of cell lysis or a reduction in cell viability. This becomes particularly apparent when compared to the list of enriched cell components obtained from analysing 24hrs versus 120hrs at 31°C where Table 3.2.21 indicates that the number of enriched entities is greatly reduced (Figure 3.2.19). This is possibly because of the high cell viability that was maintained over time in the 31°C culture. It is still perhaps worth noting that in both instances, the enriched extracellular components identified still fall well below the  $\leq 0.05$  Benjamini threshold set and that



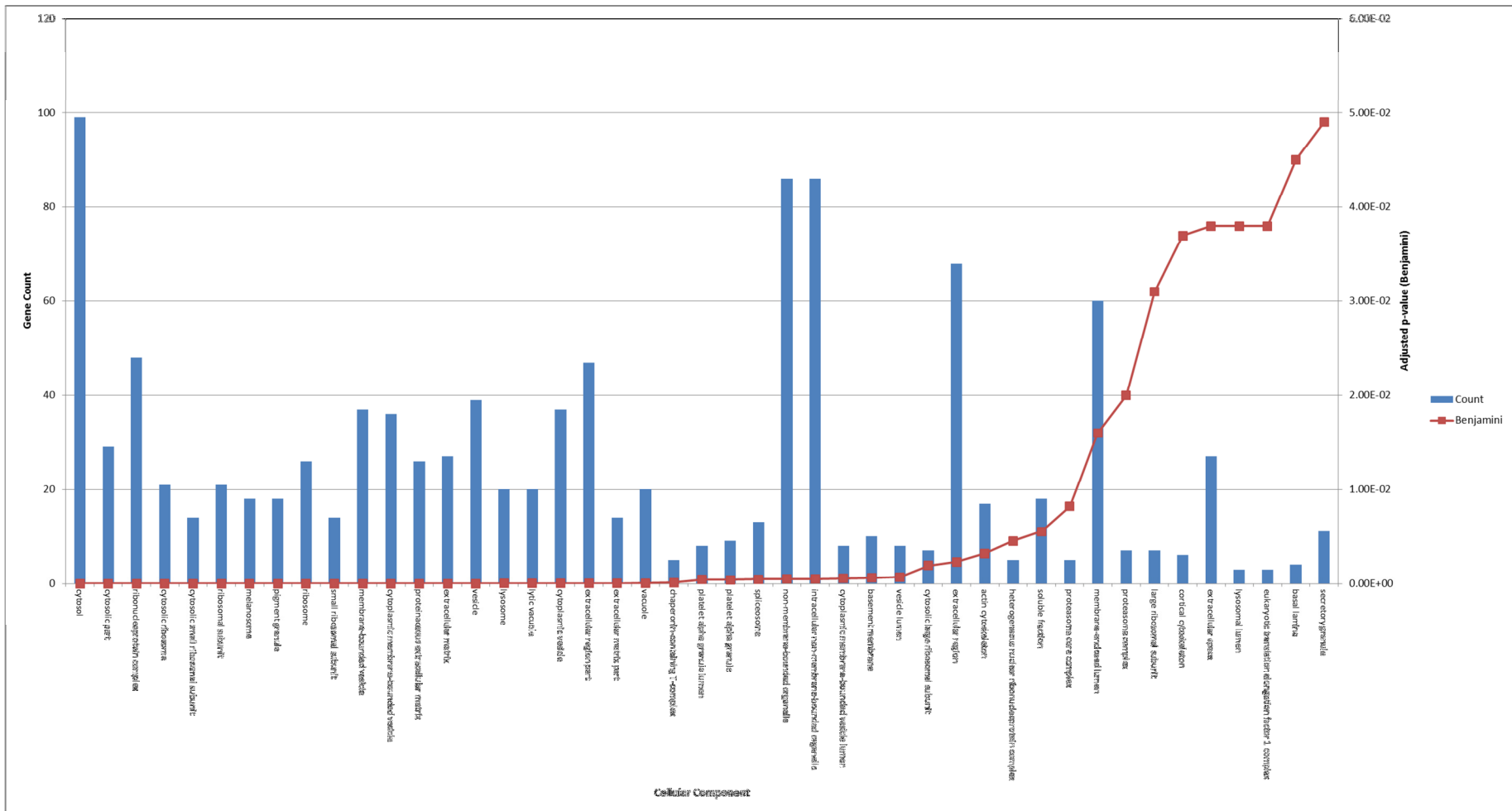
the Benjamini enrichment value is roughly the same both in 37°C and 31°C lists. Using the Uniprot protein database, all proteins classified as secreted in the lists of differentially-expressed HCPs in this study were identified (Table 3.2.19).

<b>Condition</b>	<b>Appendix Table</b>
Up at 120hrs when compared to 24hrs at 37°C	Table 3.2.22
Down at 120hrs when compared to 24hrs at 37°C	Table 3.2.23
Up at 120hrs when compared to 24hrs at 31°C	Table 3.2.24
Down at 120hrs when compared to 24hrs at 31°C	Table 3.2.25

**Table 3.2.19** List of tables containing quantitative information on proteins classified as secreted by the Uniprot protein database.

**Table 3.2.20** GO cellular component enrichment for differentially-expressed host cell proteins identified in the conditioned media of a CHO-K1 non-producing cell line at 37°C when 24hrs was compared to 120hrs (post-temperature-shift for a culture grown in parallel) analysed using quantitative label-free LC/MS. An adjusted p-value of  $\leq 0.05$  (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.

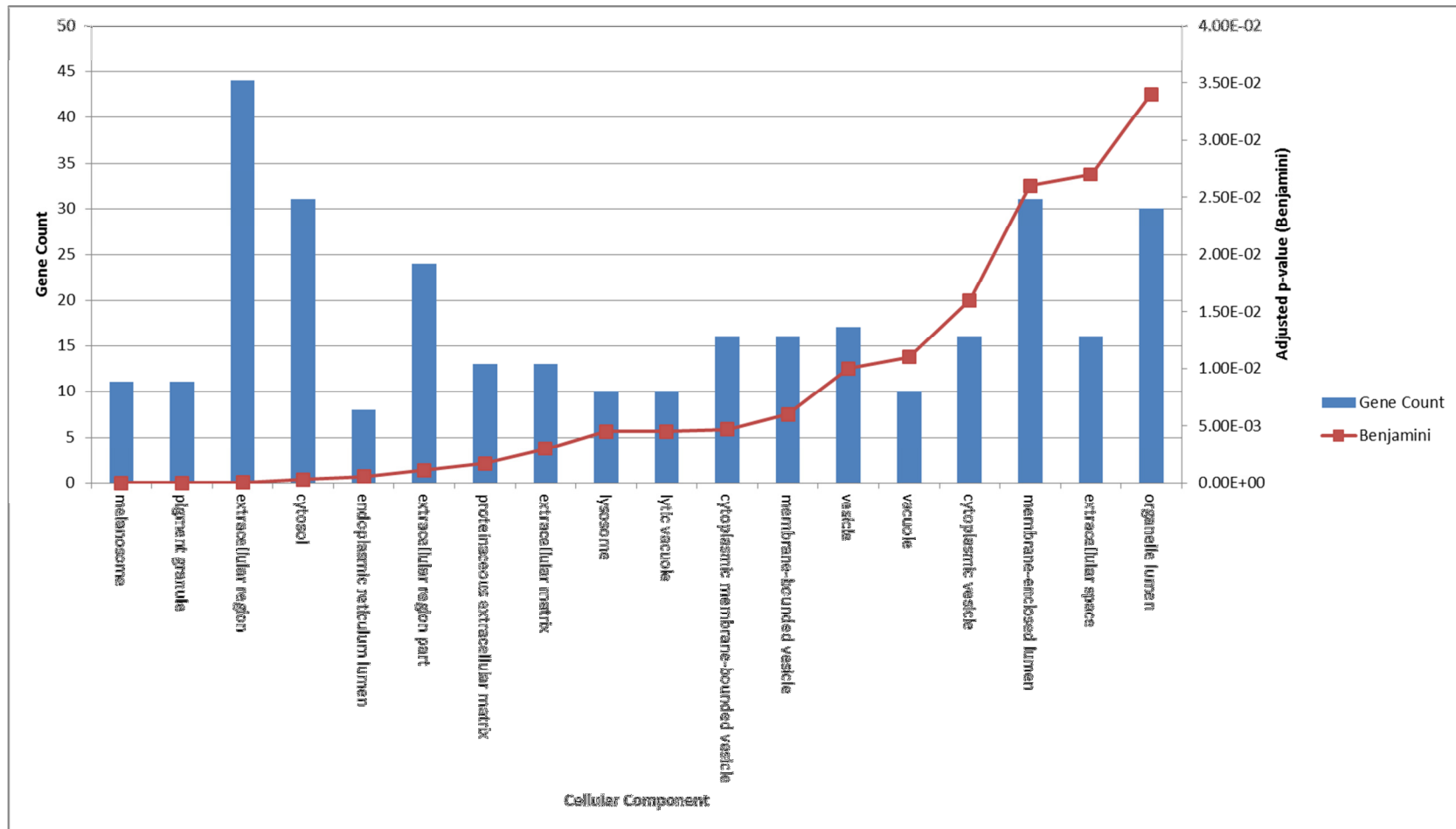
Cellular Component	Gene Count	P-Value	Benjamini
cytosol	99	3.60E-29	1.20E-26
cytosolic part	29	2.10E-18	3.40E-16
ribonucleoprotein complex	48	3.50E-17	3.80E-15
cytosolic ribosome	21	4.40E-16	3.60E-14
cytosolic small ribosomal subunit	14	1.40E-12	9.40E-11
ribosomal subunit	21	4.70E-12	2.60E-10
melanosome	18	7.00E-12	3.30E-10
pigment granule	18	7.00E-12	3.30E-10
ribosome	26	8.40E-12	3.50E-10
small ribosomal subunit	14	8.00E-10	2.90E-08
membrane-bounded vesicle	37	8.70E-09	2.80E-07
cytoplasmic membrane-bounded vesicle	36	1.30E-08	3.90E-07
proteinaceous extracellular matrix	26	3.70E-08	1.00E-06
extracellular matrix	27	4.00E-08	1.00E-06
vesicle	39	6.40E-08	1.50E-06
lysosome	20	1.90E-07	4.20E-06
lytic vacuole	20	1.90E-07	4.20E-06
cytoplasmic vesicle	37	1.90E-07	4.00E-06
extracellular region part	47	3.50E-07	6.60E-06
extracellular matrix part	14	1.60E-06	3.00E-05
vacuole	20	2.90E-06	4.90E-05
chaperonin-containing T-complex	5	7.50E-06	1.20E-04
platelet alpha granule lumen	8	2.70E-05	4.20E-04
platelet alpha granule	9	2.70E-05	4.10E-04
spliceosome	13	3.30E-05	4.60E-04
non-membrane-bounded organelle	86	3.70E-05	5.00E-04
intracellular non-membrane-bounded organelle	86	3.70E-05	5.00E-04
cytoplasmic membrane-bounded vesicle lumen	8	4.30E-05	5.60E-04
basement membrane	10	5.00E-05	6.20E-04
vesicle lumen	8	5.80E-05	7.00E-04
cytosolic large ribosomal subunit	7	1.60E-04	1.90E-03
extracellular region	68	2.00E-04	2.30E-03
actin cytoskeleton	17	2.90E-04	3.20E-03
heterogeneous nuclear ribonucleoprotein complex	5	4.30E-04	4.50E-03
soluble fraction	18	5.40E-04	5.50E-03
proteasome core complex	5	8.30E-04	8.20E-03
membrane-enclosed lumen	60	1.70E-03	1.60E-02
proteasome complex	7	2.10E-03	2.00E-02
large ribosomal subunit	7	3.40E-03	3.10E-02
cortical cytoskeleton	6	4.20E-03	3.70E-02
extracellular space	27	4.50E-03	3.80E-02
lysosomal lumen	3	4.60E-03	3.80E-02
eukaryotic translation elongation factor 1 complex	3	4.60E-03	3.80E-02
basal lamina	4	5.60E-03	4.50E-02
secretory granule	11	6.30E-03	4.90E-02



**Figure 3.2.18** Graphical representation of Table 3.2.20 depicting GO cellular component enrichment and their corresponding adjusted p-value of  $\leq 0.05$  (Benjamini) for differentially-expressed host cell proteins identified in the conditioned media of a CHO-K1 non-producing cell line at 37°C when 24hrs was compared to 120hrs (post-temperature-shift for a culture grown in parallel).

**Table 3.2.21** GO cellular component enrichment for differentially-expressed host cell proteins identified in the conditioned media of a CHO-K1 non-producing cell line at 31°C when 24hrs was compared to 120hrs (post-temperature-shift) analysed using quantitative label-free LC/MS. An adjusted p-value of  $\leq 0.05$  (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.

<b>Cellular Component</b>	<b>Gene Count</b>	<b>P-Value</b>	<b>Benjamini</b>
melanosome	11	9.90E-09	2.20E-06
pigment granule	11	9.90E-09	2.20E-06
extracellular region	44	5.10E-08	5.60E-06
cytosol	31	3.70E-06	2.70E-04
endoplasmic reticulum lumen	8	9.70E-06	5.30E-04
extracellular region part	24	2.50E-05	1.10E-03
proteinaceous extracellular matrix	13	4.60E-05	1.70E-03
extracellular matrix	13	9.50E-05	3.00E-03
lysosome	10	1.60E-04	4.50E-03
lytic vacuole	10	1.60E-04	4.50E-03
cytoplasmic membrane-bounded vesicle	16	1.90E-04	4.70E-03
membrane-bounded vesicle	16	2.70E-04	6.00E-03
vesicle	17	5.20E-04	1.00E-02
vacuole	10	6.10E-04	1.10E-02
cytoplasmic vesicle	16	9.80E-04	1.60E-02
membrane-enclosed lumen	31	1.60E-03	2.60E-02
extracellular space	16	1.90E-03	2.70E-02
organelle lumen	30	2.50E-03	3.40E-02



**Figure 3.2.19** Graphical representation of Table 3.2.21 depicting GO cellular component enrichment and their corresponding adjusted p-value of  $\leq 0.05$  (Benjamini) for differentially-expressed host cell proteins identified in the conditioned media of a CHO-K1 non-producing cell line at 31°C when 24hrs was compared to 120hrs (post-temperature-shift for a culture grown in parallel).

### **3.2.5 Qualitative analysis of proteins identified by quantitative label-free LC/MS profiling of conditioned media from a non-producing CHO-K1 cell line**

The use of Progenesis label-free software will only identify those proteins that are deemed to be differentially-expressed under a specific set of conditions. It does not provide information on those proteins whose relative abundance does not change. Knowledge of all proteins present in a proteome, not just those that are differentially-expressed, can provide useful information in aiding the identification of differences and similarities between proteomes under different environmental conditions.

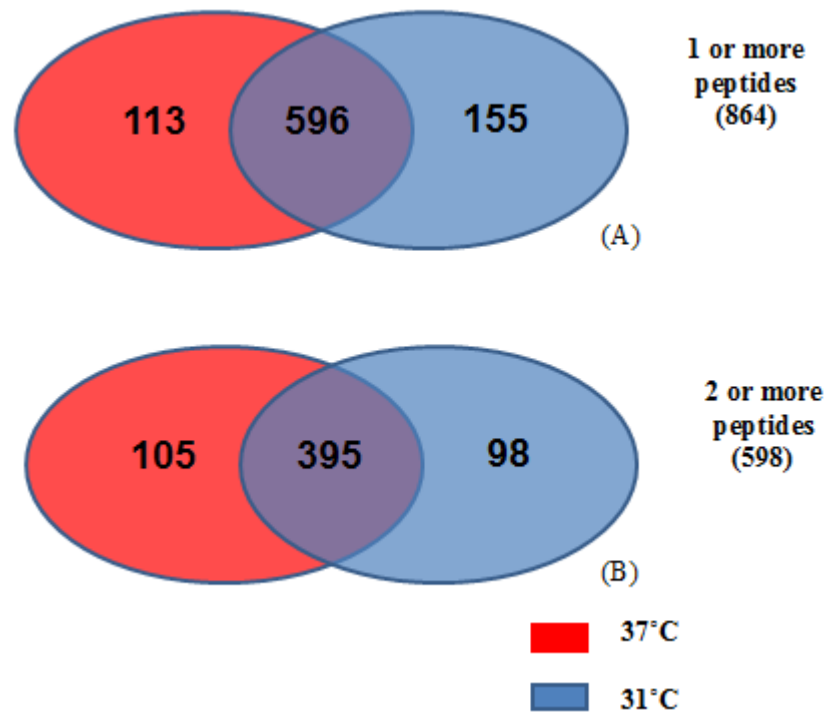
The RAW files obtained from LC/MS analysis of each conditioned media sample were searched against both the CHO NCBI and CHO BB databases using Mascot and SEQUEST search algorithms in a parallel algorithm search on Proteome Discoverer 1.4. Proteins were “identified” if they met the following criteria: SEQUEST Xcorr  $\geq 1.9/2.2/3.75$  (for charges +1/+2/+3), deltaCn  $\geq 0.1$  and/or a ion score of 40. In addition, in order for a protein to be considered “present” under a given condition, it must be identified by the Mass Spectrometer in a minimum of two biological samples. Samples were collected in biological triplicate and taken at 24, 72, 96, and 120 hours post-temperature-shift as per Figure 3.2.1.

The following comparisons were made by overlapping the lists of identified proteins:

- CHO-K1 37°C vs 31°C Cultures
  - Proteins identified with 1 or more peptides
  - Proteins identified with 2 or more peptides

### **3.2.5.1 Qualitative comparison of HCPs identified in temperature-shifted and non-temperature-shifted CHO-K1 cultures**

In order to determine whether the cell culture temperature impacted on HCP profile, qualitative lists of HCPs identified in 37°C and 31°C culture were compared. Direct comparison of these lists clearly shows that while there is a large number of proteins common to both culture conditions (596 proteins identified using a minimum of one peptide, 395 proteins identified using a minimum of two peptides) (Figure 3.2.20), there is also a subpopulation of proteins that are unique to each culture temperature. In this instance, 113 'one hit wonders' are identified at 37°C and 155 were identified at the lower culture temperature. While increasing the criteria for protein identification from one to a minimum of two peptides does slightly reduce the number of proteins unique to each condition (as expected), the numbers are still significant. 105 HCPs detected were unique to 37°C and 98 were only expressed at 31°C. This result indicates that cell culture temperature can impact on the HCP profile of a CHO-K1 cell line and that unique HCPs are expressed at lower culture temperatures. Lists of the proteins identified in Figure 3.2.20 A and Figure 3.2.20 B are available in Appendix B Table 3.2.26 - Table 3.2.30.



**Figure 3.2.20** Overlap of the list of proteins identified in the conditioned media of a non-producing CHO-K1 cell line grown at 37°C and 31°C over 168 hours in shake flask culture. Samples were collected in biological triplicate and taken at 24, 72, 96 and 120 hours post-temperature-shift. Venn diagram (A) represents those proteins identified using a minimum of 1 or more peptides. Venn diagram (B) is compiled from those proteins identified using a minimum of 2 or more peptides.



### 3.2.6 Summary

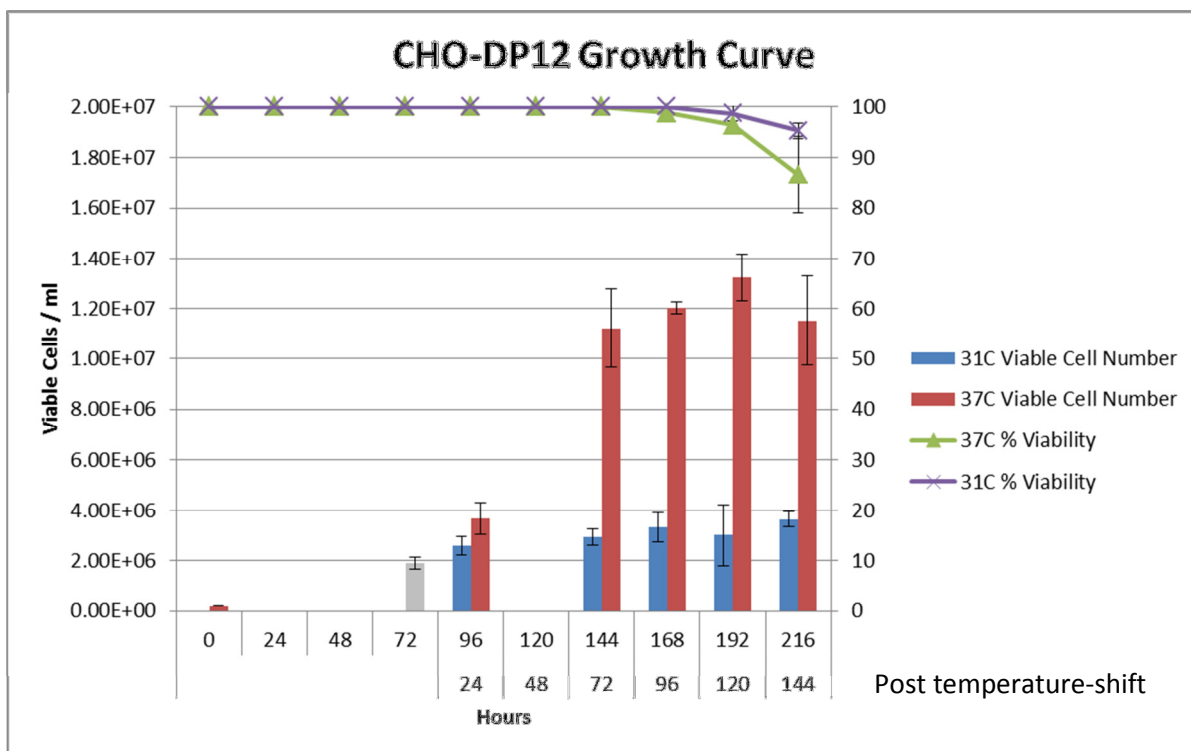
Quantitative label-free LC/MS analysis of conditioned media samples from a non-producing CHO cell line revealed that there was significant changes in the abundance of proteins over time in culture under both temperature-shifted and non-temperature-shifted conditions despite the cultures maintaining high viability. Proteins with the potential to negatively impact on product quality (proteases and glycosidases) and proteins with the potential to negatively impact on culture viability (apoptotic proteins) generally showed an increase in abundance at late stage compared to early stage culture in both temperature-shifted and non-temperature-shifted cultures. However, overall there was a reduction in the relative abundance of proteases, glycosidases and apoptotic proteins in the temperature-shifted culture compared to the non-temperature-shifted culture at 120hrs. Qualitative analysis of conditioned media samples from CHO-K1 cells revealed that while there was a lot of commonality between the HCPs found in a 37°C and a 31°C culture, there were also cohorts of HCPs that were unique to each culture temperature. In order to gather further information on the HCP profile of CHO cells and to confirm our initial findings, a second CHO cell line, this time an IgG producer, was also analysed by LC/MS as detailed in section 3.3.

### **3.3 Characterisation of the host cell protein profile of an IgG producing DP12 cell line**

In chapter 3.2 results from profiling the HCP content of conditioned media from a non-producing CHO-K1 cell line were detailed. This chapter will outline how an IgG depletion technique was coupled with a quantitative label-free LC/MS methodology to investigate the HCP profile of an IgG secreting DP12 CHO cell line. Similar to the CHO-K1 cell line, the change in the HCP profile under both temperature-shifted and non-temperature-shifted conditions over the course of the cell culture cycle, from early exponential to late stationary phase is examined. In addition to the use of the same quantitative label-free LC/MS based approach that was employed for CHO-K1 HCP analysis, the similar growth profile of the CHO-K1 and DP12 cell lines allowed for direct comparison of the results obtained from each data set as discussed later in the thesis.

#### **3.3.1 Growth curves at 31°C and 37°C of IgG producing DP12 cell line**

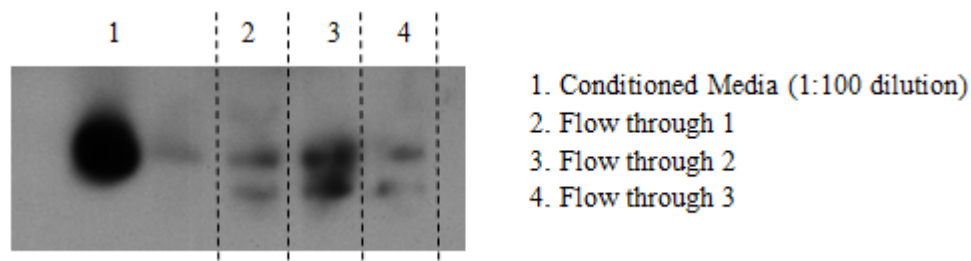
DP12 cells were grown in the same manner as the CHO-K1 cells (as detailed 3.2.1) with the exception that the temperature-shift of selected culture flasks took place 72hrs after inoculation. 72hrs was selected because the DP12 cells displayed an extended lag phase compared to the CHO-K1 cell line, however it still allowed a clear distinction to be drawn between temperature-shifted and non-temperature-shifted cultures from a biological perspective in terms of identifying a functional response to a change in the cell's environmental conditions.



**Figure 3.3.1** Growth curve of an IgG producing DP12 cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions. Temperature-shifted culture was placed at 31°C after 72hrs in culture. Cell culture supernatants were collected 24, 72, 96, 120 and 144 hours post-temperature-shift and analysed using a quantitative label-free mass spectrometry based approach.

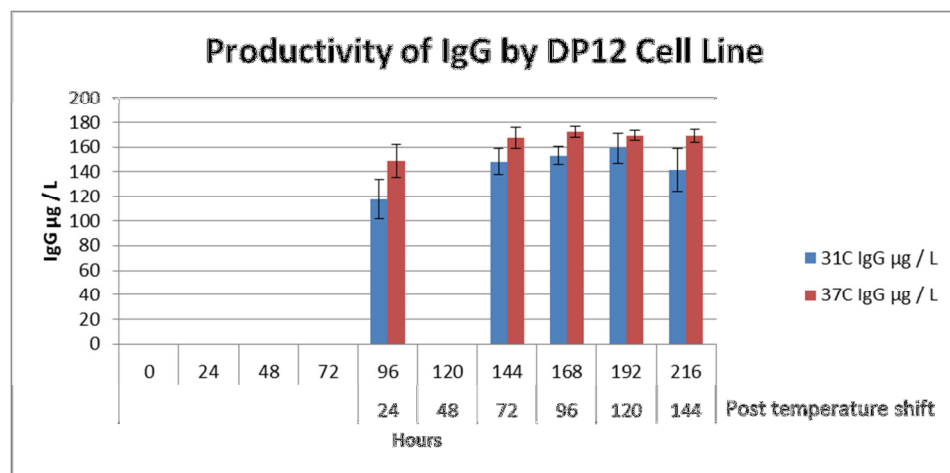
### 3.3.2 IgG depletion of conditioned media from IgG producing DP12 CHO cell line

In the analysis of cell culture supernatant by LC/MS relatively high levels of IgG (compared to the overall concentration of the other HCPs) could present a number of problems (i) high levels of IgG could potentially mask the presence of low abundant proteins in the sample when analysed by mass spectrometry (ii) differences in IgG concentration between samples would adversely affect quantitative results obtained; a basic requirement of accurate quantitative label-free LC/MS is that equal concentrations of protein are analysed in each sample, the concentration of an actively secreted IgG molecule would change both over time and under different culture conditions therefore altering the relative concentration of HCPs in each sample. In order to circumvent these problems, the use of Protein A chromatography to reduce overall IgG abundance and normalise IgG concentration between all conditioned media samples was examined. As can be seen in Figure 3.3.2, Protein A chromatography significantly reduces the levels of IgG in the cell culture supernatant of the DP12 cell line.

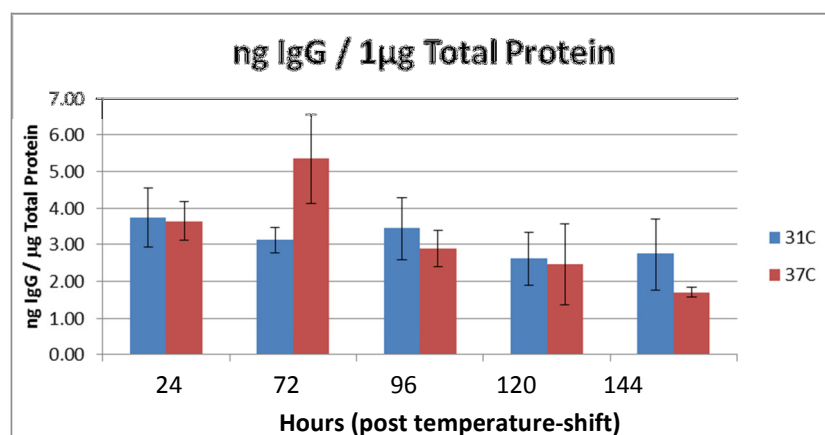


**Figure 3.3.2** Western blot image of IgG secreted by a DP12 cell line. Lane 1 contains conditioned media pre-protein A chromatography diluted 1:100. Lane 2 contains conditioned media that was passed through a Protein A column once (flow through 1). Lane 3 is the conditioned media from flow through 1 that has been passed through the same Protein A column for a second time (flow through 2). Lane 4 is flow through 2 after it has been passed through the Protein A column for a third time (flow through 3). Note that equal volumes of conditioned media sample were loaded onto a 1D-gel and immunoblotted for IgG.

Although it was expected that IgG concentration would increase over time in culture, this does not appear to be the case (Figure 3.3.3). IgG concentration remains relatively constant over time between the two culture conditions. Nonetheless, IgG depletion was carried out to ensure a reduction in the overall concentration of IgG and to further reduce any difference in IgG concentration between samples (Figure 3.3.4). Cell culture supernatants from three biological replicates at each time point under both temperature-shifted and non-temperature-shifted conditions were collected and filtered as detailed methods 2.3.2.



**Figure 3.3.3** Concentration of IgG produced by a DP12 cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions was determined by ELISA assay. Temperature-shifted culture was placed at 31°C after 72hrs in culture. Cell culture supernatants were collected 24, 72, 96, 120 and 144 hours post-temperature-shift.



**Figure 3.3.4** Concentration of IgG remaining in conditioned media sample from a DP12 cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions post IgG depletion by Protein A chromatography as determined by ELISA assay.

### **3.3.3 Quantitative label-free LC/MS analysis of conditioned media from a DP12 CHO cell line**

IgG depleted cell culture supernatants from three biological replicates at each time point under both temperature-shifted and non-temperature-shifted conditions were concentrated and prepared for LC/MS analysis as detailed methods 2.3.3. Prior to LC/MS analysis 4 pmoles of Cytochrome C was added to each sample to act as an internal control. After mass spectral analysis of all the samples was complete, inclusion lists were generated (as discussed 3.2.2).

Due to the complex sampling nature of the experiment i.e. multiple time points between two conditions, it was decided that the data would be analysed as follows:

- 24hrs vs 144hrs at 37°C
- 24hrs vs 144hrs at 31°C
- 37°C vs 31°C at 24hrs
- 37°C vs 31°C at 144hrs
- Analysis of differential regulation over 24, 72, 96, 120 and 144 hours at 37°C using Pearson correlation analysis
- Analysis of differential regulation over 24, 72, 96, 120 and 144 hours at 31°C using Pearson correlation analysis

Note: The time in hours above is with respect to temperature-shift. i.e. 24hrs is 24hrs post-temperature-shift.

Associating biological function with differentially-regulated proteins identified in this study was carried out as detailed 3.2.3.

### **3.3.4 Data analysis of differentially-regulated proteins identified by quantitative label-free LC/MS profiling of conditioned media from a DP12 CHO cell line**

Progenesis label-free software was used to analyse the data acquired following LC/MS analysis. A protein was considered to be differentially-regulated if it met the criteria as set out in 3.1.3. Although IgG levels in the conditioned media samples had been significantly reduced (Figure 3.3.4), there was still the possibility that there may not have been uniform reduction of IgG between the samples, and therefore accurate protein quantification may be distorted.

To overcome this, an internal standard, Cytochrome C was added to each sample prior to LC/MS analysis. The RAW files were initially normalised to all features, as would normally be the case for any label-free experiment. Normalisation is carried out by the software to correct for experimental variations such as deviations in the quantity of protein loaded onto the LC/MS system or differences in ionization. By default the software will normalise all the data to the peptides that don't change in abundance.

Following this analysis, the files were then normalised to IgG and the data was reanalysed. Finally the RAW files were normalised to the internal standard, Cytochrome C and the data was analysed for a third time.

The three sets of data were then compared in terms of proteins identified, number of peptides identified, fold-change, direction of change and ANOVA score. Results from this analysis indicated that when the lists of differentially-expressed proteins were compared, the files normalised to IgG was too dissimilar to the other means of normalisation and indicated that IgG could not be used as a feature to normalise the other proteins to. Although IgG had been greatly reduced in the media, the mass spectrometer would still be capable of detecting changes in the abundance of IgG between the samples.

While the overall abundance of IgG would not have been enough to interfere with LC/MS analysis, even subtle changes in the relative abundance of IgG would make it unsuitable for normalisation. As can be seen from a direct comparison between normalisation using an internal standard (Table 3.3.1) or normalising to proteins that did not change in expression (Table 3.3.2) there is great similarity in the results obtained both in terms of proteins identified, fold-change and their direction.



DP12 24hrs vs 144hrs 37C All Features Normalised to Cytochrome C												
Peptide		Peptide Quant		Score		Anova		Fold		Up in		Protein
C	N	C	N	C	N	C	N	C	N	C	N	
<b>13</b>	11	<b>13</b>	11	<b>763</b>	665	<b>0.00178</b>	0.00053	<b>16.3</b>	10.7	<b>24hrs</b>	24hrs	Complement C3
<b>12</b>	9	<b>11</b>	9	<b>895</b>	719	<b>0.01068</b>	0.02745	<b>2.9</b>	2.1	<b>24hrs</b>	24hrs	laminin subunit alpha-5 isoform 1
<b>11</b>	11	<b>9</b>	9	<b>617</b>	617	<b>7.29E-05</b>	1.14E-05	<b>8.3</b>	12.1	<b>144hrs</b>	144hrs	vimentin
<b>10</b>	10	<b>9</b>	9	<b>530</b>	530	<b>0.01183</b>	0.00311	<b>5.6</b>	3.5	<b>24hrs</b>	24hrs	Thrombospondin-1
<b>10</b>	1	<b>10</b>	1	<b>567</b>	59	<b>0.00771</b>	0.03689	<b>1.7</b>	2.7	<b>24hrs</b>	144hrs	Inter-alpha-trypsin inhibitor heavy chain H5
<b>9</b>	10	<b>9</b>	9	<b>479</b>	539	<b>0.00043</b>	0.00041	<b>5.6</b>	3.8	<b>24hrs</b>	24hrs	laminin subunit beta-1
<b>9</b>	9	<b>7</b>	7	<b>437</b>	437	<b>1.37E-05</b>	0.00038	<b>8</b>	11.1	<b>144hrs</b>	144hrs	endoplasmic
<b>9</b>	9	<b>9</b>	9	<b>540</b>	540	<b>0.0005</b>	0.00061	<b>4.3</b>	6.1	<b>144hrs</b>	144hrs	transketolase-like
<b>9</b>	9	<b>8</b>	8	<b>535</b>	535	<b>0.00205</b>	0.00188	<b>4.7</b>	6.7	<b>144hrs</b>	144hrs	Prostaglandin F2 receptor negative regulator
<b>8</b>	8	<b>6</b>	6	<b>405</b>	405	<b>0.00096</b>	1.08E-05	<b>6.1</b>	9.2	<b>144hrs</b>	144hrs	14-3-3 protein epsilon

N = All Feature Normalisation  
C = Cytochrome C normalisation

**Table 3.3.1** Top 10 differentially-regulated proteins, based on peptide number, when all features of the protein profile from 24hrs in a DP12 culture was compared to 144hrs and normalised to an internal standard, Cytochrome C. These proteins were then retrieved from the same comparison but where all features were normalised to proteins that did not change in expression. This allowed for direct comparison to be made between Cytochrome C Normalisation (C) and the Normal means of Normalisation (N), based on peptide count, ion score, ANOVA, fold-change and what condition the proteins is most highly expressed in.

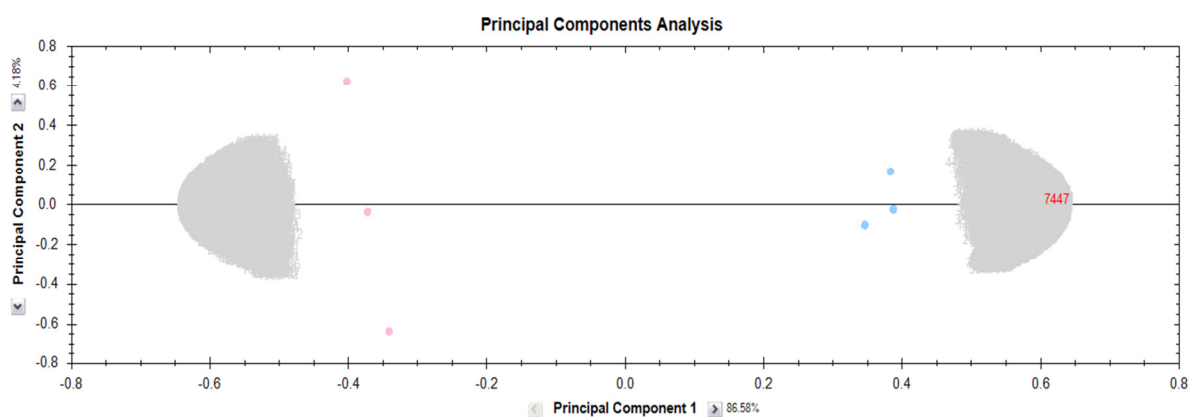
DP12 24hrs vs 144hrs 37C All Features Normalised to Features with no Fold Change												
Peptide		Peptide Quant		Score		Anova		Fold		Up in		Protein
N	C	N	C	N	C	N	C	N	C	N	C	
<b>11</b>	11	<b>9</b>	9	<b>617</b>	617	<b>1.14E-05</b>	7.29E-05	<b>12.1</b>	8.3	<b>144hrs</b>	144hrs	vimentin
<b>11</b>	13	<b>11</b>	13	<b>665</b>	763	<b>0.00053</b>	0.00178	<b>10.7</b>	16.3	<b>24hrs</b>	24hrs	Complement C3
<b>10</b>	10	<b>9</b>	9	<b>530</b>	530	<b>0.00311</b>	0.01183	<b>3.5</b>	5.6	<b>24hrs</b>	24hrs	Thrombospondin-1
<b>10</b>	9	<b>9</b>	9	<b>539</b>	479	<b>0.00041</b>	0.00043	<b>3.8</b>	5.6	<b>24hrs</b>	24hrs	Laminin subunit beta-1
<b>9</b>	12	<b>9</b>	11	<b>719</b>	895	<b>0.02745</b>	0.01068	<b>2.1</b>	2.9	<b>24hrs</b>	24hrs	laminin subunit alpha-5 isoform 1
<b>9</b>	9	<b>7</b>	7	<b>437</b>	437	<b>0.00038</b>	1.37E-05	<b>11.1</b>	8	<b>144hrs</b>	144hrs	endoplasmic
<b>9</b>	9	<b>9</b>	9	<b>540</b>	540	<b>0.00061</b>	0.0005	<b>6.1</b>	4.3	<b>144hrs</b>	144hrs	transketolase-like
<b>9</b>	9	<b>8</b>	8	<b>535</b>	535	<b>0.00188</b>	0.00205	<b>6.7</b>	4.7	<b>144hrs</b>	144hrs	Prostaglandin F2 receptor negative regulator
<b>8</b>	6	<b>7</b>	5	<b>430</b>	310	<b>0.01328</b>	0.12707	<b>2.7</b>	1.4	<b>144hrs</b>	144hrs	peroxidase homolog
<b>8</b>	7	<b>4</b>	4	<b>459</b>	416	<b>8.75E-06</b>	0.00019	<b>20.7</b>	13.9	<b>144hrs</b>	144hrs	14-3-3 protein eta

N = All Feature Normalisation  
C = Cytochrome C normalisation

**Table 3.3.2** Top 10 differentially-regulated proteins, based on peptide number, when all features of the protein profile from 24hrs in a DP12 culture was compared to 144hrs and normalised to proteins that did not change in expression. These proteins were then retrieved from the same comparison but where all features were normalised to an internal standard, Cytochrome C. This allowed for direct comparison to be made between the Normal means of Normalisation (N) and Cytochrome C Normalisation (C), based on peptide count, ion score, ANOVA, fold-change and what condition the proteins is most highly expressed in.

### 3.3.4.1 Differential regulation of HCPs 24hrs vs 144hrs at 37°C

A principal component analysis (PCA) plot generated by Progenesis during data analysis showing how the samples from 24hr and 144hr cultures at 37°C cluster is displayed in Figure 3.3.5.



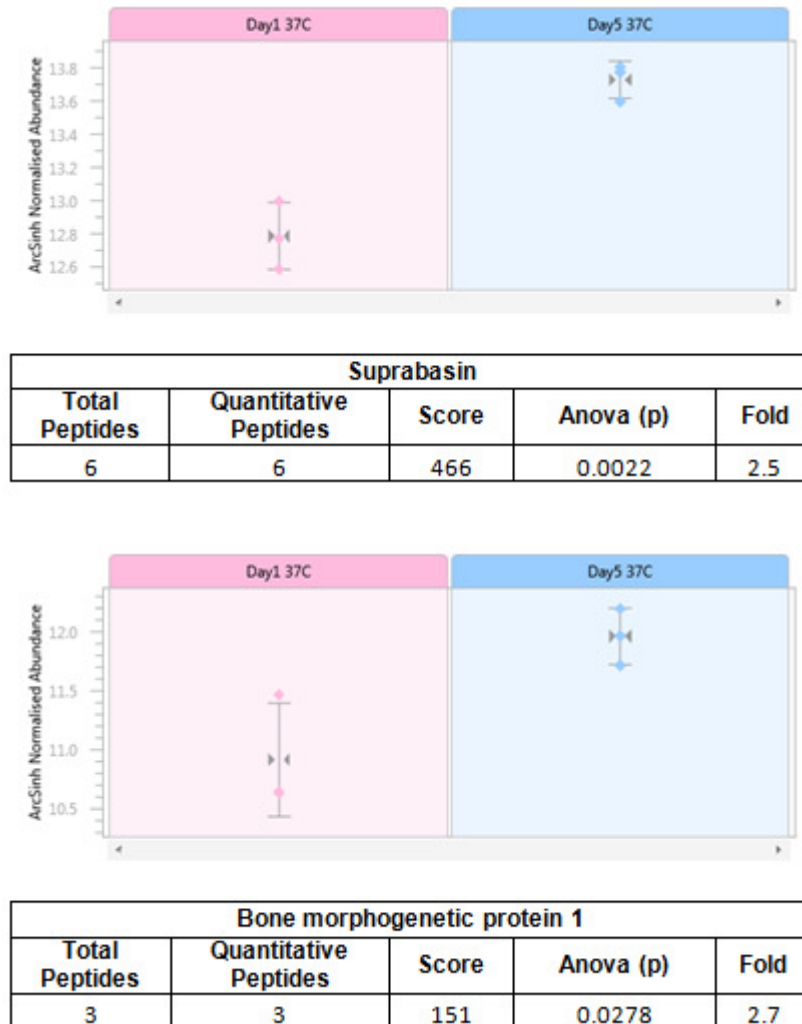
**Figure 3.3.5** PCA plot for the IgG secreting DP12 conditioned media sample groups 24hrs (pink) and 144hrs (blue) at 37°C generated during quantitative label-free LC/MS data analysis using Progenesis software.

Of the 375 proteins identified with 1 or more peptides as being differentially-expressed at 37°C between early (24hrs) and late (144hrs), 335 had increased expression at 144hr, while 40 were down-regulated (Summary Table 3.3.3). Similar to 24hrs compared to 120hrs at 37°C in the CHO-K1 cell line (section 3.2.3.2), there is an increase in the abundance of apoptotic proteins at late stage culture (Table 3.3.3). This is likely to be as a result of a decrease in culture viability at 144hrs in the DP12 cell line. Also, the majority of proteases and glycosidases, both of which have the potential to negatively impact product quality, are up regulated at the later time point. (Full lists of differentially-expressed proteins are provided in Appendix C, Table 3.3.4 and Table 3.3.5)

**Table 3.3.3** Number of differentially-expressed proteins at 24hrs compared to 144hrs at 37°C in an IgG secreting DP12 cell culture supernatant. Using Uniprot protein database, proteins of biological relevance have been identified and the number of differentially-expressed proteins under the relevant classification included in the table.

	24hrs vs 144hrs at 37°C					
	Total		Up at 144hrs		Down at 144hrs	
	≥1 Peptide	≥2 Peptides	≥1 Peptide	≥2 Peptides	≥1 Peptide	≥2 Peptides
<b>Protein Count</b>	375	162	335	143	40	19
<b>Apoptosis</b>	8	2	8	2	0	0
	GNB2L1, CLU, HMOX1, SHISA5, GAPDH, GAPDH, APP, MAPK1	GNB2L1, CLU	GNB2L1, CLU, HMOX1, SHISA5, GAPDH, GAPDH, APP, MAPK1	GNB2L1, CLU		
<b>Protease</b>	25	9	17	6	8	3
	MMP9, HTRA1, PSMB6, BMP1, PSMA7, CTSB, C1ra, PSMA5, ADAM10, CTSL1, TPP1, ADAMTS1, PLAT, P15156, PSMA3, PSMB1, PSMB8, THOP1, CNDP2, PSMB2, PARK7, ADAM17, PSMB3, PSMA6, AGA	MMP9, HTRA1, PSMB6, BMP1, PSMA7, CTSB, C1ra, PSMA5, ADAM10	PSMB6, BMP1, PSMA7, CTSB, PSMA5, ADAM10, PSMA3, PSMB1, PSMB8, THOP1, CNDP2, PSMB2, PARK7, ADAM17, PSMB3, PSMA6, AGA	PSMB6, BMP1, PSMA7, CTSB, PSMA5, ADAM10	MMP9, HTRA1, C1ra, CTSL1, TPP1, ADAMTS1, PLAT, P15156	MMP9, HTRA1, C1ra
<b>Glycosidase</b>	12	4	9	2	3	2
	GLB1, GAA, HEX, B GBA, GUSB, HEXA, GLA, IDUA, GANAB, MAN2A1, NAGLU, MAN1A1	GLB1, GAA, HEXB, GBA	GAA, GBA, HEXA, GLA, IDUA, GANAB, MAN2A1, NAGLU, MAN1A1	GAA, GBA	GLB1, HEXB, GUSB	GLB1, HEXB

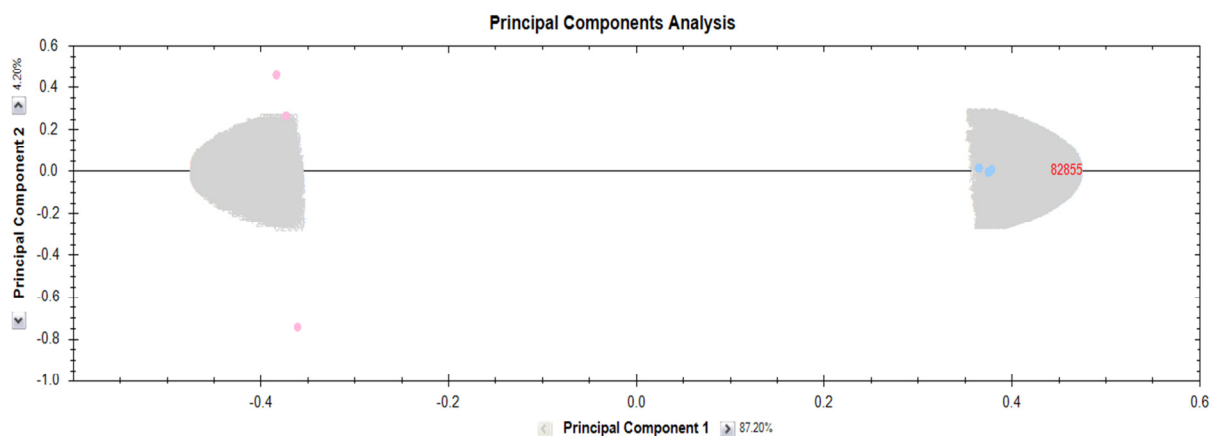
Examples of some of the differentially-regulated proteins identified in this analysis shows the up-regulation of the secreted protein Suprabasin (2.5-fold) and the protease Bone morphogenetic protein 1 (2.7-fold) (Figure 3.3.6).



**Figure 3.3.6** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the secreted protein Suprabasin and the growth factor Bone morphogenetic protein 1, identified as being differentially-expressed between 24hrs and 120hrs at 37°C in the conditioned media of a DP12 cell line.

### 3.3.4.2 Differential regulation of HCPs 24hrs vs 144hrs at 31°C

A PCA plot generated by Progenesis during data analysis illustrating how the samples from 24hr and 144hr cultures at 31°C cluster is displayed in Figure 3.3.7.



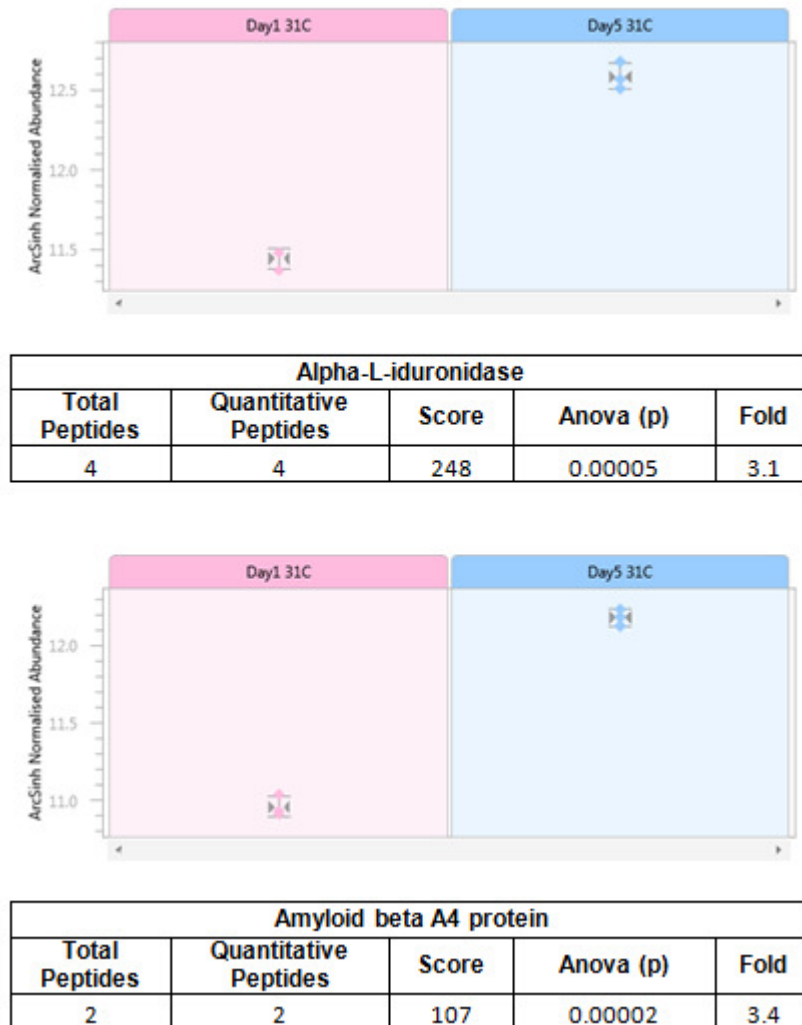
**Figure 3.3.7** PCA plot for the IgG secreting DP12 conditioned media sample groups 24hrs (pink) and 144hrs (blue) at 31°C generated during quantitative label-free LC/MS data analysis using Progenesis software.

When early (24hrs) and late (144hrs) phases of the temperature-shifted (31°C) were compared 179 were found to change in relative abundance (168 up regulated and 11 down-regulated at 144hrs with 2 or more peptides (Table 3.3.6). Similar to the 37°C culture, the majority of proteases and glycosidases are up-regulated at the later time point. Interestingly, despite the increased viability of the 31°C, there is a significant increase in the number of apoptotic related proteins at 144hrs in the 31°C compared to the 37°C (Table 3.3.6). (Full lists of differentially-expressed proteins are provided in Appendix C, Table 3.3.7 and Table 3.3.8).

**Table 3.3.6** Number of differentially-expressed proteins at 24hrs compared to 144hrs at 31°C in an IgG secreting DP12 cell culture supernatant. Using Uniprot protein database, proteins of biological relevance have been identified and the number of differentially-expressed proteins under the relevant classification included in the table.

	24hrs vs 144hrs at 31°C					
	Total		Up at 144hrs		Down at 144hrs	
	≥1 Peptide	≥2 Peptides	≥1 Peptide	≥2 Peptides	≥1 Peptide	≥2 Peptides
<b>Protein Count</b>	391	179	361	168	30	11
	7	4	7	4	0	0
<b>Apoptosis</b>	GNB2L1, HMOX1, Cyc,s APP, DNASE2, GAPDH, GAPDH	GNB2L1, HMOX1, Cycs, APP	GNB2L1, HMOX1, Cycs, APP, DNASE2, GAPDH, GAPDH	GNB2L1, HMOX1, Cycs, APP		
	24	11	19	9	5	2
<b>Protease</b>	C1ra, MMP9, PSMB1, PSMA6, PSMA3, PSMB6, DPP7, CTSB, CTSA, SCPEP1, P15156, PRCP, ADAM10, MMP19, PSMA1, PSMB2, AGA, PARK7, LAP3, PSMA7, CTSZ, PSMA5, CTSD, CNDP2	C1ra, MMP9, PSMB1, PSMA6, PSMA3, PSMB6, DPP7, CTSB, CTSA, SCPEP1, P15156	PSMB1, PSMA6, PSMA3, PSMB6, DPP7, CTSB, CTSA, SCPEP1, P15156, PSMA1, PSMB2, AGA, PARK7, LAP3, PSMA7, CTSZ, PSMA5, CTSD, CNDP2	PSMB1, PSMA6, PSMA3, PSMB6, DPP7, CTSB, CTSA, SCPEP1, P15156	C1ra, MMP9, PRCP, ADAM10, MMP19	C1ra, MMP9
	11	6	9	6	2	0
<b>Glycosidase</b>	IDUA, NAGLU, GAA, GLA, GBA, FUCA1, MAN2A1, GUSB, GANAB, NAGA, GLB1	IDUA, NAGLU, GAA, GLA, GBA, FUCA1	IDUA, NAGLU, GAA, GLA, GBA, FUCA1, GANAB, NAGA, GLB1	IDUA, NAGLU, GAA, GLA, GBA , CA1	MAN2A1, GUSB	

The Progenesis output displaying the up-regulated glycosidase Alpha-L-iduronidase (3.1-fold) and the apoptotic protein Amyloid beta A4 protein (3.4-fold) are shown in Figure 3.3.8.

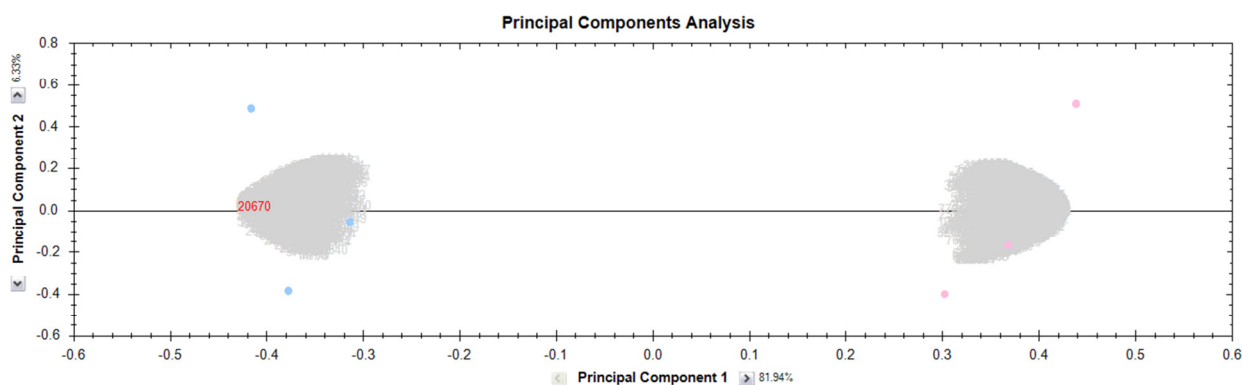


**Figure 3.3.8** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the glycosidase Alpha-L-iduronidase and the apoptotic protein Amyloid beta A4 protein, identified as being up-regulated between 24hrs and 120hrs at 31°C in the conditioned media of a DP12 cell line.



### 3.3.4.3 Differential regulation of HCPs at 24hrs 37°C vs 31°C

A PCA plot generated by Progenesis during data analysis indicating how the samples from 31°C and 37°C cultures at 24hrs cluster is displayed in Figure 3.3.9.



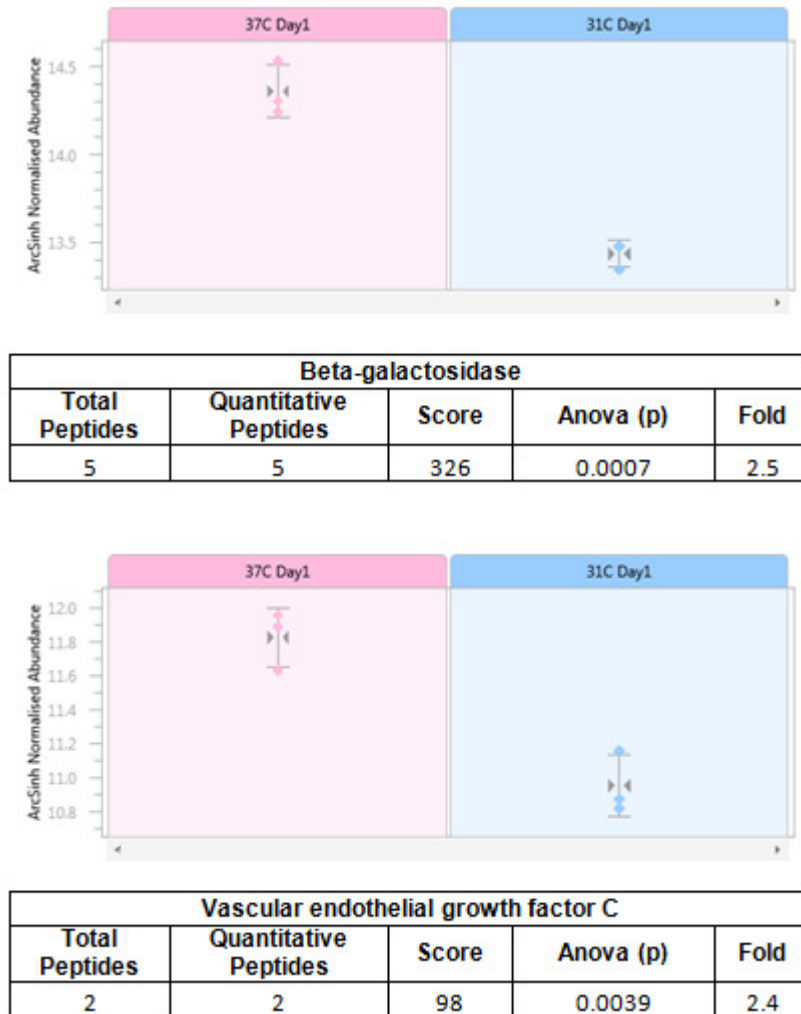
**Figure 3.3.9** PCA plot for the IgG secreting DP12 conditioned media sample groups 37°C (pink) and 31°C (blue) after 24hrs in culture generated during quantitative label-free LC/MS data analysis using Progenesis software

A comparison of 37°C and 31°C cultures 24hrs post culture temperature reduction revealed 150 HCPs with 1 or more peptides to be differentially-expressed between the two temperatures. Of these, 107 were down-regulated at 31°C, while the remaining 43 were up regulated (Table 3.3.9). Perhaps the most striking result is that all proteases, glycosidases and proteins involved in the regulation of apoptosis are down-regulated at 31°C (Table 3.3.9). (Full lists of differentially-expressed proteins are provided in Appendix C, Table 3.3.10 and Table 3.3.11)

**Table 3.3.9** Number of differentially-expressed proteins between 31°C and 37°C at 24hrs post-temperature-shift in the culture supernatant of an IgG producing DP12 cell line. Using Uniprot protein database, proteins of biological relevance have been identified and the number of differentially-expressed proteins under the relevant classification included in the table.

	37°C vs 31°C 24hrs					
	Total		Up at 31°C		Down at 31°C	
	≥1 Peptide	≥2 Peptides	≥1 Peptide	≥2 Peptides	≥1 Peptide	≥2 Peptides
<b>Protein Count</b>	150	66	43	11	107	55
<b>Apoptosis</b>	3	3	0	0	3	3
	CLU, DNASE2, APP	CLU, DNASE2, APP			CLU, DNASE2, APP	CLU, DNASE2, APP
<b>Protease</b>	20	10	0	0	20	10
	DPP7, SCPEP1, TPP1, PRCP, PRSS22, HTRA1, CTSD, CTSH, C1ra, MMP19, MMP9, ADAMTS1, ADAM9, CTSZ, BMP1, CTSA, ADAM10, CTSL1, CTSB, P15156	DPP7, SCPEP1, TPP1, PRCP, PRSS22, HTRA1, CTSD, CTSH, C1ra, MMP19			DPP7, SCPEP1, TPP1, PRCP, PRSS22, HTRA1, CTSD, CTSH, C1ra, MMP19, MMP9, ADAMTS1, ADAM9, CTSZ, BMP1, CTSA, ADAM10, CTSL1, CTSB, P15156	DPP7, SCPEP1, TPP1, PRCP, PRSS22, HTRA1, CTSD, CTSH, C1ra, MMP19
<b>Glycosidase</b>	13	6	0	0	13	6
	GLB1, GAA, MAN2A1, FUCA1, MAN2B1, HEXA, SMPD1, MAN1A1, GLA, GUSB, NEU1, HEXB, NAGA	GLB1, GAA, MAN2A1, FUCA1, MAN2B1, HEXA			GLB1, GAA, MAN2A1, FUCA1, MAN2B1, HEXA, SMPD1, MAN1A1, GLA, GUSB, NEU1, HEXB, NAGA	GLB1, GAA, MAN2A1, FUCA1, MAN2B1, HEXA

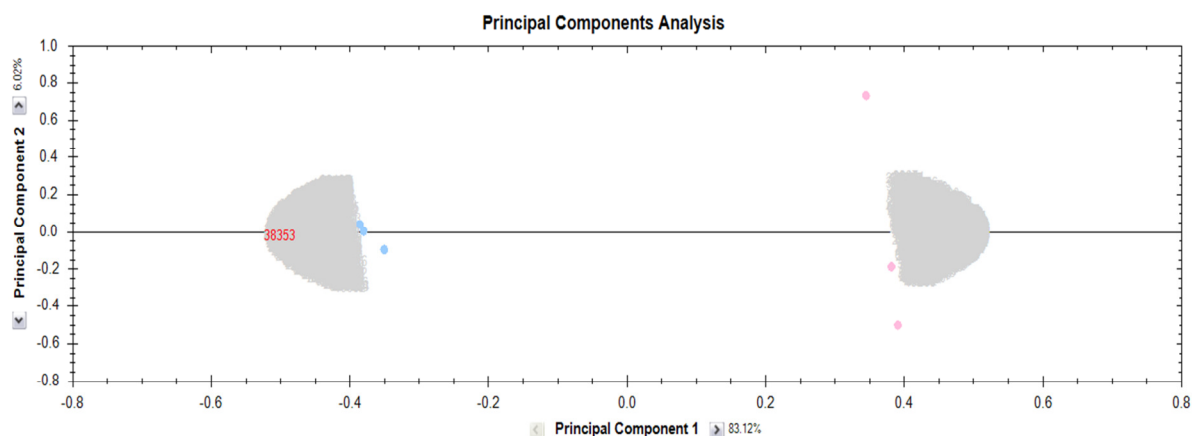
A number of interesting proteins were identified in this analysis including those with the potential to affect culture viability or negatively impact on product quality such as the growth factor Vascular endothelial growth factor C (2.4-fold) and the glycosidase Beta-galactosidase (2.5-fold). Both of which were down-regulated at 31°C compared 37°C to 24hrs (Figure 3.3.10).



**Figure 3.3.10** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the glycosidase Beta-galactosidase and the growth factor Vascular endothelial growth factor C, identified as being down-regulated at 31°C 24hrs post-temperature-shift in the conditioned media of a DP12 cell line.

### 3.3.4.4 Differential regulation of HCPs at 144hrs 37°C vs 31°C

A PCA plot generated by Progenesis during data analysis showing how the samples from 31°C and 37°C cultures at 144hrs cluster is displayed in Figure 3.3.11.



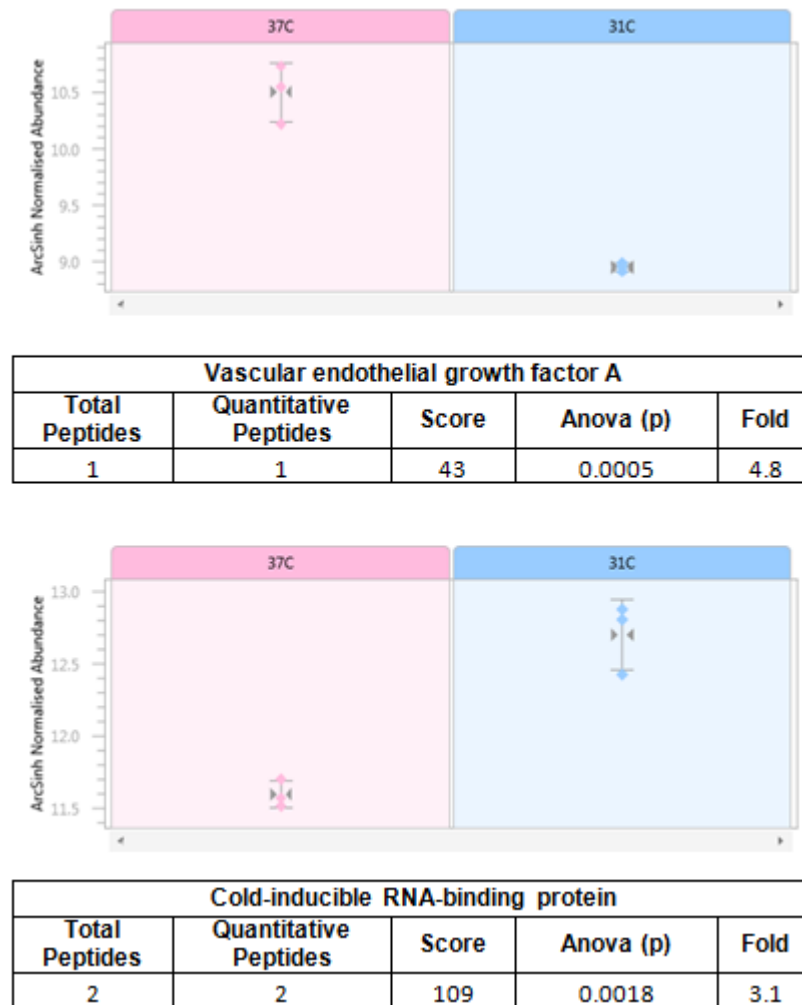
**Figure 3.3.11** PCA plot for the IgG secreting DP12 conditioned media sample groups 37°C (pink) and 31°C (blue) after 144hrs in culture generated during label-free LC/MS data analysis using Progenesis software

144hrs post-temperature-shift 222 proteins with 1 or more peptides are seen to alter their expression when temperature-shifted and non-temperature-shifted cultures are compared (Table 3.3.12) 127 proteins were up-regulated at 31°C, while 95 decreased in expression (See Appendix C, Table 3.13 and Table 3.14 for full list). As seen in Table 3.3.12, three of the four apoptotic proteins identified are down-regulated at 31°C compared to 37°C at 144hrs. Similar to the CHO-K1 cell line (section 3.2.3.5), this could be attributed to the overall increased viability of the temperature-shifted culture compared to the non-temperature-shifted culture at 144hrs. It is also worth noting that the number of differentially-expressed proteases and glycosidases are relatively evenly spread between the two culture conditions. This would indicate that enzymatic activity at late stage culture under temperature-shift or non-temperature-shift conditions could negatively impact on product quality.

**Table 3.3.12** Number of differentially-expressed proteins between 31°C and 37°C at 144hrs post-temperature-shift in the culture supernatant of an IgG producing DP12 cell line. Using Uniprot protein database, proteins of biological relevance have been identified and the number of differentially-expressed proteins under the relevant classification included in the table.

	37°C vs 31°C 144hrs					
	Total		Up at 31°C		Down at 31°C	
	≥1 Peptide	≥2 Peptides	≥1 Peptide	≥2 Peptides	≥1 Peptide	≥2 Peptides
<b>Protein Count</b>	222	77	127	46	95	31
<b>Apoptosis</b>	4	2	1	1	3	1
	CLU, HMOX1, PPID, MAPK1	CLU, HMOX1	HMOX1	HMOX1	CLU, PPID, MAPK1	CLU,
<b>Protease</b>	9	5	4	2	5	3
	BMP1, HTRA1, DPP7, P15156, THOP1, MMP19, ADAM17, PSMA5, MMP9	BMP1, HTRA1, DPP7, P15156, THOP1	DPP7, P15156, PSMA, MMP9	DPP7, P15156	BMP1, HTRA1, THOP1, MMP19, ADAM17	BMP1, HTRA1, THOP1
<b>Glycosidase</b>	9	6	5	2	4	4
	GAA, HEXA, NAGLU, FUCA1, MAN2A1, GANAB, GUSB, GLB1, HEXB	GAA, HEXA, NAGLU, FUCA1, MAN2A1, GANAB	NAGLU, FUCA1, GUSB, GLB1, HEXB	NAGLU, FUCA1	GAA, HEXA, MAN2A1, GANAB	GAA, HEXA, MAN2A1, GANAB





Among the differentially-expressed proteins identified at 31°C compared to 37°C at 144hrs was the secreted growth factor Vascular endothelial growth factor A (4.8-fold, down-regulated) and the protein Cold-inducible RNA-binding protein (CIRBP) (3.1-fold, up-regulated) (Figure 3.3.12).



**Figure 3.3.12** Shows examples of Progenesis label-free outputs of differentially-expressed HCPs, such as the secreted growth factors Bone morphogenetic protein 1 and Vascular endothelial growth factor A, identified as being down-regulated at 31°C 144hrs post-temperature-shift in the conditioned media of a DP12 cell line.

### 3.3.4.5 Western blot validation of HCPs identified as differentially-expressed following quantitative label-free LC/MS analysis of the cell culture supernatant from a CHO-K1 and DP12 cell line

Western blot analysis was used to validate the results obtained following quantitative label-free LC/MS analysis of the cell culture supernatant from CHO-K1 and DP12 cell lines. Validation was conducted in biological duplicate using 8µg of protein per sample. Western blots show the down-regulation of sulfated glycoprotein 1 at 31°C when 31°C culture is compared to 37°C culture at 24hrs and 144hrs in the DP12 cell line (Figure 3.3.13). The intracellular protein GRP78 is shown to be up-regulated at 31°C when non-temperature-shifted culture is compared to temperature-shifted culture at 24hrs and at 144hrs in the CHO-K1 cell line in Figure 3.3.14. GRP78 is also shown to be up-regulated at 144hrs when 31°C culture is compared to 37°C culture at 24hrs and when 31°C culture is compared to 37°C culture at 144hrs in the DP12 cell line (Figure 3.3.15). Validation of the differential-regulation of the proteases BMP1 (up-regulated at 120hrs in the CHO-K1 cell line when 120hrs is compared to 24hrs at 37°C) and MMP19 (down-regulated at 31°C in the CHO-K1 cell line when the 31°C culture is compared to the 37°C culture at 120hrs) are shown in Figure 3.3.16 and Figure 3.3.17 respectively. It should be noted that the text cited in Figure 3.3.13 to Figure 3.3.17 are based on the results from the label-free LC/MS analysis of the conditioned media from the respective CHO cell lines and the western blots were used to validate the direction of fold change.

Sulfated Glycoprotein 1					
DP12 cell line					
Comparison	Anova (p)	Fold	Direction of fold change	37°C	31°C
37°C vs 31°C 24hrs	0.0241	6.2	Down at 31°C		
37°C vs 31°C 144hrs	0.0231	5.0	Down at 31°C		

**Figure 3.3.13** shows a western blot confirming the down-regulation of the intracellular protein Sulfated Glycoprotein 1 at 31°C when temperature-shifted culture is compared to non-temperature-shifted culture at 24hrs and 144hrs in the conditioned media of the DP12 cell line.

GRP78					
CHO-K1 cell line					
Comparison	Anova (p)	Fold	Direction of fold change	37°C	31°C
37°C vs 31°C 24hrs	0.0022	3.3	Up at 31°C		
37°C vs 31°C 120hrs	0.0191	2.9	Up at 31°C		

**Figure 3.3.14** shows a western blot confirming the up-regulation of the intracellular protein GRP78 at 31°C when temperature-shifted culture is compared to non-temperature-shifted culture at 24hrs and 120hrs in the conditioned media of the CHO-K1 cell line.



GRP78					
DP12 cell line					
Comparison	Anova (p)	Fold	Direction of fold change	24hrs	144hrs
24hrs vs 144hrs 37°C	0.0003	39.8	Up at 144hrs		
24hrs vs 144hrs 31°C	0.00004	19.7	Up at 144hrs		

**Figure 3.3.15** shows a western blot confirming the up-regulation of the intracellular protein GRP78 at 144hrs when 144hrs is compared to 24hrs at 37°C and at 31°C in the conditioned media of the DP12 cell line.

BMP1					
CHO-K1 cell line					
Comparison	Anova (p)	Fold	Direction of fold change	24hrs	120hrs
24hrs vs 120hrs 37°C	0.0312	12.5	Up at 120hrs		

**Figure 3.3.16** shows a western blot confirming the up-regulation of the protease BMP1 at 120hrs when 120hrs is compared to 24hrs at 37°C in the conditioned media of the CHO-K1 cell line.



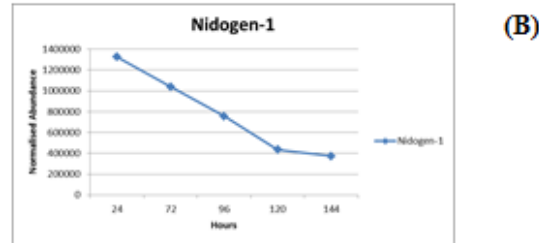
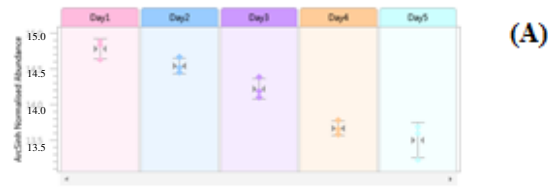
MMP19					
CHO-K1 cell line					
Comparison	Anova (p)	Fold	Direction of fold change	37°C	31°C
37°C vs 31°C 120hrs	0.0044	4.0	Down at 31°C		

**Figure 3.3.17** shows a western blot confirming the down-regulation of the protease MMP19 at 31°C when temperature-shifted culture is compared to non-temperature-shifted culture in the conditioned media of the CHO-K1 cell line.

### 3.3.4.6 Pearson correlation analysis of differentially-regulated HCPs over 24, 72, 96, 120 and 144 hours at 37°C

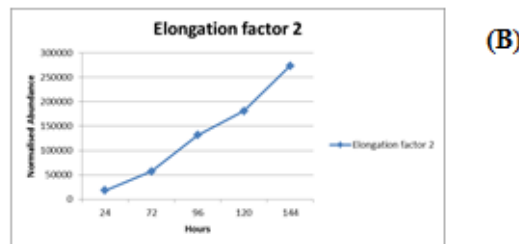
Pearson's correlation analysis was used as a means of measuring the linearity of protein expression over time in culture in the 37°C culture. An Anova cut-off of 0.05 was applied to add an additional layer of statistical rigour. Using this method of analysis a total of 227 proteins were deemed to change over time in culture. 167 correlated positively (increased expression) and 60 correlated negatively (decreased in expression) between the 24 and 144hr period. The list of negatively correlating proteins is shown in Appendix C, Table 3.3.15 and those proteins displaying an increase in expression are shown in Appendix C, Table 3.3.16.

The linear decrease in the expression of the secreted protein Nidogen-1 (Pearson correlation 0.984, p-value 0.0024) between 24hrs and 144hrs at 37°C is shown in Figure 3.3.18. The Progenesis and Pearson correlation outputs from the protein Elongation factor 2 (Pearson correlation 0.991, p-value 0.001) showing its increase in expression from 24hrs to 144hrs at 37°C are shown in Figure 3.3.19.



Nidogen-1								
Pearson's correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest	Lowest
-0.984	0.0024	5	3	270	5.75E-06	3.6	24hrs	144hrs

**Figure 3.3.18** Progenesis label-free output of the secreted protein Nidogen-1 showing its decrease in expression from 24hrs (Day1) to 144hrs (Day5) at 37°C (A). Output from excel on which Pearson correlation calculation is based (Pearson correlation 0.984, p-value 0.0024). (B)



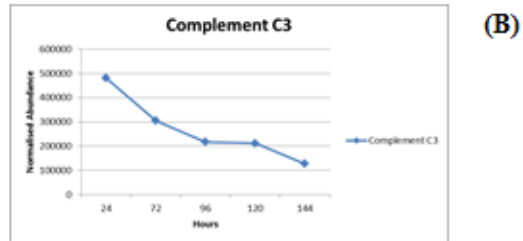
Elongation factor 2								
Pearson's correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest	Lowest
0.991	0.0010	6	6	308	1.32E-06	14.9	144hrs	24hrs

**Figure 3.3.19** Progenesis label-free output of the protein Elongation factor 2 showing its increase in expression from 24hrs (Day1) to 144hrs (Day5) at 37°C (A). Output from excel on which Pearson correlation calculation is based (Pearson correlation 0.991, p-value 0.001). (B)

#### **3.3.4.7 Pearson correlation analysis of differentially-regulated HCPs over 24, 72, 96, 120 and 144 hours at 31°C**

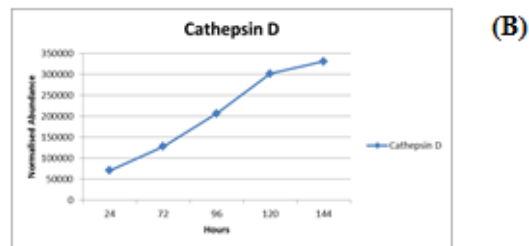
This analysis identified that 405 proteins increase expression in a linear manner over time in culture with a p-value  $\leq 0.05$  (Appendix C, Table 3.3.17), but only 2 proteins that decrease in expression meet the same criteria (Appendix C, Table 3.3.18).

Figure 3.3.20 shows the decrease in the expression of the secreted protein Complement C3 (Pearson correlation 0.892, p-value 0.042) between 24hrs and 144hrs post-temperature-shift at 31°C while Figure 3.3.21 shows the linear increase in the expression of the protease Cathepsin D (Pearson correlation 0.985, p-value 0.0022) over the same period at 31°C.



Complement C3								
Pearson's correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest	Lowest
-0.892	0.042018	11	11	619	3.95E-04	3.8	24hrs	144hrs

**Figure 3.3.20** Progenesis label-free output of the secreted protein Complement C3 showing its decrease in expression from 24hrs (Day1) to 144hrs (Day5) at 31°C (A). Output from excel on which Pearson correlation calculation is based (Pearson correlation 0.892, p-value 0.042). (B)



Cathepsin D								
Pearson's correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest	Lowest
0.985	0.0022	4	4	202	7.40E-08	5.7	144hrs	24hrs

**Figure 3.3.21** Progenesis label-free output of the protease Cathepsin D showing its increase in expression from 24hrs (Day1) to 144hrs (Day5) at 31°C (A). Output from excel on which Pearson correlation calculation is based (Pearson correlation 0.985, p-value 0.0022). (B)

#### **3.3.4.8 Comparison of differentially-regulated HCPs over time in culture at 37°C and 31°C in a DP12 cell line**

By analysing the conditioned media of a DP12 cell line over time in culture under temperature-shifted and non-temperature-shifted conditions (sections 3.3.4.1 and 3.3.4.2), and by comparing both conditions at early (24hrs) and late (144hrs) stages (sections 3.3.4.3 and 3.3.4.4), a significant number of differentially-expressed HCPs have been identified. By combining the results from 24hrs vs 120hrs at 37°C and 24hrs vs 120hrs at 31°C experiments a clearer picture of how culture temperature impacts on the HCP expression profile can be seen (Appendix C, Table 3.3.19).

Proteins identified as differentially-expressed over time in culture common to both 37°C and 31°C that may be of particular interest to the bioprocessing community are shown in Figure 3.3.22. As well as providing differential regulation information, Figure 3.3.22 shows graphs generated by Progenesis label-free software and the changes in the relative abundance of the proteases Cathepsin B and Complement C1r-A subcomponent and the glycosidase Glucosylceramidase. The enzymatic activity of proteases and glycosidases could negatively impact on the quality of products secreted into the media, information on how they change in relative abundance between early (24hrs) and late (144hrs) stages in culture under temperature-shifted and non-temperature-shifted conditions may be useful for identifying culture conditions under which a recombinant therapeutic has an increased risk of becoming degraded through the enzymatic activity of proteases and/or glycosidases.

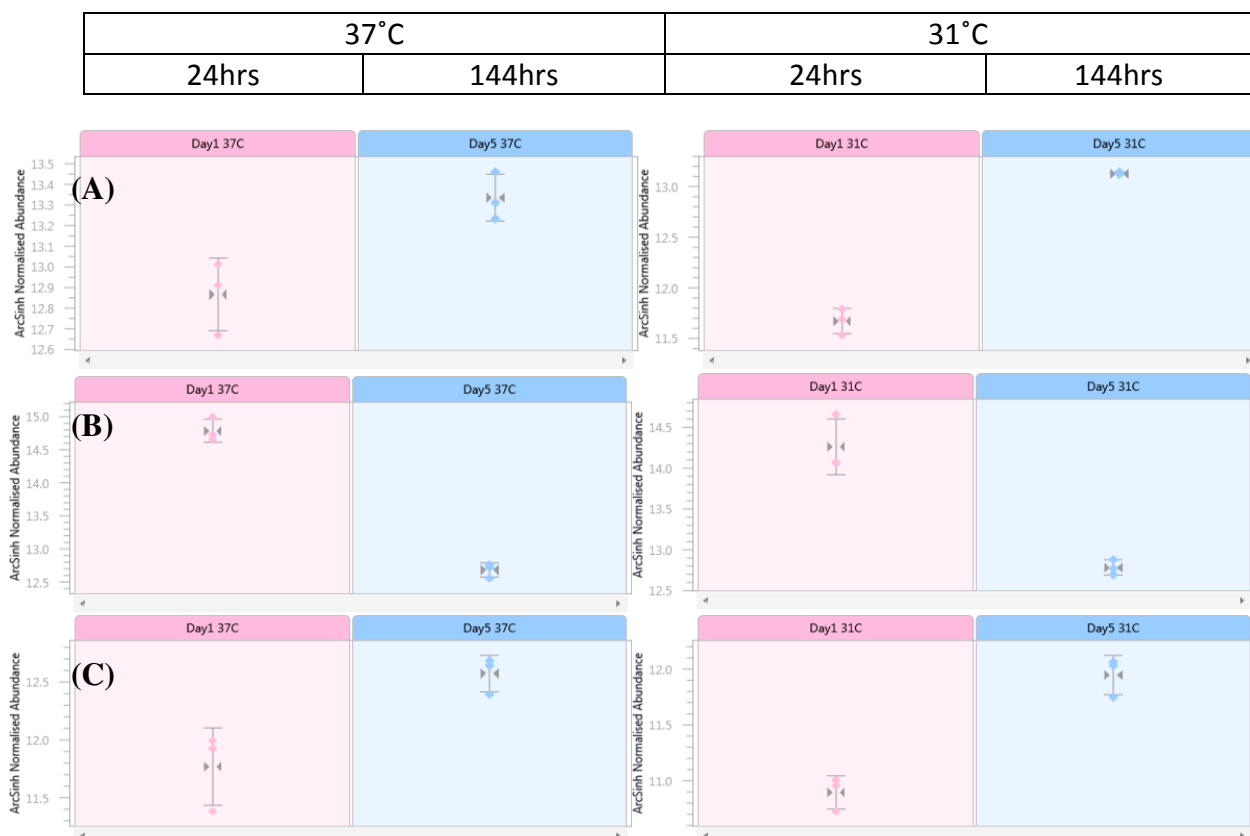


Fig.	Description	Total Peptide	Quantitative Peptides	Score	Anova (p)	Fold	Direction of fold change	Comparison
A	Cathepsin B	3	2	214	0.018	1.6	Up at 144hrs	24hrs vs 144hrs 37C
		2	2	117	3.59E-05	4.3	Up at 144hrs	24hrs vs 144hrs 31C
B	Complement C1r-A subcomponent	2	2	119	0.0001	8.3	Down at 144hrs	24hrs vs 144hrs 37C
		2	2	134	0.0019	4.6	Down at 144hrs	24hrs vs 144hrs 31C
C	Glucosylceramidase	2	2	91	0.0196	2.2	Up at 144hrs	24hrs vs 144hrs 37C
		2	2	103	0.0014	2.9	Up at 144hrs	24hrs vs 144hrs 31C

**Figure 3.3.22** Shows changes in the expression of the proteases (A) Cathepsin B and (B) Complement C1r-A subcomponent and the glycosidase (C) Glucosylceramidase between early (24hrs) and late (144hrs) stages in culture under temperature-shifted and non-temperature-shifted conditions.

As outlined in section 3.1.3.7 many of the differentially-expressed HCPs trend in the same direction under both temperature-shifted and non-temperature-shifted conditions. In addition, proteins identified as differentially-expressed that were common to 24hrs vs 144hrs 37°C and 24hrs vs 144hrs 31°C lists, were also compared to the list of differentially-expressed proteins from 37°C vs 31°C at 144hrs. A list of proteins common to all three comparisons was then compiled (Appendix C, Table 3.3.20). Thus information on how the abundance of the protein changes over time and the culture temperature at which the relative abundance of the protein is greatest is provided.

Based on this comparison, a clearer picture of how the expression of HCPs change between early (24hrs) and late (144hrs) stages in culture under temperature-shifted and non-temperature-shifted conditions as well as how they change in relative abundance between 37°C and 31°C in late culture is provided. Using the outputs from Progenesis label-free software examples of the differential expression of three secreted proteins is shown in Figure 3.3.23. Information on secreted proteases such as matrix metalloproteinase-9 (Figure 3.3.23) is of obvious benefit as such proteases may impact on product quality.

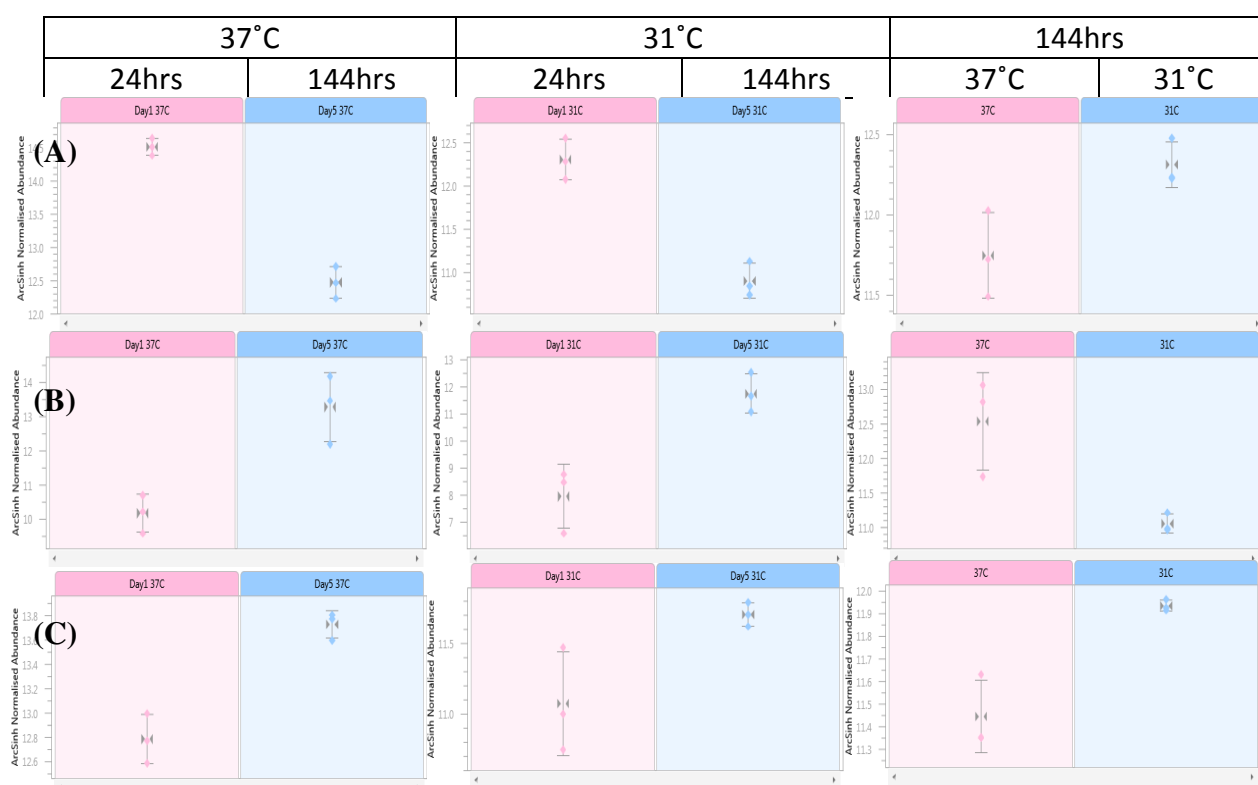


Fig.	Description	Total Peptide	Quantitative Peptides	Score	Anova (p)	Fold	Direction of fold change	Comparison
A	matrix metalloproteinase-9	4	4	269	0.0002	7.6	Down at 144hrs	24hrs vs 144hrs 37C
		2	2	89	0.0015	4.1	Down at 144hrs	24hrs vs 144hrs 31C
		1	1	70	0.0315	1.7	Up at 31C	37C vs 31C 144hrs
B	Sulfated glycoprotein 1	2	2	91	0.0094	27.1	Up at 144hrs	24hrs vs 144hrs 37C
		3	3	147	0.009	38	Up at 144hrs	24hrs vs 144hrs 31C
		2	2	126	0.0231	5	Down at 31C	37C vs 31C 144hrs
C	Suprabasin	6	6	466	0.0022	2.5	Up at 144hrs	24hrs vs 144hrs 37C
		4	4	242	0.0443	1.8	Up at 144hrs	24hrs vs 144hrs 31C
		2	2	135	0.0064	1.6	Up at 31C	37C vs 31C 144hrs

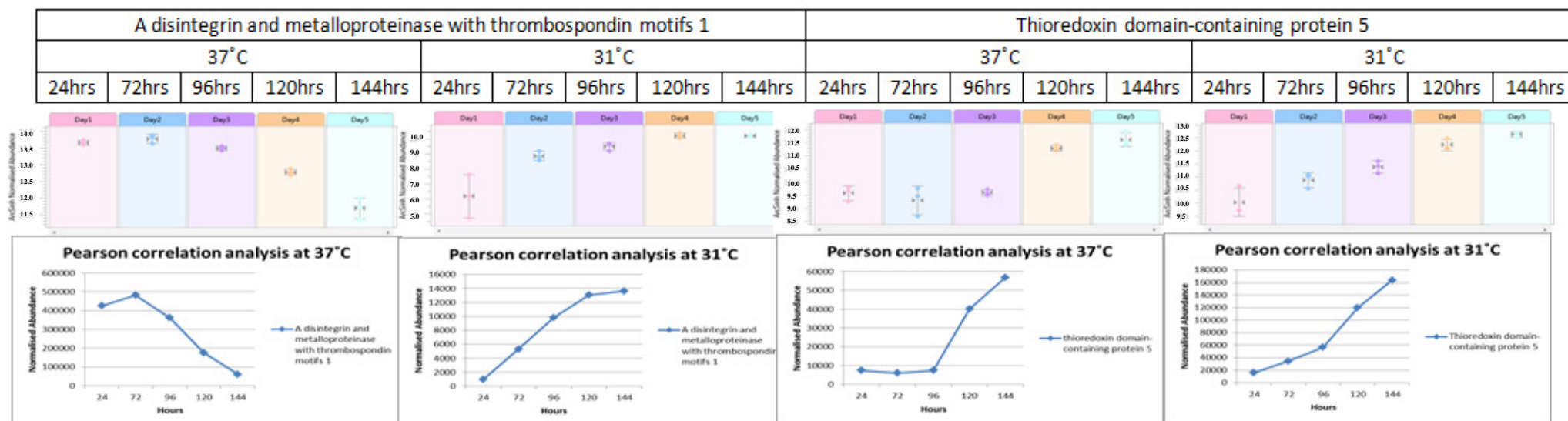
**Figure 3.3.23** Progenesis label-free outputs and information on the differential expression of the secreted proteins (A) Matrix metalloproteinase-9, (B) Sulfated glycoprotein 1 and (C) Suprabasin as they change in relative abundance between early (24hrs) and late (144hrs) stages in culture. Information from a comparison between 37°C and 31°C at 144hrs is also provided.



### **3.3.4.9 Comparison of Pearson correlation analyses of differentially-regulated HCPs over time in culture at 37°C and 31°C in a DP12 cell line**

Pearson correlations were used to examine the linearity of the expression of proteins overtime in temperature-shifted and non-temperature-shifted cultures in sections 3.2.4.5 and 3.2.4.6. Similar to section 3.1.3.7, proteins common to both 37°C and 31°C Pearson correlation lists are displayed in Appendix C, Table 3.3.21. By comparing the data in this manner the effect of culture temperature on the expression profiles of HCPs over time in 37°C and 31°C cultures can clearly be seen.

Pearson correlation plots and the outputs from Progenesis label-free software are used to show the change in abundance of some proteins from Table 3.3.21 that are of particular interest due to their potential ability to impact on cell viability or product quality. These examples include the protease ADMTS1 and the reductase TXNDC5 (Figure 3.3.24). Interestingly, the relative abundance of ADMST1 reduces over time in culture at 37°C and it increases over the same period of time at 31°C. While the linearity with which TXNDC5 increases at 37°C is poor, it can be clearly seen that between 96hrs and 144hrs there is a sharp increase in its abundance. This is compared with the 31°C where it increases in a more linear fashion over the total period in culture.

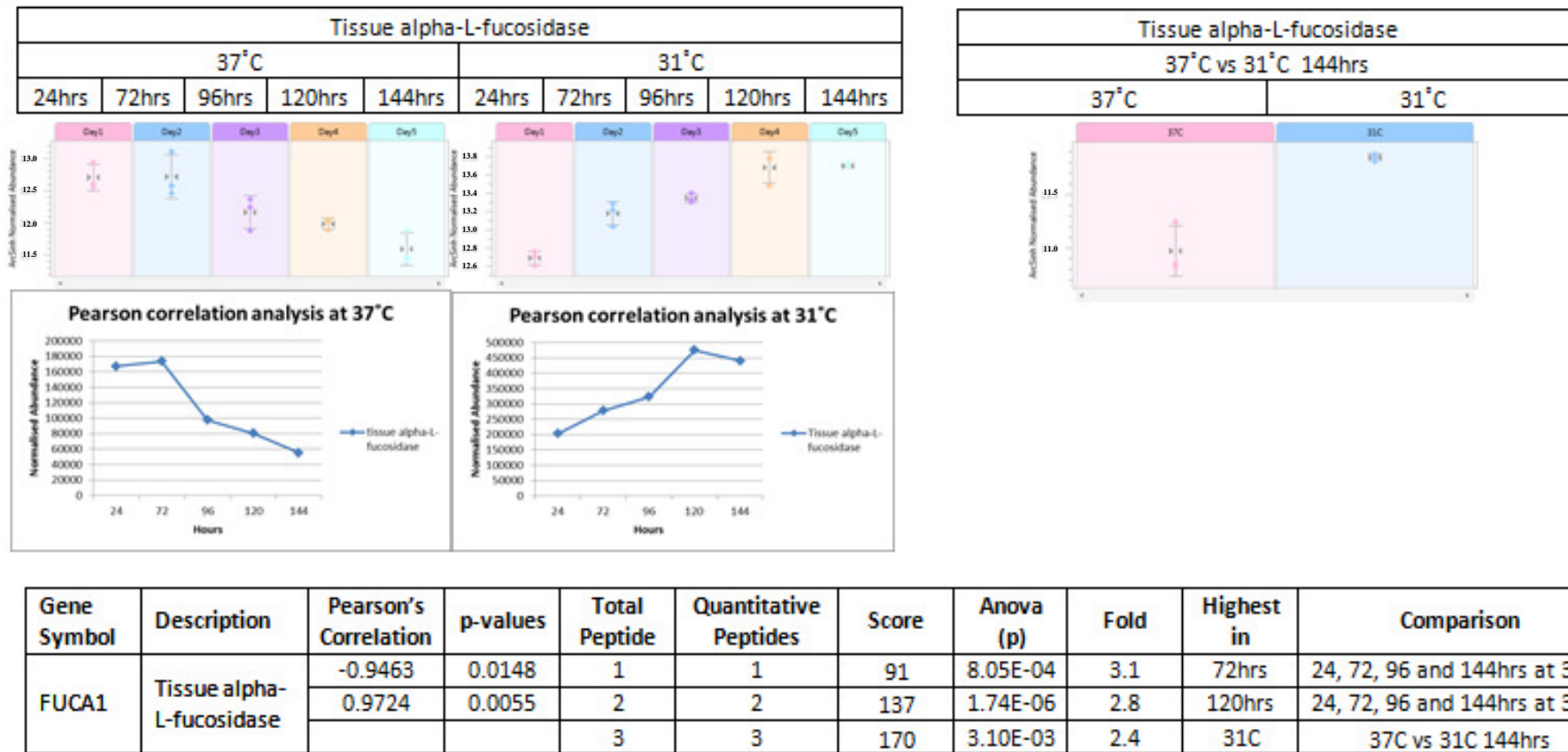


Gene Symbol	Description	Pearson's Correlation	p-values	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold	Highest at	Comparison
ADAMTS1	A disintegrin and metalloproteinase with thrombospondin motifs 1	-0.9249	0.0244	4	4	266	1.04E-07	7.9	72hrs	24hrs vs 144hrs 37C
		0.9668	0.0072	1	1	78	1.25E-04	31.8	144hrs	24hrs vs 144hrs 31C
TXNDC5	Thioredoxin domain-containing protein 5	0.8942	0.0407	2	2	96	5.63E-06	9.4	144hrs	24hrs vs 144hrs 37C
		0.9630	0.0085	2	2	106	9.46E-06	12.2	144hrs	24hrs vs 144hrs 31C

**Figure 3.3.24** Pearson correlation plots and the outputs from Progenesis label-free software of the protease ADAMTS1 (A) and the reductase TXNDC5 (B) show how they change over time in culture under both temperature-shifted and non-temperature-shifted conditions.

The proteins list obtained from the comparison of Pearson correlations analysis of the expression of proteins over time at 37°C and at 31°C was also compared to the list of differentially-expressed HCPs identified when 37°C and 31°C cultures were compared at 144hrs (section 3.2.4.4). Because many of the proteins identified using the Pearson correlation approach trend in the same direction, the use of this third list helps identify which culture has the highest abundance of a given protein. A compilation of this data obtained from the comparison of all three lists is shown in Appendix C, Table 3.3.22.

Pearson correlation plots and the outputs from Progenesis label-free software are used to show the changes in the abundance of the secreted protein the fucosidase Tissue alpha-L-fucosidase (Figure 3.3.25). While the precise function of these secreted proteins in CHO is still unclear it is obvious that temperature affects the release of Tissue alpha-L-fucosidase.



**Figure 3.3.25** Pearson correlation plots and the outputs from Progenesis label-free software of the secreted protein Tissue alpha-L-fucosidase shows how it changes over time in culture under both temperature-shifted and non-temperature-shifted conditions. Information from a comparison between 37°C and 31°C at 144hrs is also provided.

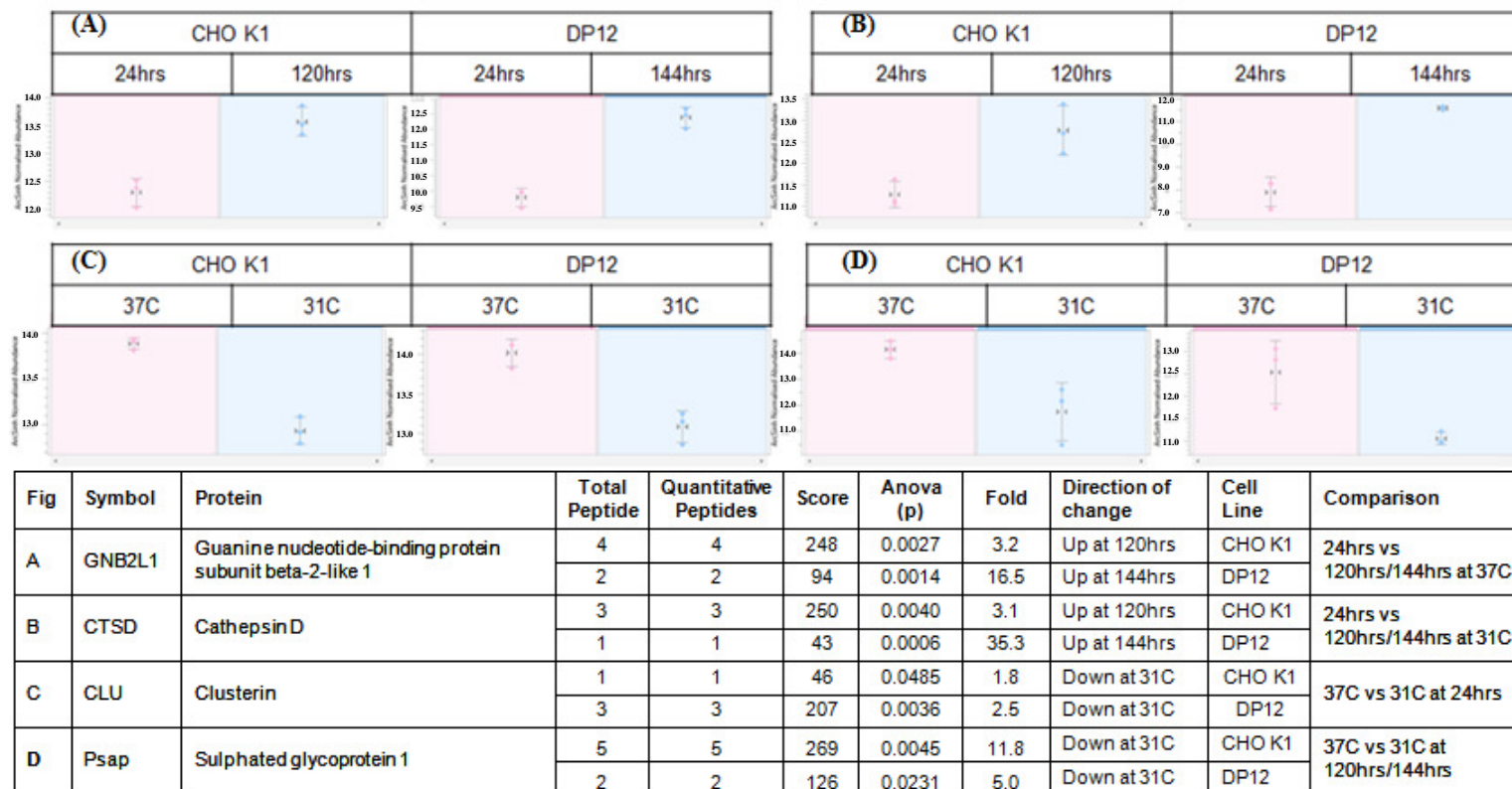
### 3.3.4.10 Comparison of differentially-regulated HCPs in the cell culture supernatant of a CHO-K1 and DP12 cell line.

A significant number of differentially-expressed HCPs were identified both over time in culture and under temperature-shifted and non-temperature-shifted conditions in the conditioned media of CHO-K1 (Section 3.2.3.1 – Section 3.2.3.4) and DP12 cell lines (Section 3.3.4.1 – Section 3.3.4.4). It was therefore decided that the list of differentially-expressed HCPs from each cell line should be compared in an effort to identify proteins of interest that are common to both cell lines. A summary of the number of differentially-expressed HCPs that are common to both cell lines compared to the total number of differentially-expressed HCPs identified in each cell line is shown in Table 3.3.23. While it can be seen that there is some overlap in the differential expression of HCPs between CHO-K1 and Dp12 cells, there is a significant proportion of differentially-expressed HCPs that are unique to each cell line.

Comparison	Number of differentially expressed HCPs in the CHO K1 cell line	Number of differentially expressed HCPs in the DP12 cell line	Number of differentially expressed HCPs common to CHO K1 and DP12 cell lines
24hrs vs 120hrs / 144hrs at 37C	322	375	174
24hrs vs 120hrs / 144hrs at 31C	146	391	50
37C vs 31C at 24hrs	211	150	20
37C vs 31C at 120hrs / 144hrs	228	222	55

**Table 3.3.23** Table indicating the number of differentially-expressed HCPs identified in a CHO-K1 and DP12 cell line compared to the number of differentially-expressed HCPs common to both cell lines. A full list of proteins common to each cell line is available in Appendix C, Table 3.3.24 - 3.3.27.

A number of examples were selected from the list of common differentially-regulated proteins and their change in expression visualised using the output from Progenesis label-free software as displayed in Figure 3.3.26. This includes the protease Cathepsin D, the secreted protein Sulphated glycoprotein 1 and the apoptotic protein GNB2L1.



**Figure 3.3.26** Progenesis label-free outputs and information on HCPs identified as being differentially-expressed in both CHO-K1 and DP12 cell lines. This includes the apoptotic protein GNB2L1, the protease Cathepsin D, the protein Clusterin which can be found both intracellularly and extracellularly and the secreted protein Sulphated glycoprotein 1.

### 3.3.4.11 Comparison of quantitative label-free LC/MS DP12 HCP data normalised to cell number compared to normalisation by protein concentration

Correction factors were applied to quantitative label-free LC/MS data on the differential expression of HCPs in the conditioned media of DP12 cells. This was carried out to account for changes in cell concentration (both viable and non-viable) over time or under temperature-shifted and non-temperature-shifted conditions (as described in section). As can be seen in Table 3.3.28, applying the correction factor resulted in the loss of a small number of proteins from three of the comparisons (24hrs vs 144hrs at 37°C, 24hrs vs 144hrs at 31°C and 37°C vs 31°C at 24hrs). In the remaining comparison however (37°C vs 31°C at 144hrs), the use of the correction factor resulted in the loss more than half the differentially-expressed proteins originally identified. Interestingly, when the same correction factor was applied to the 24hrs vs 144hrs at 31°C comparison only 26 proteins were deemed to no longer be differentially-expressed. This highlights the high fold-changes (>2.09 fold) of proteins over time in culture at 31°C and the significant number of ‘subtle’ fold-changes (<2.09 fold) between temperature-shifted and non-temperature-shifted conditions at 144hrs in the DP12 cells. Future analysis of this data could consider the use of the correction factor as a means of adding further confidence to the list of differentially-expressed proteins.

Comparison	Normalisation correction factor	Normalisation correction factor + 1.5-fold change cut-off	Number of differentially expressed proteins identified using correction factor	Number of differentially expressed proteins identified using 1.5-fold cut-off
24hrs vs 144hrs at 37°C	0.16	1.66	369	375
24hrs vs 144hrs at 31°C	0.59	2.09	365	391
37°C vs 31°C at 24hrs	0.17	1.67	138	150
37°C vs 31°C at 144hrs	0.59	2.09	105	222

**Table 3.3.28** The use of a correction factor was investigated to facilitate normalisation of the data by cell number (as opposed to normalisation by protein concentration). This table show a comparison of the number of differentially-expressed HCPs identified in the conditioned media of a DP12 cell line with and without the correction factor applied.

### 3.3.5 Bioinformatic analysis of differentially-regulated proteins

In order to gain an insight into the precise origin of the host cell proteins identified i.e. intracellular, secreted, etc. the gene symbol from each protein was searched using DAVID functional annotation software (<http://david.abcc.ncifcrf.gov/>).

As detailed in section 3.2.4, DAVID was used to classify the differentially-regulated proteins and identify enrichment of cellular locations. Protein lists obtained from comparing 24hrs to 144hrs at both 37°C and 31°C were searched against DAVID using an adjusted p-value of  $\leq 0.05$  (Benjamini) as a statistical cut-off. Similar to section 3.3.4, DAVID GO analysis of differentially-expressed proteins from the cell culture supernatant of an IgG producing DP12 cell line examining 24hrs versus 144hrs at 37°C found that the majority of the proteins identified were typically intracellular in nature and although there was enrichment of the extracellular component (adjusted p-value 0.0011), this appears to make up only a small proportion of the overall host cell protein population (Table 3.3.30 and Figure 3.3.27). Interestingly, when 24hrs versus 144hrs at 31°C was analysed using DAVID GO analysis it was found that the extracellular component was also predominantly made up by proteins of intracellular origin (Table 3.3.31 and Figure 3.3.28). This would appear to indicate that many of the host cell proteins present in the cell culture supernatant are there as a result of some sort of cell lysis, possibly as a result of cell death, particularly near the end of culture. Despite this, based on gene count alone, there is still a cohort of the host cell protein population made up by extracellular proteins, this is also evidenced by the number of proteins classified as secreted using Uniprot protein database (Table 3.3.29). It is also possible that the reason for such a large intracellular enrichment is that proteins can occupy more than one cell component, this may also give rise to the appearance of an increased proportion of intracellular proteins.

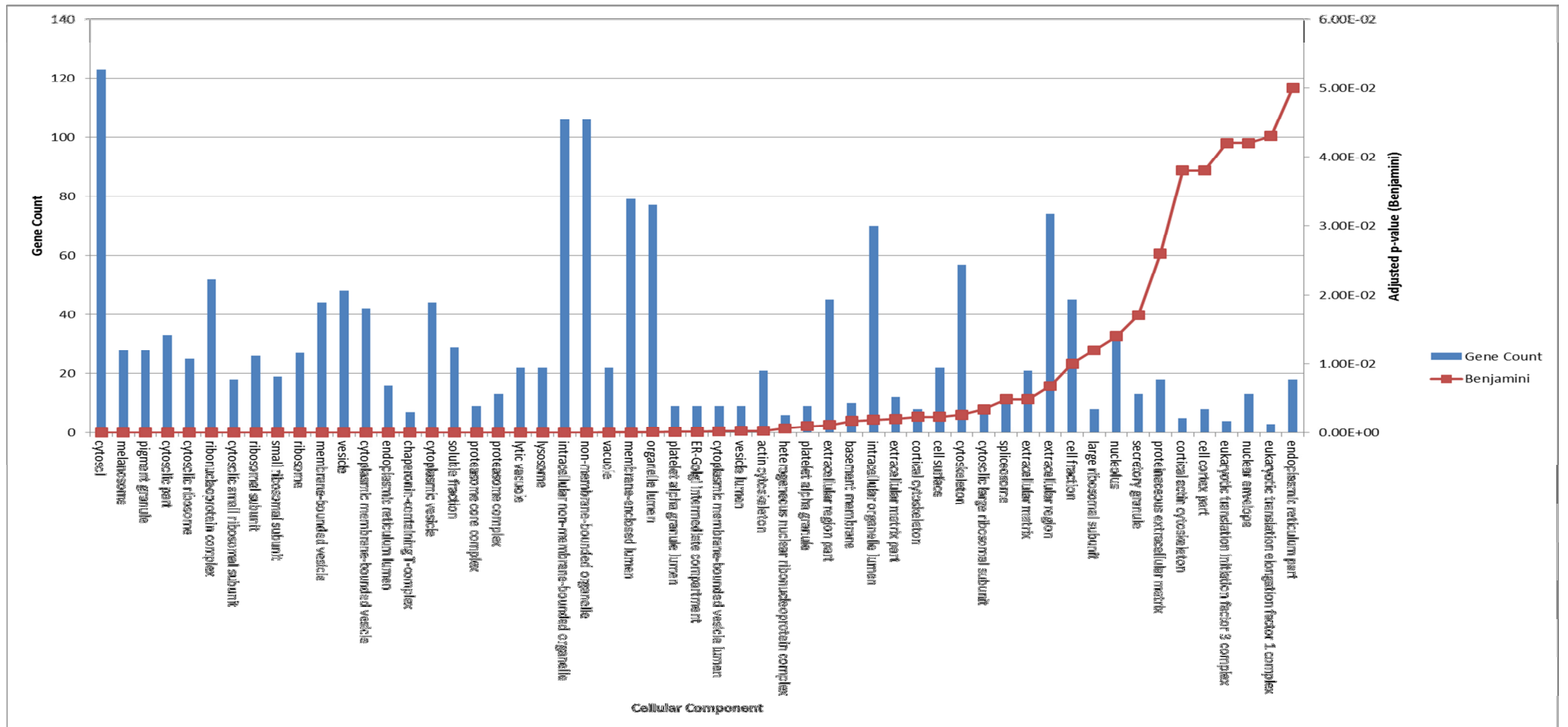
Conditions	Appendix Table
Up at 144hrs when compared to 24hrs at 37°C	Table 3.3.32
Down at 144hrs when compared to 24hrs at 37°C	Table 3.3.33
Up at 144hrs when compared to 24hrs at 31°C	Table 3.3.34
Down at 144hrs when compared to 24hrs at 31°C	Table 3.3.35

**Table 3.3.29** List of tables containing quantitative information on proteins classified as secreted by the Uniprot protein database.



**Table 3.3.30** GO cellular component enrichment for differentially-expressed host cell proteins identified in the conditioned media of a DP12 CHO cell line at 37°C when 24hrs was compared to 144hrs (post-temperature-shift for a culture grown in parallel) analysed using quantitative label-free LC/MS. An adjusted p-value of  $\leq 0.05$  (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.

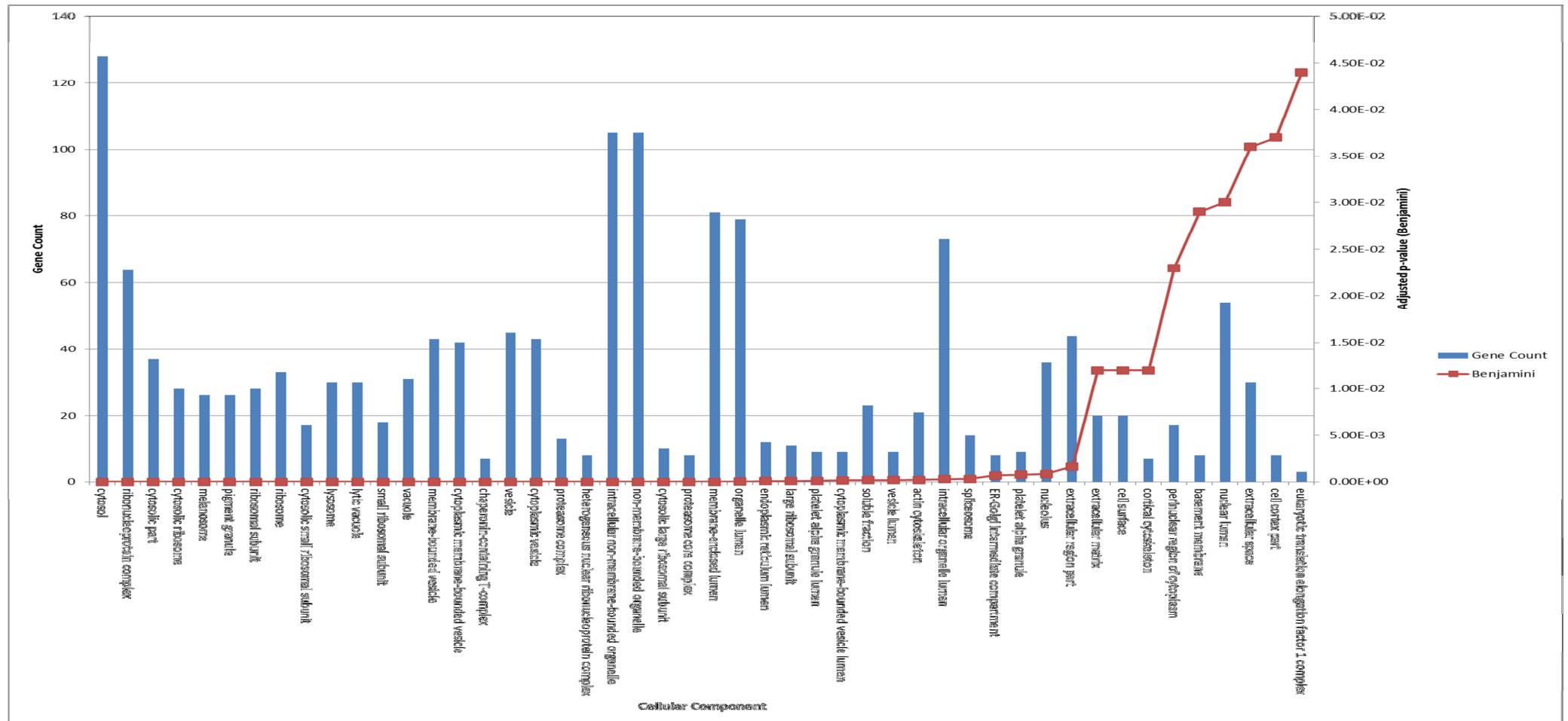
Term	Gene Count	P-Value	Benjamini
cytosol	123	1.00E-39	3.60E-37
melanosome	28	3.30E-22	5.70E-20
pigment granule	28	3.30E-22	5.70E-20
cytosolic part	33	9.00E-21	1.00E-18
cytosolic ribosome	25	1.30E-19	1.20E-17
ribonucleoprotein complex	52	2.80E-17	1.60E-15
cytosolic small ribosomal subunit	18	2.60E-17	1.80E-15
ribosomal subunit	26	1.30E-15	6.10E-14
small ribosomal subunit	19	1.10E-14	4.90E-13
ribosome	27	3.80E-11	1.40E-09
membrane-bounded vesicle	44	1.20E-10	4.10E-09
vesicle	48	2.00E-10	6.20E-09
cytoplasmic membrane-bounded vesicle	42	5.30E-10	1.50E-08
endoplasmic reticulum lumen	16	1.30E-09	3.50E-08
chaperonin-containing T-complex	7	1.80E-09	4.40E-08
cytoplasmic vesicle	44	5.20E-09	1.20E-07
soluble fraction	29	7.20E-09	1.60E-07
proteasome core complex	9	1.50E-08	3.20E-07
proteasome complex	13	3.40E-08	6.60E-07
lytic vacuole	22	9.60E-08	1.80E-06
lysosome	22	9.60E-08	1.80E-06
intracellular non-membrane-bounded organelle	106	1.80E-07	3.10E-06
non-membrane-bounded organelle	106	1.80E-07	3.10E-06
vacuole	22	1.90E-06	3.10E-05
membrane-enclosed lumen	79	3.10E-06	4.80E-05
organelle lumen	77	5.40E-06	8.20E-05
platelet alpha granule lumen	9	7.40E-06	1.10E-04
ER-Golgi intermediate compartment	9	7.40E-06	1.10E-04
cytoplasmic membrane-bounded vesicle lumen	9	1.30E-05	1.80E-04
vesicle lumen	9	1.80E-05	2.30E-04
actin cytoskeleton	21	1.80E-05	2.40E-04
heterogeneous nuclear ribonucleoprotein complex	6	5.00E-05	6.10E-04
platelet alpha granule	9	7.90E-05	9.40E-04
extracellular region part	45	9.50E-05	1.10E-03
basement membrane	10	1.60E-04	1.70E-03
intracellular organelle lumen	70	1.70E-04	1.90E-03
extracellular matrix part	12	1.90E-04	2.00E-03
cortical cytoskeleton	8	2.20E-04	2.30E-03
cell surface	22	2.30E-04	2.30E-03
cytoskeleton	57	2.70E-04	2.60E-03
cytosolic large ribosomal subunit	7	3.60E-04	3.40E-03
spliceosome	12	5.40E-04	4.80E-03
extracellular matrix	21	5.30E-04	4.80E-03
extracellular region	74	7.80E-04	6.70E-03
cell fraction	45	1.20E-03	1.00E-02
large ribosomal subunit	8	1.50E-03	1.20E-02
nucleolus	32	1.80E-03	1.40E-02
secretory granule	13	2.20E-03	1.70E-02
proteinaceous extracellular matrix	18	3.40E-03	2.60E-02
cortical actin cytoskeleton	5	5.20E-03	3.80E-02
cell cortex part	8	5.10E-03	3.80E-02
eukaryotic translation initiation factor 3 complex	4	5.90E-03	4.20E-02
nuclear envelope	13	6.20E-03	4.20E-02
eukaryotic translation elongation factor 1 complex	3	6.10E-03	4.30E-02
endoplasmic reticulum part	18	7.60E-03	5.00E-02



**Figure 3.3.27** Graphical representation of Table 3.3.30 depicting GO cellular component enrichment and their corresponding adjusted p-value of  $\leq 0.05$  (Benjamini) for differentially-expressed host cell proteins identified in the conditioned media of a DP12 cell line at 37°C when 24hrs was compared to 144hrs (post-temperature-shift for a culture grown in parallel).

**Table 3.3.31** GO cellular component enrichment for differentially-expressed host cell proteins identified in the conditioned media of a DP12 cell line at 31°C when 24hrs was compared to 144hrs (post-temperature-shift) analysed using quantitative label-free LC/MS. An adjusted p-value of  $\leq 0.05$  (Benjamini) was used as a statistical cut off to generate the list of enriched cellular components.

Term	Gene Count	P-Value	Benjamini
cytosol	128	1.10E-43	3.50E-41
ribonucleoprotein complex	64	2.50E-26	4.10E-24
cytosolic part	37	3.50E-25	3.90E-23
cytosolic ribosome	28	1.80E-23	1.50E-21
melanosome	26	9.10E-20	6.00E-18
pigment granule	26	9.10E-20	6.00E-18
ribosomal subunit	28	9.40E-18	5.20E-16
ribosome	33	3.70E-16	1.60E-14
cytosolic small ribosomal subunit	17	7.00E-16	2.80E-14
lysosome	30	9.00E-14	3.30E-12
lytic vacuole	30	9.00E-14	3.30E-12
small ribosomal subunit	18	1.70E-13	5.70E-12
vacuole	31	1.50E-12	4.60E-11
membrane-bounded vesicle	43	3.70E-10	1.00E-08
cytoplasmic membrane-bounded vesicle	42	4.90E-10	1.20E-08
chaperonin-containing T-complex	7	1.70E-09	4.10E-08
vesicle	45	5.60E-09	1.20E-07
cytoplasmic vesicle	43	1.40E-08	3.00E-07
proteasome complex	13	3.30E-08	6.40E-07
heterogeneous nuclear ribonucleoprotein complex	8	9.90E-08	1.80E-06
intracellular non-membrane-bounded organelle	105	2.90E-07	5.10E-06
non-membrane-bounded organelle	105	2.90E-07	5.10E-06
cytosolic large ribosomal subunit	10	3.30E-07	5.50E-06
proteasome core complex	8	3.70E-07	5.80E-06
membrane-enclosed lumen	81	6.90E-07	1.00E-05
organelle lumen	79	1.30E-06	1.80E-05
endoplasmic reticulum lumen	12	5.20E-06	7.10E-05
large ribosomal subunit	11	6.60E-06	8.70E-05
platelet alpha granule lumen	9	7.20E-06	9.20E-05
cytoplasmic membrane-bounded vesicle lumen	9	1.30E-05	1.50E-04
soluble fraction	23	1.60E-05	1.90E-04
vesicle lumen	9	1.80E-05	1.90E-04
actin cytoskeleton	21	1.70E-05	2.00E-04
intracellular organelle lumen	73	2.70E-05	2.90E-04
spliceosome	14	3.00E-05	3.10E-04
ER-Golgi intermediate compartment	8	6.80E-05	6.80E-04
platelet alpha granule	9	7.70E-05	7.50E-04
nucleolus	36	8.70E-05	8.20E-04
extracellular region part	44	1.80E-04	1.60E-03
extracellular matrix	20	1.30E-03	1.20E-02
cell surface	20	1.40E-03	1.20E-02
cortical cytoskeleton	7	1.40E-03	1.20E-02
perinuclear region of cytoplasm	17	2.70E-03	2.30E-02
basement membrane	8	3.60E-03	2.90E-02
nuclear lumen	54	3.80E-03	3.00E-02
extracellular space	30	4.70E-03	3.60E-02
cell cortex part	8	5.00E-03	3.70E-02
eukaryotic translation elongation factor 1 complex	3	6.10E-03	4.40E-02



**Figure 3.3.28** Graphical representation of Table 3.3.31 depicting GO cellular component enrichment and their corresponding adjusted p-value of  $\leq 0.05$  (Benjamini) for differentially-expressed host cell proteins identified in the conditioned media of a DP12 cell line at 31°C when 24hrs was compared to 144hrs (post-temperature-shift for a culture grown in parallel).

### **3.3.6 Qualitative analysis of proteins identified by label-free LC/MS profiling of conditioned media from a DP12 CHO cell line**

As detailed in section 3.3.4, the lists of proteins identified by Progenesis label-free software are those that are deemed to change in relative abundance under a given set of conditions. However by generating a list of all expressed proteins identified in the conditioned media, it is possible to identify differences and similarities between proteomes under different environmental conditions or from different cell lines. This could have important implications for the ability of generic anti-HCP ELISAs to detect contaminating HCPs in the final drug substance of products manufactured using different CHO cell lines or culture conditions.

The RAW files obtained from LC/MS analysis of each conditioned media sample were searched against both the CHO NCBI and CHO BB databases using Mascot and SEQUEST search algorithms as detailed in section 3.3.5.

Protein lists obtained as a result of these searches were then combined and, as per section 3.3.5, in order for a protein to be considered “present” under a given condition, it had to be identified in a minimum of two biological samples by the Mass Spectrometer.

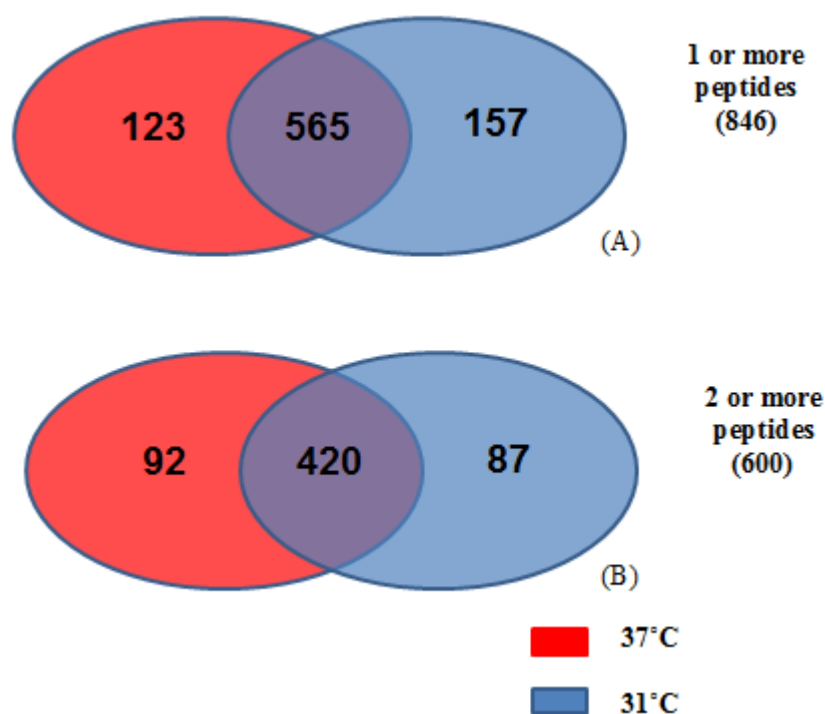
The following comparisons were made by overlapping the lists of identified proteins:

- DP12 37°C vs 31°C Cultures
  - Proteins identified with 1 or more peptides
  - Proteins identified with 2 or more peptides

#### **3.3.6.1 Qualitative comparison of HCPs identified in temperature-shifted and non-temperature-shifted DP12 cultures**

When a comparison of proteins identified in temperature-shifted and non-temperature-shifted cultures from a DP12 cell line is made it can be seen that while there is a large number of proteins common to culture conditions, there is also a cohort of proteins that are also unique to each culture condition (Figure 3.3.29 A).

In order to further investigate this observation a second comparison was made based on proteins identified using two or more peptides. The use of a minimum of two peptides to identify a protein further strengthens the basis on which the identification was made. This second comparison (Figure 3.3.29 B) shows a similar result to that observed in Figure 3.3.29 A and again identifies a subpopulation of proteins unique to both temperature-shifted and non-temperature-shifted cultures. This result indicates that cell culture conditions can impact on the HCP profile expressed in a DP12 cell line. A full list of all the proteins identified in Figure 3.3.29 A and Figure 3.3.29 B are provided in Appendix C, Tables 3.3.36 - 3.3.40.



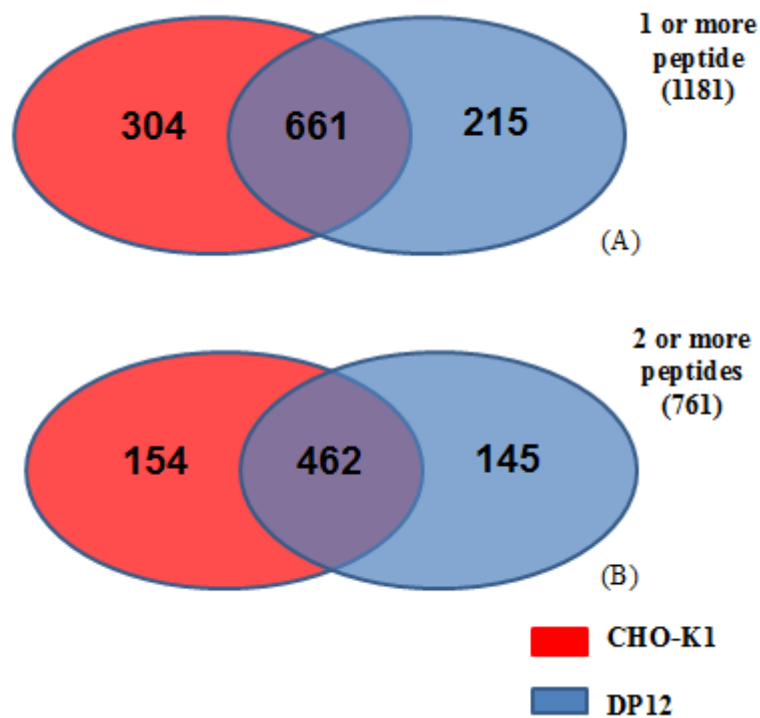
**Figure 3.3.29** Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line grown at 37°C and 31°C over 216 hours in shake flask culture. Samples were collected in biological triplicate and taken at 24, 72, 96, 120 and 144 hours post-temperature-shift. Venn diagram (A) represents those proteins identified using a minimum of 1 or more peptides. Venn diagram (B) is compiled from those proteins identified using a minimum of 2 or more peptides.

### 3.3.7 Qualitative comparison of HCPs identified in a CHO-K1 and DP12 cell line

A list of all HCPs identified in the cell culture supernatant of each cell line was generated using both CHO NCBI and CHO BB databases to identify proteins expressed under both temperature-shifted and non-temperature-shifted conditions. Due to the similar viabilities of CHO-K1 and DP12 cells at 37°C, as well as CHO-K1 and DP12 cells at 31°C, it was felt that protein lists generated from these cultures were comparable. With this in mind, both 37°C and 31°C lists were compared to identify proteins are common or unique to each cell line as follows:

- DP12 vs CHO-K1
  - Proteins identified with 1 or more peptides
  - Proteins identified with 2 or more peptides

Based on proteins identified using a minimum of one peptide for identification, 1,181 proteins were identified between both cell lines (Figure 3.3.30 A). Of these, 662 HCPs were common to both cell lines, 304 were unique to CHO-K1 and the DP12 cell line expressed 215 unique HCPs. A similar trend was observed when proteins whose identification was based on a minimum of two peptides was used to form the basis of the comparison (Figure 3.3.30 B). While 462 proteins were common to both cell lines, 154 HCPs were unique to CHO-K1 and 145 were unique to the DP12 cell line. Based on this set of results it appears that while there is a large number of HCPs expressed by both cell lines, there is also a subpopulation of HCPs that are unique to both CHO-K1 and DP12 cell lines. A full list of all the proteins identified in Figure 3.3.30 A and Figure 3.3.30 B are provided in Appendix C, Tables 3.3.41 - 3.3.45.



**Figure 3.3.30** Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line and a non-producing CHO-K1 cell line grown at 37°C and 31°C in shake flask culture. Samples were collected in biological triplicate from both cell lines and taken at 24, 72, 96, 120 and 144 hours post-temperature-shift from the DP12 cell line and 24, 72, 96 and 120hrs from the CHO-K1 cell line. Venn diagram (A) represents those proteins identified using a minimum of 1 or more peptides. Venn diagram (B) is compiled from those proteins identified using a minimum of 2 or more peptides.



### **3.3.7.1 Qualitative comparison of HCPs identified in a CHO-K1 and a DP12 cell line under temperature-shifted and non-temperature-shifted conditions**

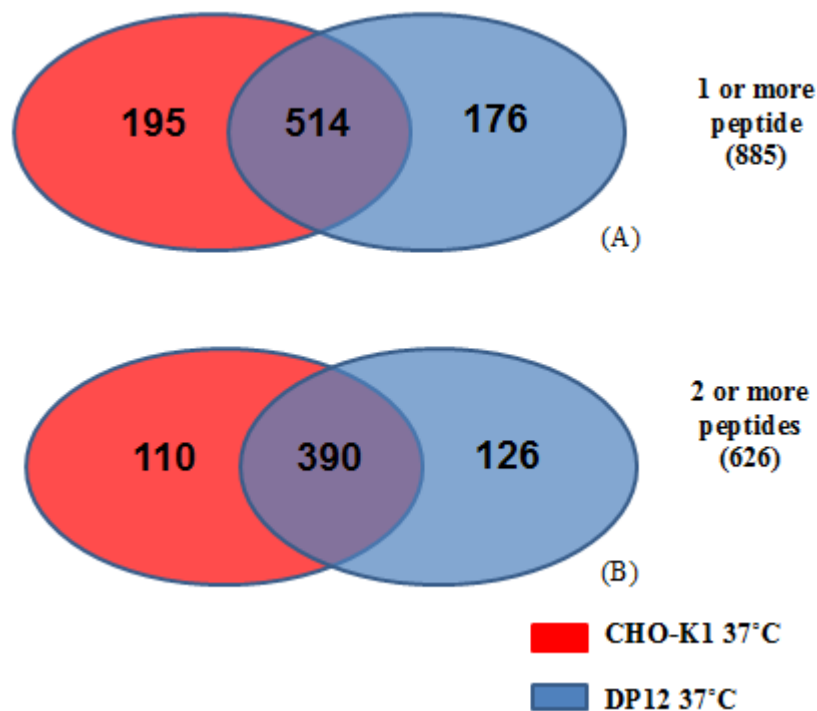
In order to decipher the findings in section 3.3.7, further comparisons were made between both cell lines based on proteins identified at (i) 37°C (ii) 31°C.

These comparisons were carried out as follows:

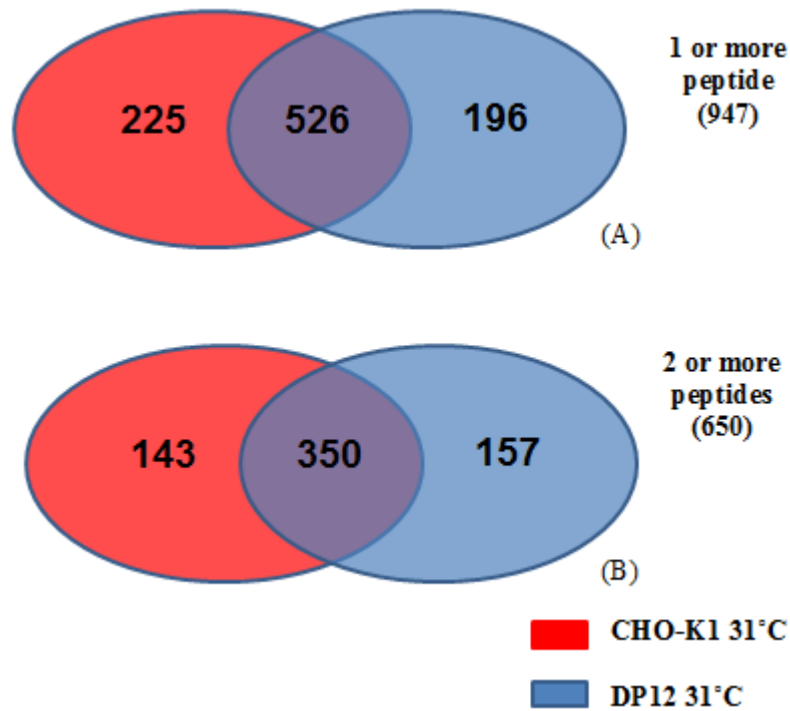
- DP12 vs CHO-K1 37°C
  - Proteins identified with 1 or more peptides
  - Proteins identified with 2 or more peptides
- DP12 vs CHO-K1 31°C
  - Proteins identified with 1 or more peptides
  - Proteins identified with 2 or more peptides

Based on proteins identified using a minimum of two peptides, 390 proteins were common to both CHO-K1 and DP12 cell lines at 37°C (Figure 3.3.31). 110 were found to be unique to CHO-K1 and 126 unique to DP12. The second comparison overlapped the lists of HCPs identified in CHO-K1 and DP12 grown under temperature-shift conditions. This comparison found that of those proteins identified using two or more peptides, 157 proteins were unique to the DP12 cell line and 143 were unique to CHO-K1 while 350 were common to both cell lines (Figure 3.3.32).

Taken together, this set of results appears to indicate that there are subpopulations of HCPs that are unique to different cell lines. In addition to this, these subpopulations can be influenced by the environmental conditions under which the cells are grown. A full list of all the proteins identified in Figure 3.3.31 and Figure 3.3.32 are provided in Appendix C, Tables 3.3.46 - 3.3.55.



**Figure 3.3.31** Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line and a non-producing CHO-K1 cell line grown at 37°C in shake flask culture. Samples were collected in biological triplicate from both cell lines and taken at 24, 72, 96, 120 and 144 hours post-temperature-shift from the DP12 cell line and 24, 72, 96 and 120hrs from the CHO-K1 cell line. Venn diagram (A) represents those proteins identified using a minimum of 1 or more peptides. Venn diagram (B) is compiled from those proteins identified using a minimum of 2 or more peptides.



**Figure 3.3.32** Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line and a non-producing CHO-K1 cell line grown at 31°C in shake flask culture. Samples were collected in biological triplicate from both cell lines and taken at 24, 72, 96, 120 and 144 hours post-temperature-shift from the DP12 cell line and 24, 72, 96 and 120hrs from the CHO-K1 cell line. Venn diagram (A) represents those proteins identified using a minimum of 1 or more peptides. Venn diagram (B) is compiled from those proteins identified using a minimum of 2 or more peptides.

### **3.3.8 Utilisation of CHO specific proteomic databases in the identification of HCPs in conditioned media samples from a non-producing CHO-K1 cell line and an IgG secreting DP12 cell line.**

The work detailed in this thesis has benefited from the relatively recent publication of two CHO specific proteomic databases (Meleady et al. 2012; Xu et al. 2011). As discussed in section 1.5.5.1, the CHO NCBI proteomic database was generated from CHO-K1 genome information while the CHO BB database is primarily based on information from cDNA expressed in a variety of CHO cell lines grown under different culture conditions, therefore, the use of both databases should result in increased coverage of the CHO proteome (Meleady et al. 2012). This section outlines the advantages of using both databases in the examination of mass spectral data acquired during the proteomic analysis of CHO conditioned media.

A list of proteins identified in the cell culture supernatant of a CHO-K1 and a DP12 cell line was compiled following mass spectral analysis. In order for a protein to be considered “identified” a search of the RAW data against one of the CHO databases previously mentioned must detect that protein in a minimum of two biological samples from both 37°C and 31°C cultures. The protein also had to be detected in twice in a given database in order to be considered “identified”, proteins that were detected only once were removed from further analysis. The protein samples were collected in biological triplicate and taken at 24, 72, 96, 120 and 144 hours post-temperature-shift as per (Figure 3.3.1).

Both databases were compared overlapping the lists of identified proteins in the CHO-K1 and DP12 cell line as follows:

- CHO-K1
  - Proteins identified with only 1 peptide
  - Proteins identified with 2 or more peptides
  - Proteins identified with 1 or more peptides
- DP12
  - Proteins identified with only 1 peptide
  - Proteins identified with 2 or more peptides
  - Proteins identified with 1 or more peptides

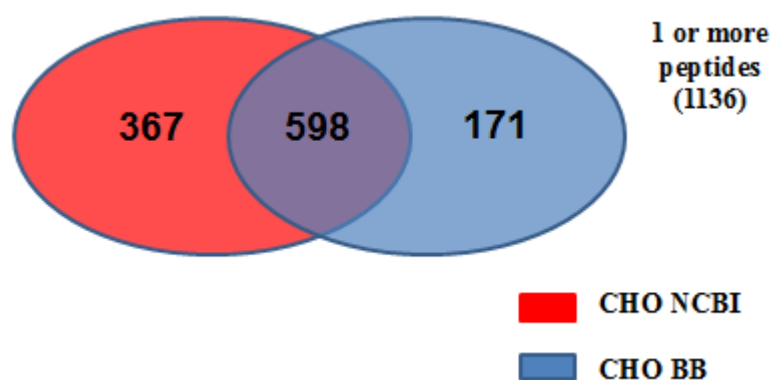
### **3.3.8.1 Comparison of CHO NCBI and CHO BB proteomic databases in the identification of HCPs from a non-producing CHO-K1 cell line**

As can be seen in Table 3.3.56, when the RAW files obtained from mass spectral analysis were searched against the CHO NCBI proteomic database a total of 966 proteins were identified in the conditioned media of a CHO-K1 cell line. This included 364 proteins identified using one peptide and 602 identified using two or more peptides. A search of the same files against the CHO BB database identified 769 proteins, of which 225 were detected by one peptide and 544 were identified using two or more peptides.

A second comparison was then made based on overlapping the list of proteins identified in each database irrespective of how many peptides were used in the identification of that protein. A comparison between databases solely comparing 'one hit wonders' or proteins identified using two or more peptides was not carried out as there may be instances where a protein was identified by two or more peptides in one database and that same protein identified by one peptide in the other database. This would have led to an inaccurate tabulation of the proteins identified. When the lists of proteins identified using the CHO NCBI and CHO BB databases were overlapped 598 proteins were found to have been common to both databases (Figure 3.3.33). 368 proteins identified were unique to the CHO NCBI database and 171 were only detected using the CHO BB database. A full list of all the proteins identified in Figure 3.3.33 are provided in Appendix C Tables 3.3.57 - 3.3.59.

	1 Peptide	2 or more peptides	Total
CHO NCBI	364	602	966
CHO BB	225	544	769

**Table 3.3.56** Table displaying the number of proteins identified in the conditioned media of a CHO-K1 cell line grown under temperature-shift and non-temperature-shift cultures with samples taken at 24, 72, 96 and 120 hours post-temperature-shift. RAW files from the mass spectral analysis were searched against the CHO NCBI and CHO BB protein databases. Columns are divided into those proteins identified by 1 peptide and proteins identified by 2 or more peptides.



**Figure 3.3.33** Overlap of the list of proteins identified in the conditioned media of a non-producing CHO-K1 cell line. Samples were collected from both temperature-shift and non-temperature-shift cultures. Samples were taken at 24, 72, 96 and 120 hours post-temperature-shift. RAW files from mass spectral analysis were searched against CHO NCBI and CHO BB proteomic databases. Only proteins that were detected in at least two biological samples were counted. The Venn diagram represents those proteins identified using a minimum of 1 or more peptides.

### **3.3.8.2 Comparison of CHO NCBI and CHO BB proteomic databases in the identification of HCPs from an IgG secreting DP12 cell line**

Similar to section 3.3.8.1, RAW files obtained during mass spectral analysis of cell culture supernatant from a DP12 cell line were searched against the CHO NCBI and CHO BB proteomic databases. A total of 845 protein identifications were retrieved from the CHO NCBI database, including 261 proteins identified using one peptide and 584 identified using two or more peptides. A search against the CHO BB database identified 204 proteins with one peptide and 532 were identified using two or more peptides. This resulted in 736 proteins being identified using the CHO BB database (Summary Table 3.3.60).

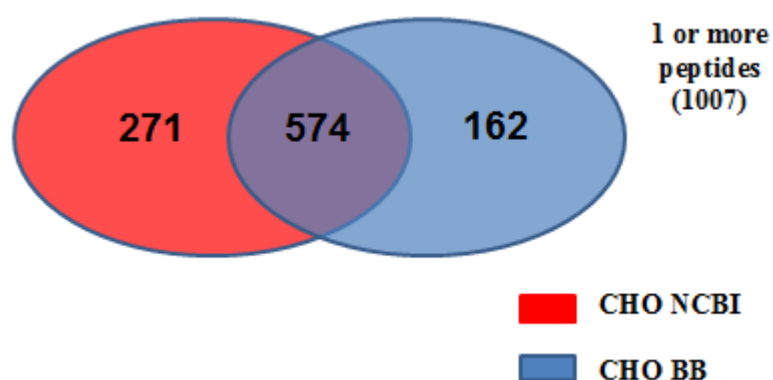
The number of proteins identified that were common or unique to each database was also compared. As per section 3.3.8.1 the comparison was made based on overlapping the list of proteins identified in each database irrespective of how many peptides were used in the identification of that protein. When the lists of proteins identified using the CHO NCBI and CHO BB databases were overlapped 574 proteins were found to have been common to both databases (Figure 3.3.34). 271 proteins identified were unique to the CHO NCBI database and 162 were only detected in the CHO BB database. This result shows that the availability of a second CHO database has greatly increased the number of HCPs identified in the sample set.

The number of proteins identified that are common or unique to each database were also compared. As per section 3.3.8.1 the comparison was made based on overlapping the list of proteins identified in each database irrespective of how many peptides were used in the identification of that protein. This comparison found that 574 of the proteins identified were common to both the CHO NCBI and CHO BB databases (Figure 3.3.34). 271 proteins were identified as being unique to the CHO NCBI database and 162 were only detected using the CHO BB database.

These results, coupled with those from section 3.3.8.1 indicate show that the use of a second database has increased the number of protein identifications by a minimum of 171 and 162 proteins respectively (depending on which database is being used). This shows that this work has clearly benefited from the availability of a second CHO database by greatly increasing the number of HCPs identified in each sample set.

	1 Peptide	2 or more peptides	Total
CHO NCBI	261	584	845
CHO BB	204	532	736

**Table 3.3.60** Table displaying the number of proteins identified in the conditioned media of a DP12 cell line grown under temperature-shift and non-temperature-shift cultures with samples taken at 24, 72, 96, 120 and 144 hours post-temperature-shift. RAW files from the mass spectral analysis were searched against the CHO NCBI and CHO BB protein databases. Columns are divided into those proteins identified by 1 peptide and proteins identified by 2 or more peptides.



**Figure 3.3.34** Overlap of the list of proteins identified in the conditioned media of an IgG secreting DP12 CHO cell line. Samples were collected from both temperature-shift and non-temperature-shift cultures. Samples were taken at 24, 72, 96, 120 and 144 hours post-temperature-shift. RAW files from mass spectral analysis were searched against CHO NCBI and CHO BB proteomic databases. Only proteins that were detected in at least two biological samples were counted. The Venn diagram represents those proteins identified using a minimum of 1 or more peptides. A full list of all the proteins identified in this figure can be found in Appendix C, Table 3.3.61 - 3.3.63.



### **3.3.9 Summary**

In accordance with our initial findings in the non-producing CHO-K1 cell line, quantitative label-free LC/MS analysis of an IgG producing DP12 cell line identified an increase in relative abundance of proteins with the potential to negatively affect product quality (proteases/glycosidases) at late stage culture compared to early stage culture under both temperature-shifted and non-temperature-shifted conditions. In addition, proteins with the ability to impact cell viability (apoptotic proteins) were found to be down-regulated under temperature-shifted conditions compared to non-temperature-shifted conditions at 144hrs. Furthermore, it was also shown that substantial differences exist in the HCP profile from producing and non-producing CHO cells when compared under both temperature-shifted and non-temperature-shifted conditions. This analysis utilised two CHO specific proteomic databases that were recently released. The use of both databases provided increased coverage of the secretome of both CHO cell lines allowing a greater number of HCPs to be identified than if only one database was available.

## 4.0 Phosphoproteomic Analysis of CHO cells

### Chapter overview

There have been a number of studies to profile the proteome of the CHO cell that have provided insights into the changes that cells undergo in response to temperature-shift. Through profiling the phosphoproteome however it was hoped that the cell signalling events that result in these changes would be revealed.

Initially an antibody array based approach was utilised to investigate a panel of kinase for changes in phosphorylation under temperature-shifted conditions compared to non-temperature-shifted conditions. This array contained antibodies specific to the phosphorylation sites of 26 kinases including ERK, AKT, JNK and p53, all of which are known to play key roles in controlling cellular function.

During this study, advances in the sequencing of the CHO genome facilitated the release of two CHO proteomic databases. This enabled a more global approach to profiling the phosphoproteome to be pursued using quantitative label-free LC/MS.

Quantitative label-free LC/MS analysis of phosphoprotein enriched CHO SEAP whole cell lysates collected from cells under temperature-shifted and non-temperature shifted conditions was conducted. In order to obtain site specific phosphorylation data, this experiment was followed by quantitative label-free LC/MS analysis of a Gallium oxide phosphopeptide enrichment of the phosphoprotein enriched samples.

While these label-free experiments provided some information on the differential-expression of phosphopeptides from proteins involved in cell cycle and translation, the data obtained was limited due to the low number of differentially-expressed phosphorylation sites identified.

To increase coverage of the CHO phosphoproteome, phosphopeptide enrichment of CHO SEAP whole cell lysate samples using titanium dioxide and iron oxide columns were assessed (i) qualitatively for the number of phosphopeptides with site specific information identified (ii) quantitatively for the differential expression of phosphopeptides.

The results show that this approach provided significantly enhanced coverage of the phosphoproteome. It was also shown that the availability of two CHO protein databases contributed to increased coverage of the phosphoproteome.

In addition to the identification of the differential expression of phosphopeptides from a number of interesting proteins such as NDRG1, eIF4G3 and DNA excision repair protein ERCC-6, GO analysis of quantitative data from the three phosphopeptide enrichment strategies was pooled together and identified enrichment of processes including translation, cell division and membrane organisation in temperature-shifted CHO cells.

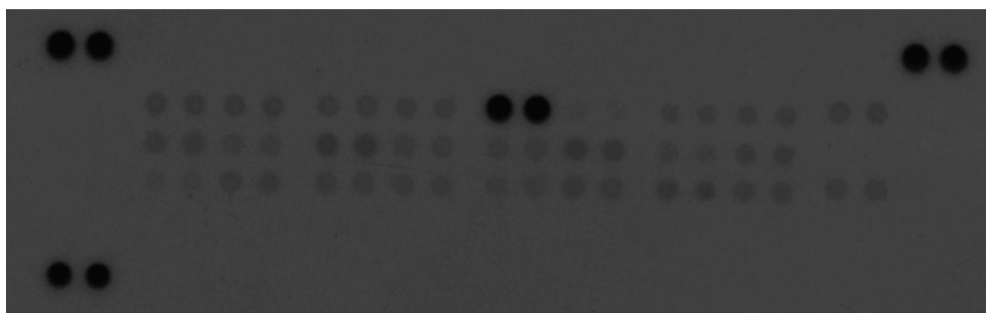
The work presented in this thesis is the first time that profiling of the CHO phosphoproteome by mass spectrometry has been reported. While future work to definitively determine total and phosphorylated protein abundance may be required, the application of phosphoproteomic profiling by mass spectrometry to CHO cells under temperature-shift has provided useful insights into the active cellular response to reduced culture temperature conditions.

## **4.1 Phospho-Kinase Array Analysis of CHO SEAP cells subject to temperature-shift**

In order to identify candidates potentially involved in the cellular response to temperature-shift an array based technology was used to assess a panel of kinases for changes in phosphorylation. This array contained antibodies for 26 different kinases, of which 9 were MAP kinases (Figure 4.1.1). (A full list of the kinases detected by the array is given in section 2.5.6.)

In this particular profiling experiment the array was used to detect changes in phosphorylation of the different kinases in CHO SEAP cells that had undergone temperature-shift (37°C vs 31°C). Samples were collected 72hrs after the cell culture flasks were seeded, 24hrs post-temperature shift (Figure 4.1.2). This time point was selected as it is still relatively early into temperature-shift and was more likely to capture changes in phosphorylation that were a direct result of culture temperature reduction. Using a later time point as the basis for analysis may have been complicated by other signalling events associated with cells entering into stationary or death phase. By using this early time point, while the cells were still in exponential growth, it was hoped that the signalling event that triggered growth arrest and subsequent increase in productivity might be captured. Samples used in this experiment for array analysis (CHO SEAP) and validation by western bolt (CHO SEAP and DP12) were grown for up to 96 hours post-temperature shift (Figure 4.1.2 and Figure 4.1.3). The increased productivity in temperature-shifted CHO SEAP cells can be seen in Figure 4.1.4. Western blots were carried out on biological duplicates of CHO SEAP cells and once on samples from the DP12 cell line.

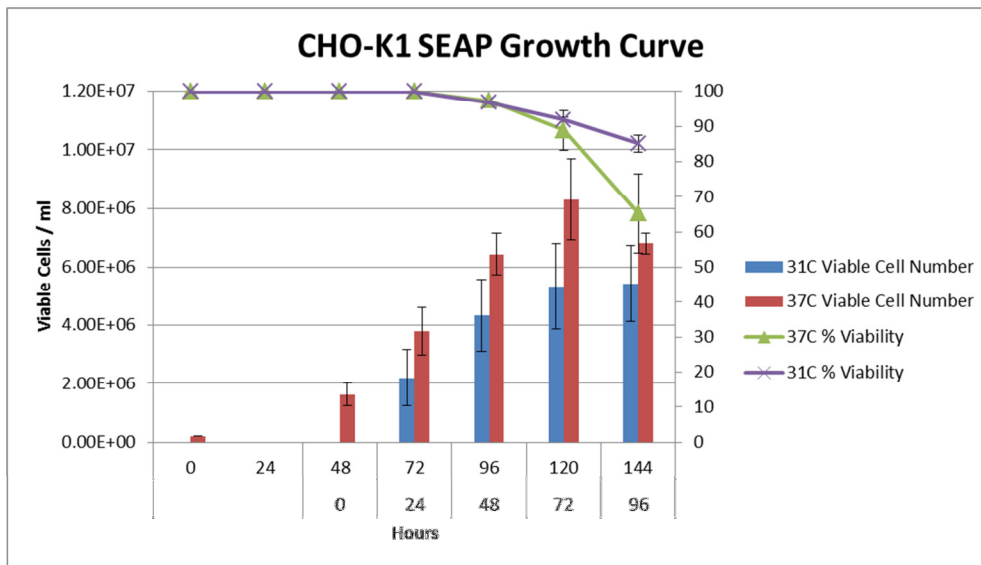
The experiment was conducted as a pair-wise comparison using biological duplicates. This was done to ensure robustness and reproducibility of the array. When a fold-change cut-off of 1.5 was applied to the results, 12 different kinases were deemed to have a potential change in phosphorylation status after temperature-shift (Table 4.1.1).



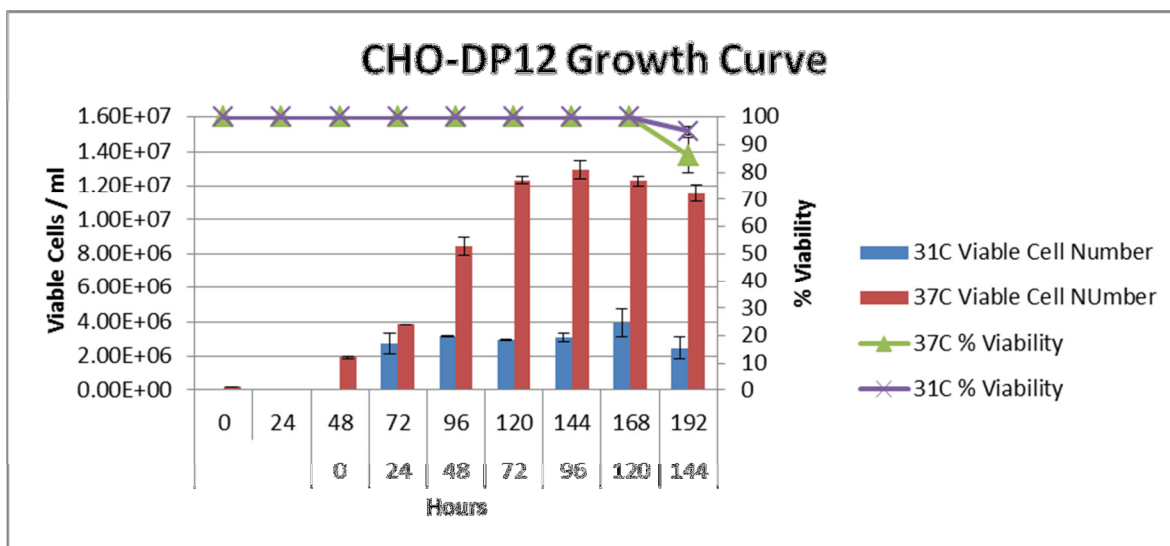
**Figure 4.1.1** Image of kinase array containing 26 different kinases including 9 MAP kinases. The array was used to look at the change in phosphorylation of the different kinases in CHO SEAP cells that had undergone temperature-shift (37C vs 31C) 24 hrs post-temperature shift.

	Phosphorylation Site	Fold Change Up at 31° C
<b>MSK2</b>	S360	1.59
<b>MKK6 (MEK6)</b>	S207/T211	1.93
<b>Akt 2</b>	S474	1.74
<b>ERK 1</b>	T202/Y204	3.17
<b>HSP27</b>	S78/S82	2.21
<b>JNK pan</b>	T183/Y185, T221/Y223	2.23
<b>JNK1</b>	T183/Y185	2.29
<b>JNK3</b>	T221/Y223	2.05
<b>p38<math>\alpha</math></b>	T180/Y182	2.42
<b>p38<math>\beta</math></b>	T180/Y182	2.09
<b>p38<math>\delta</math></b>	T180/Y182	1.74
<b>p53</b>	S46	2.97

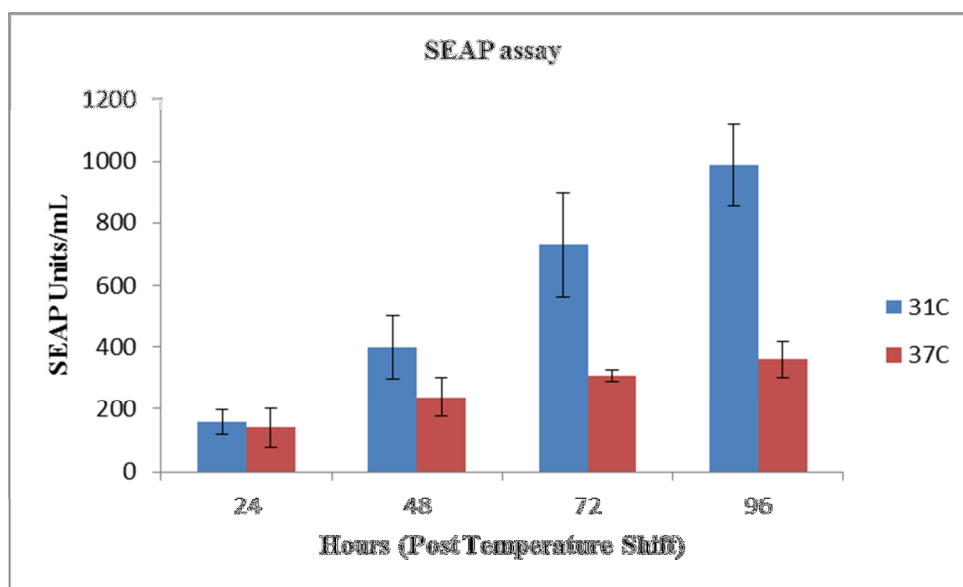
**Table 4.1.1** 12 Kinases identified using a Phosphokinase array as having a potential change in phosphorylation status in CHO SEAP cell 24hrs post-temperature shift. This table represents the average fold change observed from duplicate biological samples analysed on the array (n=2).



**Figure 4.1.2** Growth curve of CHO SEAP cell line under both temperature-shifted and non-temperature shifted conditions.



**Figure 4.1.3** Growth curve of DP12 cell line under both temperature-shifted and non-temperature shifted conditions.



**Figure 4.1.4** Graph showing the final titre of SEAP secreting CHO cells grown under temperature-shifted (31°C) and non-temperature shifted (37°C) conditions. Temperature-shifted culture was placed at 31°C after 48hrs in culture. Cells were collected for proteomic analysis after 24hrs, 48hrs, 72hrs and 96hrs at 31°C.

From this panel of 12 kinases, two (P38 and P53) had previously been identified as being involved in the cellular response to temperature-shift (Roobol et al. 2011). This provided an interesting partial validation of this array as the results obtained were in agreement with those obtained by Roobol et. al. As both P38 and P53 proteins had already been reported, this left eight candidates that were of interest, due to a potential change in their phosphorylation status as a result of temperature-shift (after removing P38  $\alpha$ ,  $\beta$ , and  $\gamma$ , and P53). From these, ERK, JNK and Mek6 were initially selected for validation by western blot. The antibodies used for validation of each phospho-kinase were specific to the same phosphorylation site as the antibody on the array. In fact, the primary reason for the omission of MSK6 from validation was that there was no antibody available to detect the phosphorylation of S360 on MSK6. Antibodies against the unphosphorylated form of ERK, JNK and MEK6 were used to detect the relative abundance of each kinase. From this it could be determined whether it was a change in the expression of the kinase or a change in phosphorylation that accounted for the results obtained using the array.

#### **4.1.1 Western blot validation of Total and phosphorylated ERK**

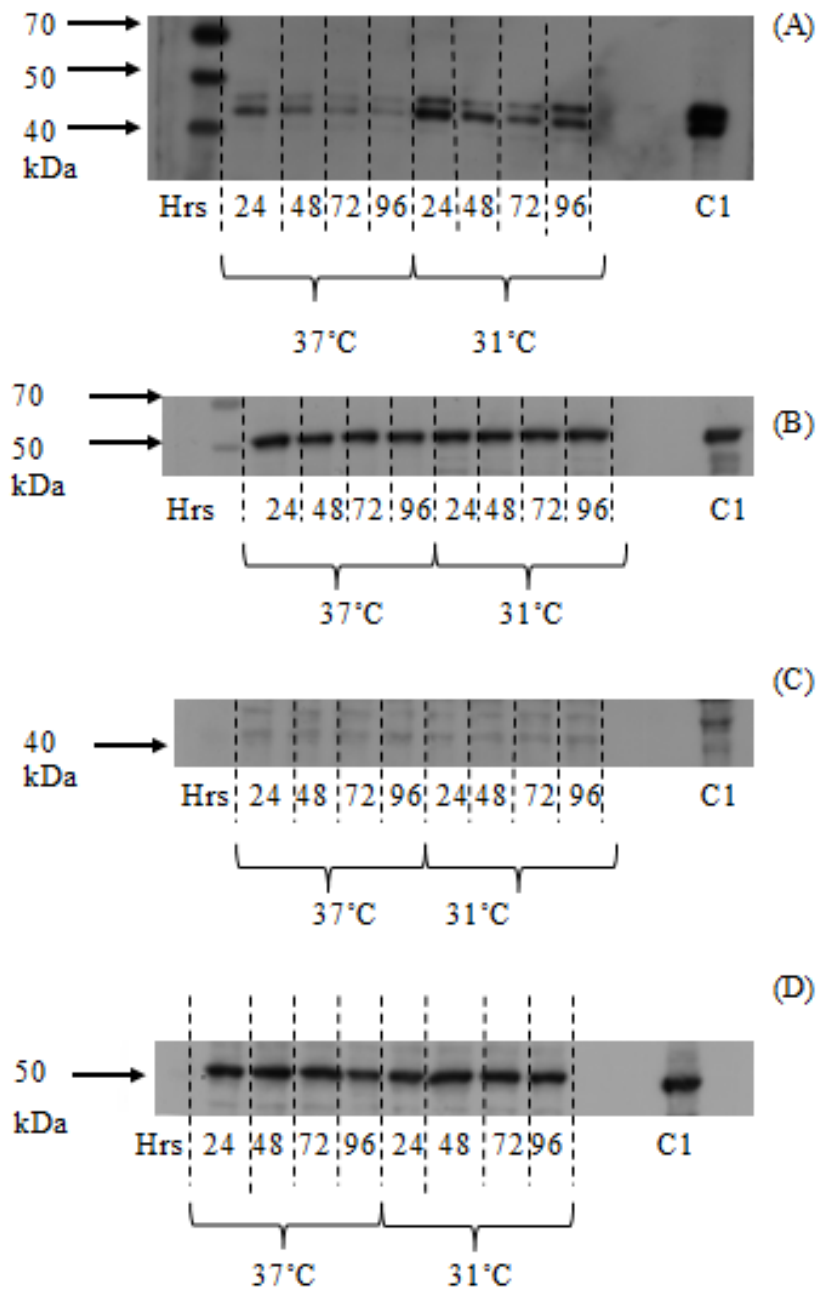
Based on the results from the array, of all the kinase targets, ERK1 showed the greatest increase in phosphorylation. Western blots were conducted in duplicate using four time points post-temperature shift, each at 24hr intervals, on CHO SEAP whole cell lysates. As can be seen in Figure 4.1.5 (A), the level of phosphorylation in ERK1 is clearly increased on day1 at 31°C compared to the same time point at 37°C. Interestingly, the level of phosphorylation remains increased over the four time points at 31°C compared to 37°C where it decreases over time. Although the levels of phosphorylated ERK at 37°C are greatly reduced, their decrease appears to coincide with the reduction in culture viability. A blot for the total ERK1 appears to indicate that its overall levels within the cell remains constant throughout the culture (Figure 4.1.5 C). This confirms that in the blot for the phosphorylated form of ERK1, it is in fact an increase in phosphorylation that is being observed and not just an increase in total protein.

It should however be noted that the expected molecular weight of total ERK1 does not correspond with the band detected. It was expected that a band would be detected at ~44kDa, in Figure 4.1.5 (C) it is just above 50kDa. A band at the same molecular weight is also detected in the control (MCF7 whole cell lysate). In order to see if this response was cell line specific, a western blot of phosphorylated and total ERK was also carried out on an IgG secreting CHO cell line (DP12). Although samples had been collected at all the time points indicated on the DP12 growth curve (Figure 4.1.3), samples used for western blot analysis spanned just the 24-96hr range, similar to the CHO SEAP samples. The behaviour of phosphorylated ERK in the DP12 cell line completely contrasted that observed in CHO SEAP. Phosphorylated ERK remained constant over the four time points at 37°C. After 24hrs at the reduced culture temperature, phosphorylated ERK appears slightly reduced compared to the same time point at 37°C (Figure 4.1.6 A). Over time at 31°C levels of phosphorylated ERK appears decrease. It should be noted that the two bands in Figure 4.1.5 A and Figure 4.1.6 A correspond to ERK1 (p44) (upper band) and ERK2 (p42) (lower band).

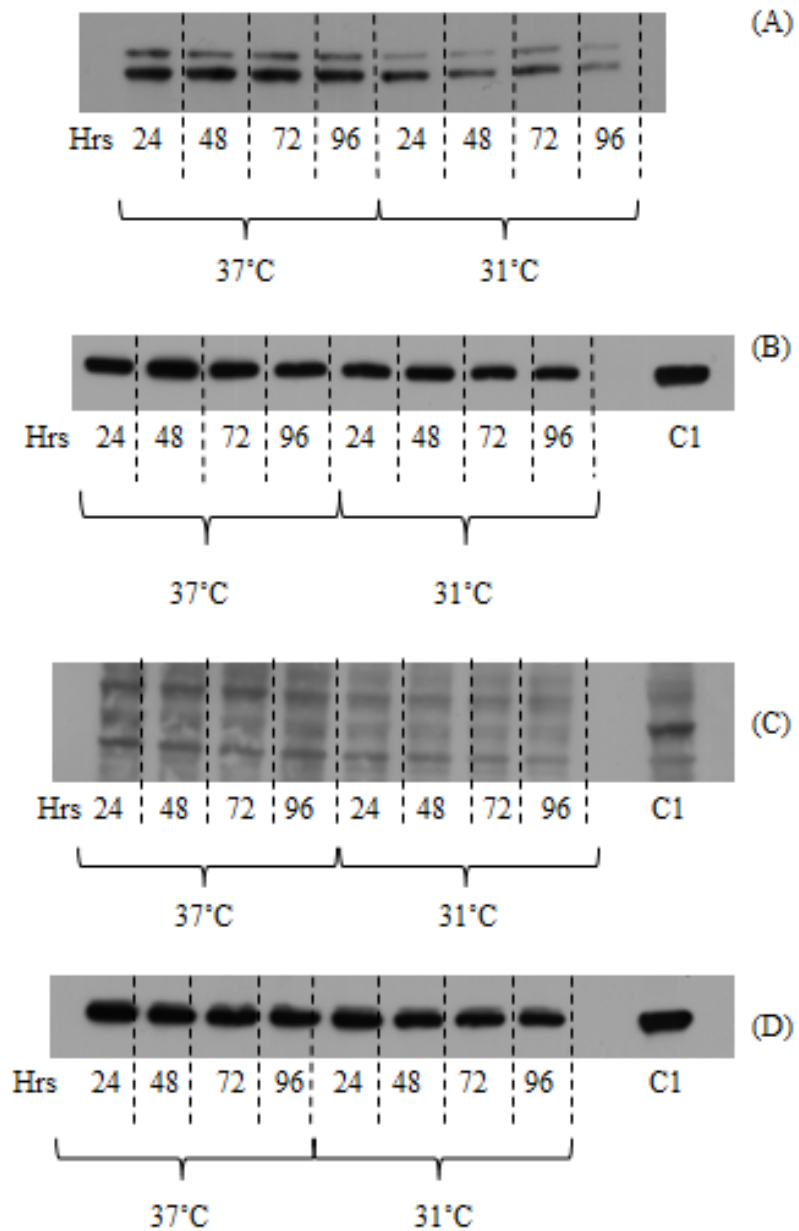


It is interesting to see that in both figures, the level of phosphorylation of ERK1 and ERK2 remains relatively similar indicating that the effect of temperature-shift does not appear to cause the cell to preferentially signal through ERK1 or ERK2. Of course, only knockout or inhibition based studies could definitively confirm this.

Although viability remained high over the 24-96hr period investigated, the growth rate of the DP12 at 31°C is almost completely stifled, this is particularly apparent when compared to the growth rate of the CHO SEAP cells at the same temperature.

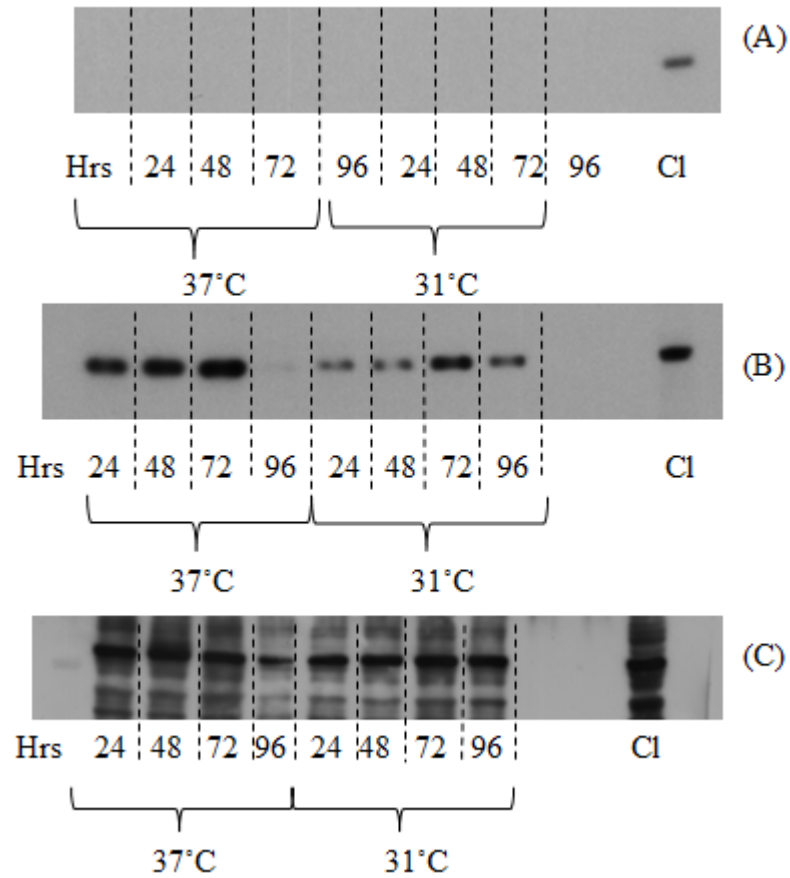


**Figure 4.1.5** Western blot image of Phosphorylated ERK (A) and total ERK (C) in a CHO SEAP cell line over time in culture under both temperature-shifted (31°C) and non-temperature (37°C) shifted conditions. Alpha tubulin (B and D) was used as a loading control. Time is given in hours post culture temperature reduction and C1 is whole cell lysate from MCF7 human cancer cell line as a control.



**Figure 4.1.6** Western blot image of Phosphorylated ERK1/2 (A) and total ERK1/2 (C) in a DP12 cell line over time in culture under both temperature-shifted (31°C) and non-temperature (37°C) shifted conditions. Alpha tubulin (B and D) was used as a loading control. Time is given in hours post culture temperature reduction and C1 is whole cell lysate from MCF7 human cancer cell line as a control.

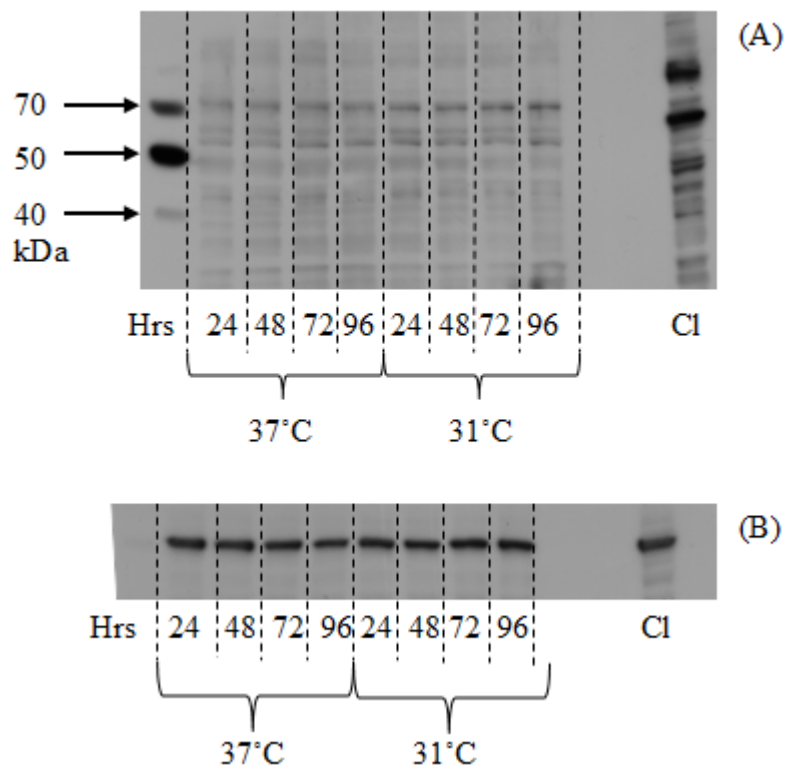
The different responses of the two cell lines to reduced culture temperature could underpin the different expression profiles of phosphorylated ERK. To further investigate change in phosphorylation of ERK in temperature-shifted CHO SEAP cells a western blot was conducted on MEK1/2, a key upstream activator of ERK1. Unfortunately, MEK1/2 could not be detected in the CHO SEAP cell line (Figure 4.1.7 A). A band detected at the expected molecular weight of 45kDa in the control MCF7 cell line indicated that the antibody was working. It was hypothesized that the failure to detect anything in CHO was possibly due to the antibody being raised against the Human version of MEK1/2 which may not have a similar amino acid sequence around the same phosphorylation site in CHO. However, by comparing the amino acid sequence of Human Mek1/2 against the CHO version it was discovered that the sequence for up to 11 amino acids either side of Ser217/221 was completely homologous. It therefore remains unclear why Mek1/2 was not detected in CHO. Immunoblotting for total MEK1/2 appeared to show the opposite trend to that seen in phosphorylated ERK1 i.e. an decrease in abundance at 31°C compared to 37°C (Figure 4.1.7 C). Without an immunoblot of the active form of MEK1/2 one cannot draw any firm conclusions. Due to the absence of an antibody that detects MEK1/2 in CHO, western blots for phosphorylated and total MEK1/2 were only conducted once.



**Figure 4.1.7** Western blot image of Phosphorylated Mek1/2 (A) and total Mek1/2 (B) in a CHO SEAP cell line over time in culture under both temperature-shifted (31°C) and non-temperature (37°C) shifted conditions. Alpha tubulin (C) was used as a loading control. Time is given in hours post culture temperature reduction and Cl is whole cell lysate from MCF7 human cancer cell line as a control.

#### 4.1.2 Western blot validation of Total and phosphorylated MEK6

Immunoblotting was also used to confirm the differential expression of phosphorylation on MEK6. Kinase array analysis indicated that phosphorylation of MEK6 increased 1.9-fold at 31°C. Analysis by western blot was less conclusive however. It was expected that a distinct band would be detected between 37-40kDa. As can be seen in Figure 4.1.8 (A), despite altering blocking, incubation and washing conditions, multiple bands were detected in CHO whole cell lysate samples, while western bolts of total MEK6 failed to detect any bands in CHO lysates (data not shown).



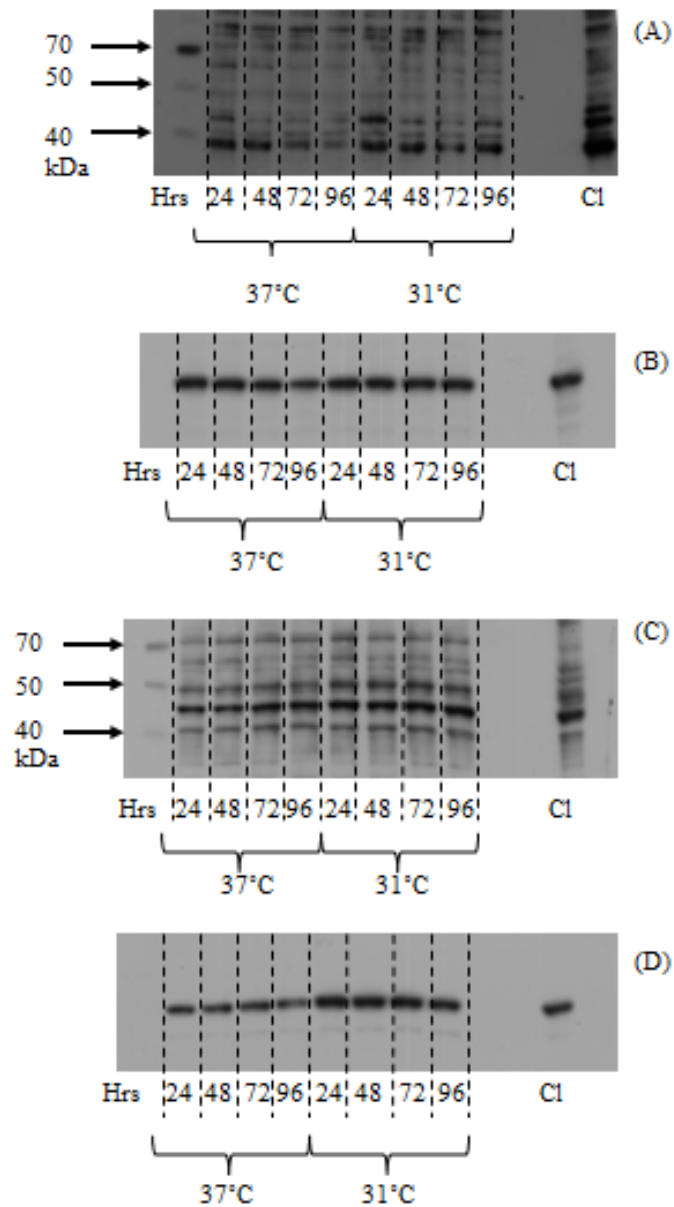
**Figure 4.1.8** Western blot image of Phosphorylated Mek6 (A) in a CHO SEAP cell line over time in culture under both temperature-shifted (31°C) and non-temperature (37°C) shifted conditions. Alpha tubulin (B) was used as a loading control. Time is given in hours post culture temperature reduction and Cl is whole cell lysate from MCF7 human cancer cell line as a control.

### 4.1.3 Western blot validation of Total and phosphorylated JNK

Western blots were conducted on total JNK1 and the phosphorylated form of JNK1 to determine whether changes in the levels of phosphorylation in JNK reported in the kinase array results were as a result of changes in the abundance of the kinase as opposed to changes in the levels of actual phosphorylation (Figure 4.1.9). The kinase array detected a 2.29-fold increase in phosphorylated JNK1 (T183/Y185), a 2.09-fold increase in phosphorylated JNK3 (T221/Y223) while the antibody specific for pan-phospho-JNK (T183/Y185 and T221/Y223) identified a 2.23-fold upregulation for the phosphorylated forms of JNK1/JNK2/JNK3. In Figure 4.1.9 (A) it can be seen that bands were detected at multiple molecular weights. The most prominent bands are located at ~37kDa, ~43kDa and, to a lesser extent > 70kDa. The predicted molecular weight of phosphorylated JNK1 is 48kDa, however distinct bands were detected at this molecular weight. Similarly in the immunoblot for total JNK1 numerous bands were detected at different molecular weights (Figure 4.1.9 C). It was expected that two well defined bands at 46kDa and 54kDa would be observed in accordance with the predicted molecular weight of the kinase. While there is a prominent band at ~46kDa, other distinct bands at 40kDa and 49kDa raise questions as to the specificity of the antibody. Assuming that the band detected at 46kDa corresponds to p46 JNK, then relative abundance of the kinase appears to remain constant between both temperatures over all four time points. In the immunoblot of phosphorylated JNK (Figure 4.1.9 A) the band at ~43kDa appears to be greatest at 24hrs in 31°C when compared to its counterpart at 37°C.

Since the kinase array was used to detect samples taken at the 24hr time point, it is possible that the increased abundance of this detected band accounts for the 2.3-fold increase in JNK observed in the initial profiling experiment. However given the amount of non-specific binding seen on some of the immunoblots it is possible that some of the results obtained were as a result of non-specific binding on the array.

It should be noted that none of the non-specific binding observed in the JNK or MEK6 immunoblots are believed to be caused by the secondary antibody. This is because blots incubated with secondary antibody only did not indicate any binding was taking place (data not shown).

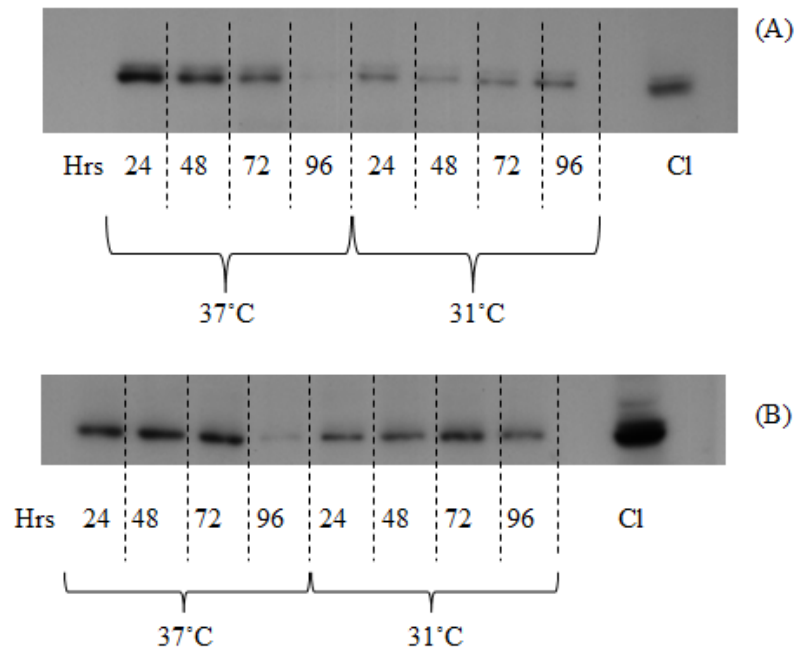


**Figure 4.1.9** Western blot image of Phosphorylated JNK (A) and total JNK (C) in a CHO SEAP cell line over time in culture under both temperature-shifted (31°C) and non-temperature (37°C) shifted conditions. Alpha tubulin (B and D) was used as a loading control. Time is given in hours post culture temperature reduction and Cl is whole cell lysate from MCF7 human cancer cell line as a control.

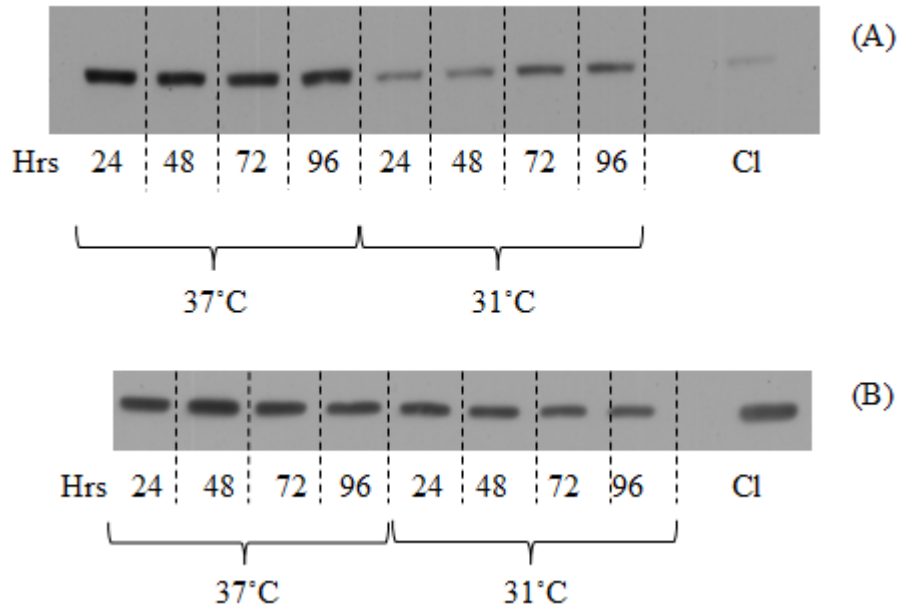


#### **4.1.4 Western blot validation of Total and phosphorylated AKT1**

Although AKT1 did not show any differential expression in phosphorylation during array profiling, it was nonetheless worth investigating to see if the results by western blot reflected those obtained by the array i.e. no fold change. It was also used to determine if AKT1 changed in phosphorylation at a time point beyond that measured by the array. Western blots of AKT1 were conducted in both CHO SEAP and DP12 cell lines as one-off experiments. As can be seen in Figure 4.1.10 the phosphorylation of AKT1 remains constant at 24, 48 and 72hrs but then greatly reduces at 96hrs at 37°C. This is in contrast to 31°C where the phosphorylation of AKT1 remains reduced over the 96hrs post-temperature shift. However when the total abundance of AKT1 is examined (Figure 4.1.10 B), it can be seen that it is in fact a change in the abundance of AKT that is giving rise to the appearance of a change in the phosphorylation of the kinase. Similar results were observed in the DP12 cell line (Figure 4.1.11). Phosphorylation of AKT1 is greatly reduced for the 96hrs post- temperature shift at 31°C compared to 37°C (Figure 4.1.11 A) and relative abundance of total AKT1 remains constant over the four time points at both temperatures in the DP12 cell line (Figure 4.1.11 B). This result indicates that there is a specific reduction in the phosphorylation of AKT1 in temperature-shifted DP12 cells. Taken together, both results provide some evidence that either as a consequence of reduced phosphorylation or a reduction in total kinase abundance that there may be some specific signalling taking place via AKT1 in recombinant protein producing CHO cells as a direct response to temperature-shift.



**Figure 4.1.10** Western blot image of Phosphorylated AKT1 (A) and total AKT1 (B) in a CHO SEAP cell line over time in culture under both temperature-shifted (31°C) and non-temperature (37°C) shifted conditions. Time is given in hours post culture temperature reduction and Cl is whole cell lysate from MCF7 human cancer cell line as a control.



**Figure 4.1.11** Western blot image of Phosphorylated AKT1 (A) in a DP12 cell line over time in culture under both temperature-shifted (31°C) and non-temperature (37°C) shifted conditions. Alpha tubulin (B) was used as a loading control. Time is given in hours post culture temperature reduction and Cl is whole cell lysate from MCF7 human cancer cell line as a control.

#### 4.1.5 Summary

As a result of this study, a number of kinases including ERK, JNK and AKT were identified as being potentially involved in the cellular response to temperature-shift. Despite optimising western blot conditions through altering blocking, incubation and wash times, efforts to definitively validate these results by western blot were hampered however by the non-specific binding of detection antibodies causing multiple bands to appear on the blots. While this approach could provide a useful platform to analyse changes in the phosphorylation profile of kinases involved in a wide range of cellular functions, the reliance on anti-human antibodies to detect kinases in CHO has proven to be sub-optimal. The recent sequencing of the CHO genome which heralded the advent of a CHO specific database enabled an alternative approach to be explored; the use of Mass Spectrometry to analyse the phosphoproteome of CHO cells.

## **4.2 Phosphoproteomic profiling of CHO SEAP cells subject to temperature-shift**

Phosphoproteomic profiling of CHO SEAP cells subject to temperature-shift was conducted in an effort to identify pathways involved in the cellular response to reduced culture temperature. Understanding such mechanisms could potentially aid in the identification of cell engineering targets to extend cellular viability, increase productivity or influence cellular growth rate.

This chapter details the use of phosphoprotein and phosphopeptide enrichment strategies coupled with LC/MS based technology to obtain both qualitative and quantitative phosphoproteomic information on CHO SEAP cells under reduced culture temperature conditions.

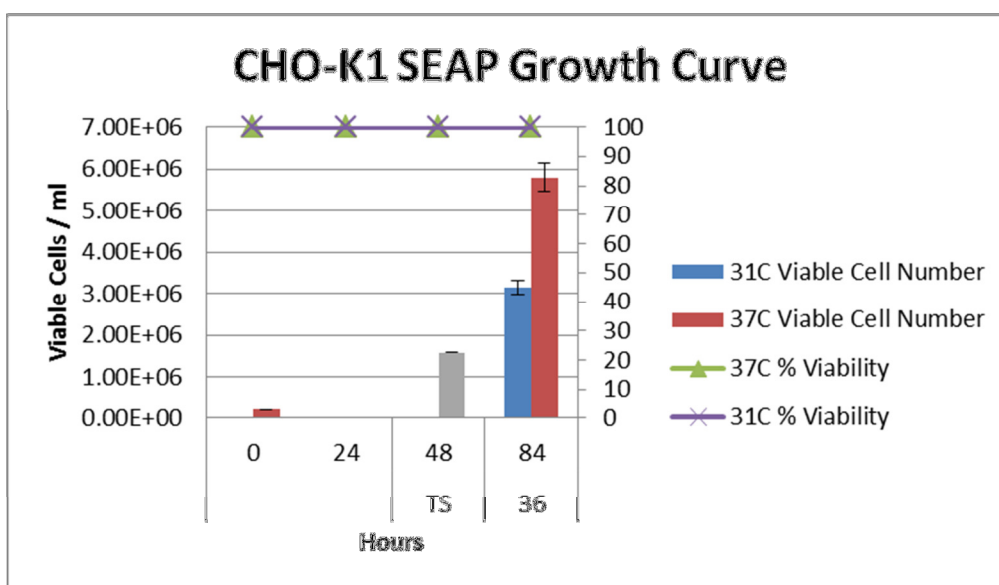
CHO SEAP cells collected from 37°C and 31°C cultures 36hrs post-temperature-shift were lysed and prepared for LC/MS analysis as follows:

- Whole cell lysate
- Phosphoprotein enrichment of whole cell lysate
- Gallium phosphopeptide enrichment of phosphoprotein enrichment
- Titanium dioxide phosphopeptide enrichment of whole cell lysate
- Iron oxide phosphopeptide enrichment of whole cell lysate

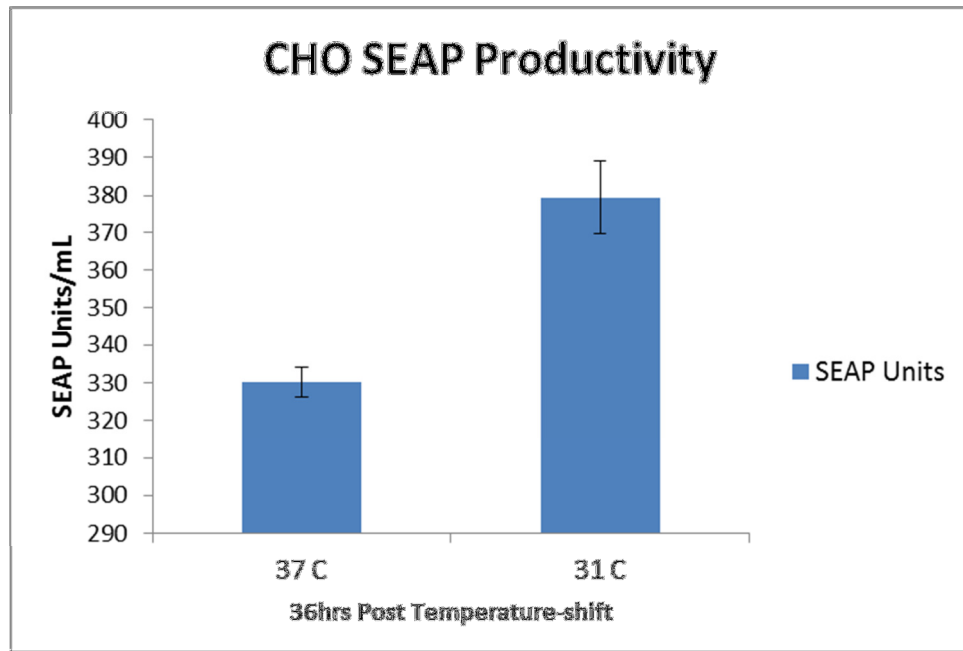
### **4.2.1 Growth of CHO SEAP cells subject to temperature-shift**

Eight shake flasks of CHO SEAP cells seeded at  $2 \times 10^5$  cell/ml were set up in a 50ml volume of serum free CHO-S-SFMII media in a batch culture. After 48 hours in culture, four flasks were selected for temperature-shift and were placed in a 31°C incubator for the remainder of the culture period. After 36hrs at 31°C CHO SEAP cells were collected from both temperature-shifted and non-temperature-shifted cultures as detailed in methods section 2.3.1 (Figure 4.2.1). 36hrs was selected because this is more than sufficient time for consistent phosphorylation events in direct response to culture temperature reduction to take place and allow any changes in phosphorylation that may be as a result of shock to subside.

Based on other growth profiling experiments, 36hrs post-temperature-shift is also still within the exponential phase of growth for this cell line. It was felt that this was a good point at which to collect cell samples as cell signalling events as a result of temperature-shift may become difficult to decipher from other signalling events later in the culture, e.g., reduced growth rate due to cells entering stationary phase. The increased productivity in temperature-shifted CHO SEAP cells can be seen in Figure 4.2.2.



**Figure 4.2.1** Growth curve of SEAP-secreting CHO cell line grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions. Temperature-shifted culture was placed at 31°C after 48hrs in culture. Cells were collected for proteomic analysis after 36hrs at 31°C.



**Figure 4.2.2** Graph showing the final titre of SEAP-secreting CHO cells grown under temperature-shifted (31°C) and non-temperature-shifted (37°C) conditions. Temperature-shifted culture was placed at 31°C after 48hrs in culture. Cells were collected for proteomic analysis after 36hrs at 31°C.

#### **4.2.2 Qualitative analysis of phosphopeptide enrichment of CHO SEAP**

In order to obtain further information on the phosphoproteome of temperature-shifted CHO cells it was decided that a more global profiling approach would be required. To achieve this, LC/MS was coupled with phosphoprotein and phosphopeptide enrichment strategies. It is only through the creation of two CHO specific databases as a result of recent advances in sequencing the CHO genome that phosphoproteomic profiling of CHO cells by Mass Spectrometry has been enabled. This section assesses the different phosphopeptide enrichment strategies used in this study, namely;

- Gallium phosphopeptide enrichment of phosphoprotein-enriched CHO whole cell lysate samples.
- Titanium dioxide phosphopeptide enrichment of CHO whole cell lysates.
- Iron oxide phosphopeptide enrichment of CHO whole cell lysates.

This section also details the contributions made to this study through the use of two CHO specific proteomic databases.

Following mass spectral analysis of all samples enriched using titanium, gallium and iron phosphopeptide enrichment methods, a list of phosphopeptides identified was compiled and compared as outlined above. Each compilation was based on the use of the SEQUEST algorithm to search all RAW data against both CHO NCBI and CHO BB databases. SEQUEST search criteria used as detailed in 2.4.6. An algorithm called PhosphoRS was used for the assignment of phosphorylation site localisation (Taus et al. 2011). Searches performed with PhosphoRS used a site probability of 0.99 (this corresponds to a false localization rate (FLR) of 1%), peptide to spectrum match (PSM) ambiguity was defined as unambiguous and peptide probability was <0.05.

Tallying of phosphopeptides and phosphorylation sites identified was conducted using a conservative counting method such that different charge states, oxidation of methionine residue and mis-cleaved variations were not considered as unique peptides. Unless otherwise stated these are the search parameters used to identify phosphopeptides in the comparisons that follow.

#### **4.2.2.1 Comparison of enrichment methods - Qualitative analysis of all phosphopeptides and phosphosites identified**

The use of different enrichment strategies to profile the phosphoproteome of CHO cells in this work significantly increased the number of phosphopeptides identified with site-specific information. This section catalogues all the phosphopeptides identified in this work and compares the different methods used for their enrichment.

Lists of identified phosphopeptides were generated by searching the RAW files obtained from LC/MS analysis of phosphopeptide enriched CHO SEAP samples as detailed in section 4.2.2. In order for a phosphopeptide to be considered identified it must have been detected in a minimum of two sample replicates for the respective enrichment method from either temperature-shifted or non-temperature-shifted cultures. Lists obtained were then compared to determine what phosphopeptides were common or unique to each enrichment method.

As a result of this comparison;

- 1,307 phosphopeptides were identified resulting in the detection of 1,480 phosphorylation sites
- Iron oxide enriched 960 phosphopeptides (1,076 phosphosites)
- Titanium dioxide enriched 867 phosphopeptides (1,001 phosphosites)
- Gallium oxide enriched 105 phosphopeptides (120 phosphosites)

Although Gallium oxide was by far the poorest enrichment method, it should be remembered that the starting material used for gallium oxide enrichment were pre-enriched phosphoprotein samples and not whole cell lysates as was the case with titanium or iron enrichment methods. It is therefore difficult to make a direct comparison between Gallium oxide and Titanium dioxide / Iron oxide enrichment methods.



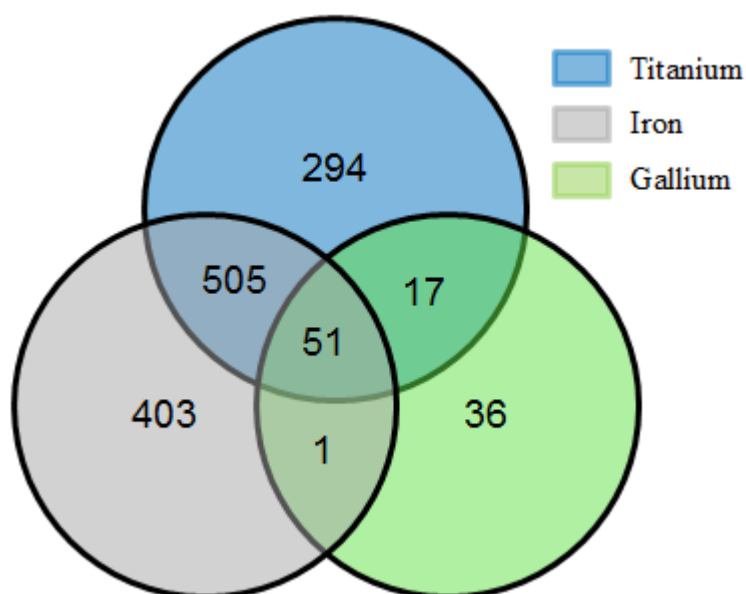
Nonetheless, for the purpose of understanding the contributions made by the three enrichment methods, Figure 4.2.3 shows the number of phosphopeptides that are common and unique to the different enrichment chemistries.

- Titanium and iron have the greatest number of shared phosphopeptide identification with 556 common to both enrichment methods.
- 402 phosphopeptides were uniquely identified by iron enrichment, followed by titanium with 294 and gallium with 36.
- 51 phosphopeptides were enriched by all three methods.

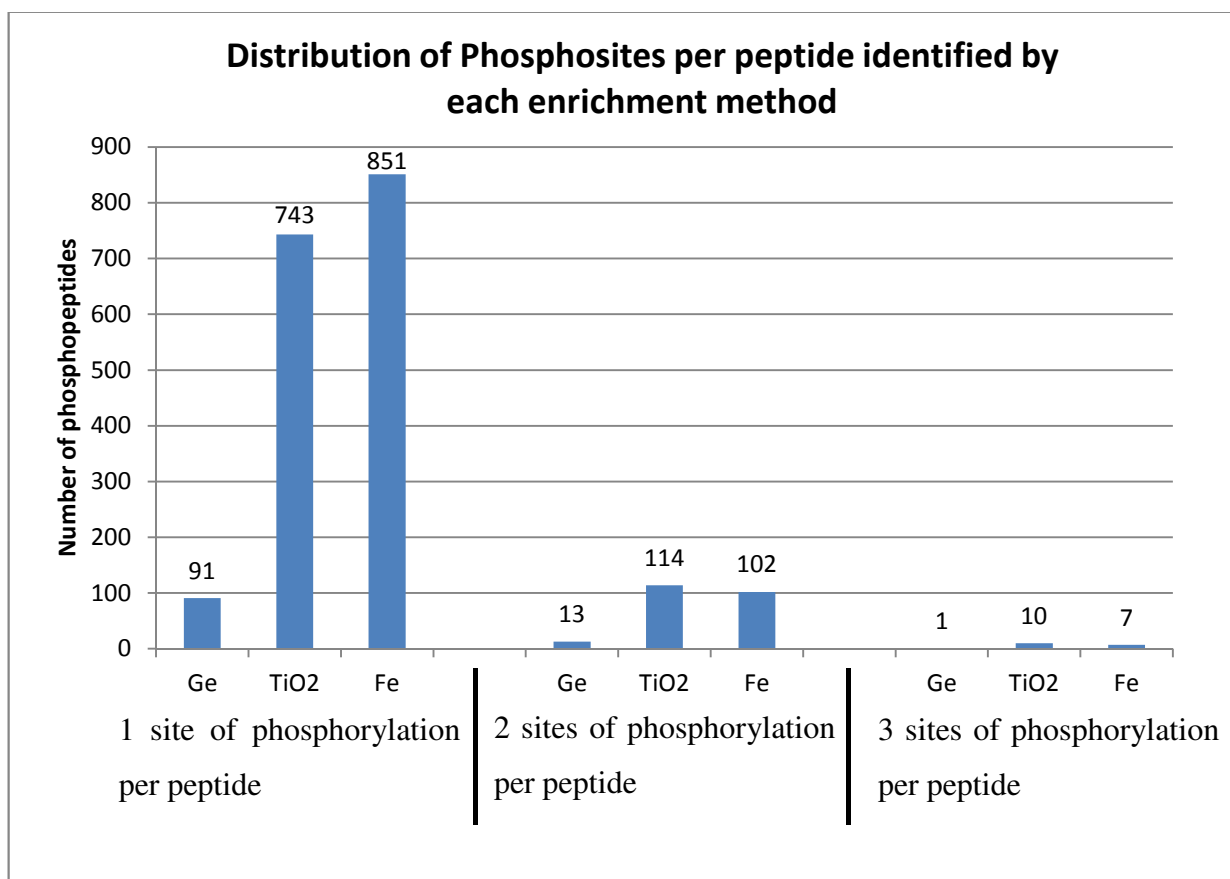
A full list of the phosphopeptides identified in Figure 4.2.3 is available in Appendix D Table 4.2.2 – Table 4.2.8. Also identified in this investigation were a number of multiply-phosphorylated peptide species (Figure 4.2.4). Singly phosphorylated peptides are typically identified using Gallium and Titanium enrichment methods while multiply-phosphorylated species are usually enriched using Iron based methods. In this instance however no such bias was observed with both methods enriching similar numbers of singly and multiply phosphorylated phosphopeptides.

Enrichment Method	Total number of phosphopeptides identified
Iron Oxide	960
Titanium Dioxide	867
Gallium Oxide	105

**Table 4.2.1** Table indicating the total number of phosphopeptides identified using each phosphopeptide enrichment method following enrichment of CHO SEAP cells grown under temperature-shifted and non-temperature-shifted conditions



**Figure 4.2.3** Venn diagram of the number of phosphopeptides that are common and unique to each phosphopeptide enrichment method following enrichment of CHO SEAP cells grown under temperature-shifted and non-temperature-shifted conditions. A full list of the phosphopeptides identified are available in Appendix D, Table 4.2.2 – Table 4.2.8.



**Figure 4.2.4** Graph of the number of phosphorylation sites per peptide as identified by titanium, iron and gallium phosphopeptide enrichment methods following LC/MS analysis of CHO SEAP cells subject to temperature-shift.

#### **4.2.2.2 Utilisation of CHO specific proteomic databases in the analysis of LC/MS data in a phosphoproteomic study**

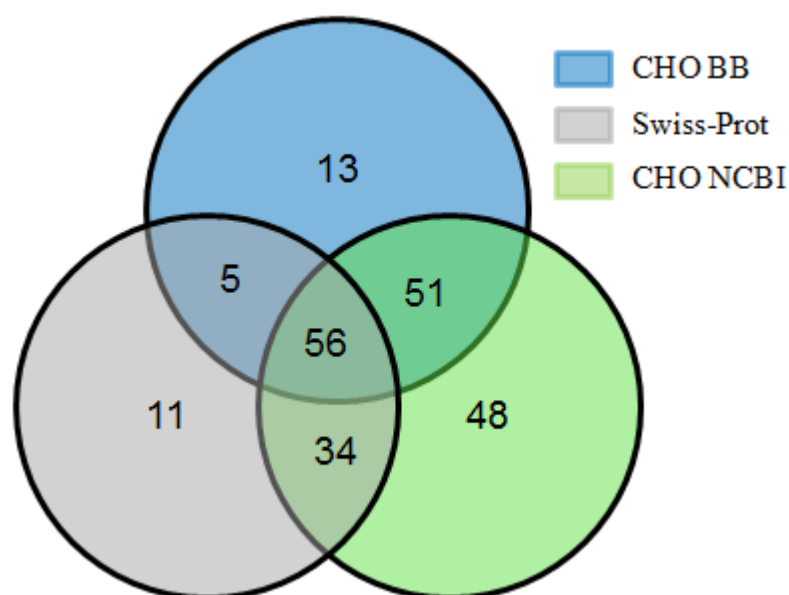
The relatively recent publication of two CHO specific proteomic databases (Meleady et al. 2012; Xu et al. 2011) has enabled the identification of 1,307 phosphopeptides and 1,480 sites of phosphorylation in this investigation. Here, the improvement in the number of phosphopeptides identified through the use of a CHO specific database is examined. In addition to this, the advantages of using a second database are also explored.

A simple comparison was set up in order to determine whether there were any significant differences in the number of phosphopeptides identified using the CHO databases compared to searching the data against human, mouse, rat and CHO protein information available in the UniProt/Swiss-Prot database. It should be noted that the CHO element of the UniProt/Swiss-Prot database is sequence information that was available prior to the advent of either the CHO NCBI or CHO BB database. As outlined in introduction section 1.5.5.1, the key difference between the two CHO databases is that the CHO NCBI proteomic database is primarily based on CHO-K1 genome information while the CHO BB database originates from cDNA expressed in a variety of CHO cell lines grown under different culture conditions (Meleady et al. 2012). The comparison was based on searching a raw data file from the LC/MS analysis of an Iron phosphopeptide enrichment of cells grown at 37°C against both CHO NCBI and CHO BB databases and against a UniProt/Swiss-Prot database containing protein sequence information for human, mouse, rat and CHO. At the time of analysis, the CHO NCBI database contained 24,383 sequences, CHO BB contained 15,285 sequences while 20,273 Human, 16,642 Mouse and 7,875 Rat sequences were contained in the UniProt/Swiss-Prot database. A list of identified phosphopeptides was generated by searching the RAW file as detailed in section 4.2.2. As mentioned previously, this file was also searched against a UniProt/Swiss-Prot database, this search was based on the same search criteria as section 4.2.2.

As a result of this search the following were identified;

- 208 phosphopeptides using both CHO NCBI and CHO BB databases
- 107 phosphopeptides using the UniProt/Swiss-Prot database
- 112 phosphopeptides were unique to the CHO specific databases

This result clearly shows that use of the CHO specific proteomic databases vastly improves the number of phosphopeptides identified (Figure 4.2.5). Although 11 phosphopeptides were unique to the UniProt/Swiss-Prot database, the use of both CHO databases identified 95 of the 106 phosphopeptides identified by the UniProt/Swiss-Prot database. This indicates that the use of CHO specific proteomic databases is not only comparable in the identifications obtained from a search that is primarily based on sequence homology, but it significantly increases phosphoproteomic coverage.



**Figure 4.2.5** Venn diagram showing the number of phosphopeptides identified when the raw file from LC/MS analysis of a phosphopeptide enriched CHO SEAP sample was searched against CHO BB, CHO NCBI and Swiss-Prot (human, mouse, rat and CHO {downloaded Jan 11}) databases. A full list of the phosphopeptides identified are available in Appendix D, Table 4.2.9 – Table 4.2.15.

Figure 4.2.6 (A) provides an example of the use of BLAST analysis to identify a protein based on sequence homology in a related species whose genomes are sequenced and so have a well annotated database. Identified following a search of the CHO databases, a phosphopeptide sequence from the protein Transcription activator BRG1 is searched against BLAST and is identified as having 100% sequence homology with human, mouse and rat. Using traditional search methodologies this protein would almost certainly have been identified if the LC/MS data file was searched against a database containing sequence information from any of these species. A BLAST search of a phosphopeptide sequence from the protein “LIM domain only protein 7” shows that while the peptide shares some sequence homology with the human version, there are still a number of amino acid substitutions which may lead to the peptide not being identified depending on database search cut-offs, see Figure 4.2.6(B). Another example of a phosphopeptide searched against BLAST is that of a peptide from Eukaryotic translation initiation factor 4 gamma 3. Although there is still significant sequence homology between the human and CHO version of the protein, an amino acid substitution of a serine for an alanine may have lead to an incorrect phosphorylation assignment on the peptide Figure 4.2.6 (C). (Symbols for the amino acids Serine, Tyrosine and Threonine are shown in Table 4.2.2.) These examples provide further evidence of the benefits of availing of a CHO-specific database when using an LC/MS based approach to characterise the phosphoproteome of CHO cells.

Amino Acid	Symbol
Serine	S
Tyrosine	Y
Threonine	T

**Table 4.2.16** Symbols denoting the amino acids Serine, Tyrosine and Threonine.

Alignment statistics for match #1					
Score	Expect	Identities	Positives	Gaps	
38.0 bits(82)	8e-05	11/11(100%)	11/11(100%)	0/11(0%)	
CHO	1	EVDYSDSLTEK	11		
		EVDYSDSLTEK			
Human	1376	EVDYSDSLTEK	1386		

Alignment statistics for match #1					
Score	Expect	Identities	Positives	Gaps	
38.0 bits(82)	8e-05	11/11(100%)	11/11(100%)	0/11(0%)	
CHO	1	EVDYSDSLTEK	11		
		EVDYSDSLTEK			
Mouse	1343	EVDYSDSLTEK	1353		

(A)

Alignment statistics for match #1					
Score	Expect	Identities	Positives	Gaps	
24.0 bits(49)	3.5	11/15(73%)	11/15(73%)	0/15(0%)	
CHO	1	KLSGDQLAAGAPESEK	15		
		KLSGDQ-AA-P--SK			
Human	1348	KLSGDQPAARTPRSEK	1362		

Alignment statistics for match #1					
Score	Expect	Identities	Positives	Gaps	
27.8 bits(58)	0.20	11/15(73%)	11/15(73%)	1/15(6%)	
CHO	1	KLSGDQLAAGAPESEK	15		
		KLSGD-L-AGAP--K			
Mouse	1189	KLSGD-LEAGAPKMK	1202		

(B)

Alignment statistics for match #1					
Score	Expect	Identities	Positives	Gaps	
29.9 bits(63)	0.041	10/14(71%)	11/14(78%)	0/14(0%)	
CHO	1	SRSTTELNDRATEK	14		
		SRSTTEL+D--T-K			
Human	1421	SRSTTELDYDSTYK	1434		

(C)

**Figure 4.2.6** Example of the use of BLAST analysis to identify a protein based on sequence homology in a related species with a well annotated database. (A) shows 100% sequence homology when a peptide from the CHO protein “Transcription activator BRG1” is searched against Human and mouse databases. (B) shows that although the correct protein “LIM domain only protein 7” is identified, a number of amino acid substitutions means there is less than 100% homology between the peptides. (C) indicates the potential for incorrect assignment of phosphorylation due to amino acid substitutions surrounding Threonine 12 (T12).

While the improvement in the number of phosphopeptides identified through the use of CHO specific databases have just been displayed, the importance of the use of the two available databases (from genomic DNA and expressed DNA) will now be outlined.

Lists of phosphopeptides were generated by searching the RAW files obtained from LC/MS analysis of phosphopeptide enriched CHO SEAP samples as detailed in section 4.2.2. In order for a phosphopeptide to be considered identified in a given database it had to be detected in a minimum of two independent samples for the respective enrichment method from either temperature-shifted or non-temperature-shifted cultures. Lists of phosphopeptides identified were then compared to determine what phosphopeptides were common or unique to each CHO database.

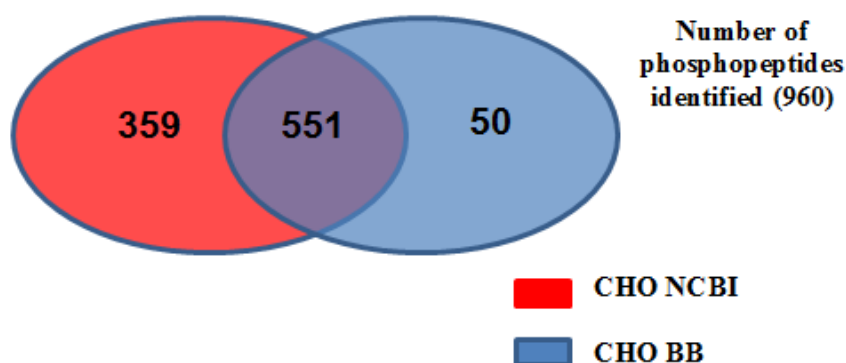
The following comparisons were made by overlapping the lists of identified phosphopeptides:

- CHO NCBI vs CHO BB database
  - Gallium phosphopeptide enrichment
  - Titanium phosphopeptide enrichment
  - Iron phosphopeptide enrichment

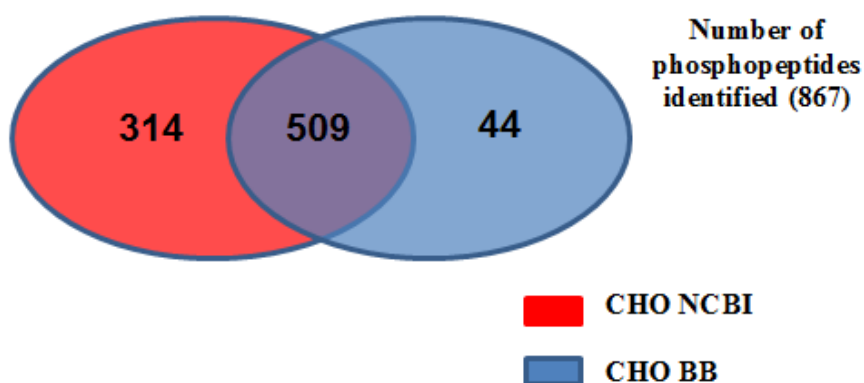
Figure 4.2.7 shows the distribution of phosphopeptide identifications between the CHO NCBI and CHO BB databases when LC/MS data from CHO SEAP whole cell lysate samples were enriched using iron oxide spin columns. Figure 3.3.34 shows the distribution of phosphopeptides from the same databases when titanium dioxide spin columns were used for phosphopeptide enrichment. Both comparisons show that the CHO BB identified a smaller number of unique phosphorylated peptides than the CHO NCBI database. Nonetheless, its use contributed an additional 50 phosphopeptide identifications in the case of iron oxide enrichment and increases titanium dioxide enrichment identifications by 44. In both instances the CHO BB database was capable of identifying >60% of all phosphopeptides identified. Although the number of phosphopeptides identified using gallium oxide enrichment is significantly smaller than the two previously mentioned methods, the use of the CHO BB database increases the phosphopeptide identification rate by just over 8% (Figure 4.2.9).



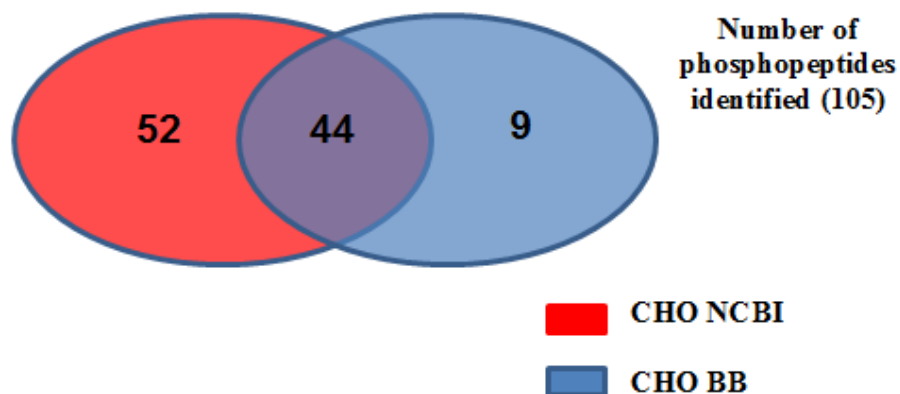
As evidenced by these examples, the use of a second database has provided a number of additional identifications that would otherwise not have been obtained. The resulting increase in proteome coverage is of obvious benefit for a study such as this. A full list of the phosphopeptides identified in Figure 4.2.7, Figure 3.3.34 and Figure 4.2.9 are available in Appendix D, Table 4.2.17 – Table 4.2.25.



**Figure 4.2.7** Overlap of the list of phosphopeptides identified after the iron oxide enrichment of a whole cell lysate from a CHO SEAP cell line. Four samples were collected from both temperature-shift and non-temperature-shift cultures 36hrs post-temperature-shift (eight samples total). All eight RAW files from mass spectral analysis were searched against CHO NCBI and CHO BB proteomic databases. Only phosphopeptides that were detected in at least two independent samples were considered to be identified. A full list of the phosphopeptides identified are available in Appendix D, Table 4.2.17 - Table 4.2.19.



**Figure 4.2.8** Overlap of the list of phosphopeptides identified after the titanium dioxide enrichment of a whole cell lysate from a CHO SEAP cell line. Four samples were collected from both temperature-shift and non-temperature-shift cultures 36 hours post-temperature-shift (eight samples total). All eight RAW files from mass spectral analysis were searched against CHO NCBI and CHO BB proteomic databases. Only phosphopeptides that were detected in at least two independent samples were considered to be identified. A full list of the phosphopeptides identified are available in Appendix D, Table 4.2.20 - Table 4.2.22.



**Figure 4.2.9** Overlap of the list of phosphopeptides identified after the gallium oxide enrichment of a whole cell lysate pre-enriched for phosphoproteins from a CHO SEAP cell line. Three samples were collected from both temperature-shift and non-temperature-shift cultures 36hrs post-temperature-shift (six samples total). The six RAW files from mass spectral analysis were searched against CHO NCBI and CHO BB proteomic databases. Only phosphopeptides that were detected in at least two independent samples were considered to be identified. A full list of the phosphopeptides identified are available in Appendix D, Table 4.2.23 - Table 4.2.25.

#### **4.2.2.3 Phosphopeptides uniquely identified in temperature-shifted or non-temperature-shifted cultures**

Qualitative lists obtained from the titanium dioxide and iron oxide phosphopeptide enrichment of CHO SEAP whole cell lysate samples and gallium oxide phosphopeptide enrichment of whole cell lysate samples pre-enriched for phosphoproteins were compiled. By comparing these lists it was hoped that phosphopeptides whose phosphorylation may be unique to temperature-shifted or non-temperature-shifted cultures could be identified. These lists may provide information on phosphopeptides that are uniquely phosphorylated at either culture temperature. This could indicate that a protein and, as a consequence, a pathway is turned 'on' or 'off' at a given temperature.

Lists of identified phosphopeptides were generated by searching the RAW files obtained from the LC/MS analysis of both culture temperatures and each enrichment method. In order for a phosphopeptide to be considered as unique to a given culture temperature it had to be identified in a minimum of three out of the four samples for that temperature for titanium and iron enriched samples and in a minimum of two out of the three samples in the case of gallium enriched samples. The lists of phosphopeptides that met these criteria were then compared to lists of all the phosphopeptides identified by all three enrichment methods at the opposing temperature. If any phosphopeptide from the 'potentially unique' list was found in any of the lists from all the phosphopeptides identified using the three enrichment methods at the opposing temperature it was no longer considered 'potentially unique' and removed from further analysis.

In addition to this, if a protein of one of the phosphopeptides from the 'potentially unique' list was identified as a protein of one of the phosphopeptides identified in the three enrichment methods at the opposing temperature then this phosphopeptide was no longer considered 'potentially unique' and removed from further analysis.

Based on these criteria, 64 phosphopeptides were identified as being ‘potentially unique’ in the 37°C culture with slightly less than half that number identified in the 31°C culture (Table 4.2.26). Interestingly, when Table 4.2.27 and Table 4.2.29 are compared it can be seen that titanium and iron enrichment methods both contained phosphopeptides from the proteins TBC1 domain family member 17 and PERQ amino acid-rich with GYF domain-containing protein 1 that are considered to be unique in the 37°C culture. Similarly, a phosphopeptide from the protein 60S ribosomal protein L12, considered to be unique in the 31°C culture, was also enriched using both titanium and iron enrichment methods (Table 4.2.28 and Table 4.2.30).

This analysis also identified a number of translation initiation factors as being unique to different culture temperatures, such as transcription initiation factor TFIID subunit 6-like and activating transcription factor 7-interacting protein 1 at 37°C and transcription factor jun-B-like, transcription factor Dp-1 isoform 1 and eukaryotic translation initiation factor 3 subunit A-like at 31°C.

Enrichment Method	Temperature	Number of unique phosphopeptides	Table
Titanium	37°C	28	Table 4.2.27
	31°C	13	Table 4.2.28
Iron	37°C	22	Table 4.2.29
	31°C	12	Table 4.2.30
Gallium	37°C	14	Table 4.2.31
	31°C	6	Table 4.2.32

**Table 4.2.26** This table shows the number of phosphopeptides that are considered to be ‘potentially unique’ to temperature-shifted and non-temperature-shifted cultures using gallium, titanium and iron enrichment methods. This table also contains the table number for the list of potentially unique phosphopeptides from each enrichment method and each culture temperature.

**Table 4.2.27** List of phosphopeptides that were only identified in CHO SEAP cells at 37°C using titanium dioxide phosphopeptide enrichment. Database denotes which CHO proteomic database the phosphopeptide was identified in (N= CHO NCBI) (BB=CHO BB). The table also indicates the number of independent biological replicates that the phosphopeptide was identified in. For this list each phosphopeptide was identified in a minimum of three biological replicates.

Database	Number of times identified	Peptide	Protein Name	Gene symbol
N/BB	3/4	GSQPGPAVESQS <b>S</b> PR	Band 4.1-like protein 2	EPB41L2
N/BB	4	TI <b>S</b> DGTISASK	Formin-binding protein 1-like	FNBP1L
N/BB	4	SD <b>S</b> LLSFR	Cdc42 effector protein 1	CDC42EP1
N/BB	4	IHRAS <b>S</b> DPGLPAEEP	CAD protein	CAD
N	4	EPSSP <b>S</b> PPISPLPLSTR	TBC1 domain family member 17	TBC1D17
N	4	SALLAQNGT <b>S</b> SLPR	Pleckstrin-likey-like domain family B member 1	PHLDB1
N	4	LK <b>S</b> PPHSPGNLmDN	Snurportin-1	SNUPN
BB	4	ESVPDFPL <b>S</b> PPK	Stathmin	STMN1
N/BB	3	SKDHFGLLEGDEESTmLEDSV <b>S</b> PK	Talin-1	TLN1
N/BB	3	GDLS <b>S</b> DVEEEEEEmDVDEATGAPK	Paired amphipathic helix protein Sin3a	SIN3A
N/BB	3	DTPTSAGPN <b>S</b> FNK	PEST proteolytic signal-containing nuclear protein	PCNP
N/BB	3	DTSTEALLK <b>S</b> SPEER	Transcriptional repressor p66 alpha	GATAD2A
N/BB	3	<b>S</b> PGGPGPLTLK	Zyxin	ZYX
N/BB	3	LSVHDLKPLD <b>S</b> PGR	Thyroid receptor-interacting protein 11	TRIP11
N/BB	3	AG <b>S</b> QLGLDHLR	Rab3 GTPase-activating protein catalytic subunit	RAB3GAP1
N/BB	3	<b>S</b> mEVDLNQAHmEDTPK	Maternal embryonic leucine zipper kinase	MELK
N/BB	3	VP <b>S</b> VESLFR	Golgin subfamily A member 4	GOLGA4
N/BB	3	GP <b>S</b> SVEDIK	Nucleophosmin	NPM1
N/BB	3	LLHEDLDES <b>S</b> DDDVDEK	Arf-GAP with SH3 domain, ANK repeat and PH domain-containing protein 2	ASAP2
N	3	VSALES <b>S</b> GPHQ <b>S</b> PALSK	FERM, RhoGEF and pleckstrin domain-containing protein 1-like	FARP1
N	3	IAS <b>S</b> ESDEQQWPEEK	Sister chromatid cohesion protein PDS5-like B	PDS5B
N	3	NP <b>S</b> PNPVTALPK	Zinc finger protein 280C	ZNF280C
N	3	KH <b>S</b> DPHLLER	Serine/threonine-protein kinase Nek10	NEK10
N	3	<b>S</b> PHQLLSPSSFS <b>S</b> PSAAPSQK	DNA polymerase alpha subunit B-like	POLA2
N	3	QDGEKEE <b>S</b> LSGV	high mobility group nucleosome-binding domain-containing protein 5-like	HMG5
N	3	ASIS <b>S</b> PmDEVPDSE <b>S</b> PIEK	Microtubule-associated protein 1B	MAP1B
N	3	SAQASS <b>S</b> SPALPR	Tyrosine-protein kinase ABL2	ABL2
N	3	<b>S</b> IEEGDGAFGR	PERQ amino acid-rich with GYF domain-containing protein 1	GIGYF1

**Table 4.2.28** List of phosphopeptides that were only identified in CHO SEAP cells at 31°C using titanium dioxide phosphopeptide enrichment. Database denotes which CHO proteomic database the phosphopeptide was identified in (N= CHO NCBI) (BB=CHO BB). The table also indicates the number of independent biological replicates that the phosphopeptide was identified in. For this list each phosphopeptide was identified in a minimum of three biological replicates.

Database	Number of times identified	Peptide	Protein Name	Gene symbol
N/BB	4	SV <u>S</u> VATGLNmmK	Brain-specific angiogenesis inhibitor 1-associated protein 2-like protein 1	BAIAP2L1
N	4	mQF <u>S</u> FEGPEK	IQ motif and SEC7 domain-containing protein 1	IQSEC1
N/BB	4	A <u>S</u> DPGLPAEEPK	CAD protein	CAD
N/BB	4	GNK <u>S</u> <u>S</u> PSPPDGSPAATPEIR	Myc box-dependent-interacting protein 1	BIN1
N/BB	3	VK <u>S</u> LEDEEAPR	tRNA pseudouridine synthase A, mitochondrial	PUS1
N	3	SY <u>S</u> LSELSVLQAK	SH3 domain-binding protein 4-like	SH3BP4
N	3	VEKPATIAGENG <u>S</u> PVLR	Inner centromere protein	INCENP
BB	3	TA <u>S</u> GSSVTSLEGPR	Protein NDRG1	NDRG1
N/BB	3	GHL <u>S</u> EGLVTK	Mitogen-activated protein kinase 6	MAPK6
N/BB	3	SF <u>S</u> PK <u>S</u> PLELGEK	NGFI-A-binding protein 2	NAB2
N/BB	3	<u>S</u> QVINVIGSER	Dedicator of cytokinesis protein 1	DOCK1
N/BB	3	IGPLGL <u>S</u> PK	60S ribosomal protein L12 -	RPL12
N/BB	3	IT <u>S</u> PLmEPSSIEK	Proteasome subunit alpha type-5	PSMA5

**Table 4.2.29** List of phosphopeptides that were only identified in CHO SEAP cells at 37°C using iron oxide phosphopeptide enrichment. Database denotes which CHO proteomic database the phosphopeptide was identified in (N= CHO NCBI) (BB=CHO BB). The table also indicates the number of independent biological replicates that the phosphopeptide was identified in. For this list each phosphopeptide was identified in a minimum of three biological replicates.

Database	Number of times identified	Peptide	Protein Name	Gene symbol
N/BB	4	RG <u>S</u> ISQEITK	Biorientation of chromosomes in cell division protein 1-like	BOD1L1
N/BB	4	<u>S</u> PLGGGGGSGASSQAACLK	Male-specific lethal 1-like	MSL1
N/BB	4	VIAISGDAE <u>S</u> PAK	Retinoblastoma-like protein 1	RBL1
N	4	EPSSP <u>S</u> PPISPLPLSTR	TBC1 domain family member 17	TBC1D17
N	4	AVSTVVVTTAP <u>S</u> PK	Nuclear factor related to kappa-B-binding protein	NFRKB
N/BB	3	AR <u>S</u> PVQDVLK	TRAF-interacting protein	TRAIIP
N/BB	3	A <u>S</u> DLEDEENATR	zinc finger CCCH domain-containing protein 18 isoform 2	ZC3H18
N/BB	3	H <u>S</u> PIKEEPcGSLSETVcK	Male-specific lethal 1-like	MSL1
N/BB	3	QEAGD <u>S</u> PPPAPGtPK	transcription initiation factor TFIID subunit 6-like	TAF6
N/BB	3	QDSSASLA <u>S</u> PVEIAGK	NK-tumor recognition protein	NKTR
N/BB	3	DLG <u>S</u> ELVR	Hepatocyte growth factor receptor	MET
N/BB	3	VDPSLMED <u>S</u> DDGPSLPTK	protein RER1-like	RER1
N	3	Ic <u>S</u> DEEEDEEKGGR	ubiquitin-1	UBN1
N	3	SE <u>S</u> SMPVQLNK	TBC1 domain family member 5	TBC1D5
N	3	KAAS <u>S</u> LTEDR	Eukaryotic translation initiation factor 4 gamma 1	EIF4G1
N	3	IQIISTDSAVA <u>S</u> PQR	nuclear receptor subfamily 2 group C member 2	Nr2c2
N	3	TAVSVAQGGH <u>S</u> R	Pre-B-cell leukemia transcription factor 2	PBX2
N	3	IV <u>S</u> PSGAAVScK	Filamin-A	FLNA
N	3	<u>S</u> IEEGDGAFGR	PERQ amino acid-rich with GYF domain-containing protein 1	GIGYF1
N	3	RED <u>S</u> PGPEVQPMDK	tripartite motif-containing protein 3	TRIM3
N	3	SLDRD <u>S</u> PGVENK	B-cell CLL/lymphoma 9 protein	BCL9
BB	3	AQEP <u>S</u> PVPSK	Replication factor C subunit 2	RFC2

**Table 4.2.30** List of phosphopeptides that were only identified in CHO SEAP cells at 31°C using iron oxide phosphopeptide enrichment. Database denotes which CHO proteomic database the phosphopeptide was identified in (N= CHO NCBI) (BB=CHO BB). The table also indicates the number of independent biological replicates that the phosphopeptide was identified in. For this list each phosphopeptide was identified in a minimum of three biological replicates.

Database	Number of times identified	Peptide	Protein Name	Gene symbol
N/BB	4	LE <u>S</u> LNIQR	eukaryotic translation initiation factor 3 subunit A-like	EIF3A
N	4	DA <u>T</u> PPV <u>S</u> PINMEDQER	transcription factor jun-B-like	JUNB
N/BB	3	DLDPDMV <u>T</u> EDEDDPGSQR	bromodomain adjacent to zinc finger domain protein 1A	BAZ1A
N/BB	3	IN <u>S</u> DDKNLYLTASK	ras-related protein Rab-34-like isoform 1	RAB34
N/BB	3	LVD <u>S</u> SPVSADALR	Protein Spindly	SPDL1
N/BB	3	VFIDQNL <u>S</u> PGK	transcription factor Dp-1 isoform 1	TFDP1
N/BB	3	LGML <u>S</u> PEGTcK	fatty acid synthase	FASN
N/BB	3	IGPLGL <u>S</u> PK	60S ribosomal protein L12	RPL12
N/BB	3	GL <u>S</u> EDTTEETLK	nucleolin	NCL
N	3	GGGN <u>S</u> EGELETTVDQGGGENQAR	Ras and Rab interactor 1	RIN1
N	3	APLTLAG <u>S</u> PTPK	Protein Wiz	WIZ
N	3	DNNT <u>S</u> LTALEK	hypothetical protein I79_019528	N/A



**Table 4.2.31** List of phosphopeptides that were only identified in CHO SEAP cells at 37°C using gallium oxide phosphopeptide enrichment of phosphoprotein-enriched CHO SEAP whole cell lysate samples subject to temperature-shift. Database denotes which CHO proteomic database the phosphopeptide was identified in (N= CHO NCBI) (BB=CHO BB). The table also indicates the number of independent biological replicates that the phosphopeptide was identified in. For this list each phosphopeptide was identified in a minimum of three biological replicates.

Database	Number of times identified	Peptide	Protein Name	Gene symbol
N	3	TLSNAEDYIDEEDS	osteoclast-stimulating factor 1-like	OSTF1
N	3	GGNVFEALIQDESEEEEEEEEEK	ATP-binding cassette sub-family F member 1-like	ABCF1
N	3	VDGSDVALKDPDQVAQS	Protein transport protein Sec31A	SEC31A
N	3	GD	Kinesin-like protein KIF1B	KIF1B
N	2	EAFVLV	Activating transcription factor 7-interacting protein 1	ATF7IP
N	2	SLM	Microtubule-associated protein 1B	MAP1B
BB	2	SLDDLEIGLGHLS	LIM domain and actin-binding protein 1	LIMA1
BB	2	NA	DnaJ homolog subfamily C member 2	DNAJC2
BB	2	LDGPVENV	La-related protein 1B	LARP1B
BB	2	VQGEAVSNIQENTQTPTVQEE	Nascent polypeptide-associated complex subunit alpha	NACA
BB	2	S	Ras GTPase-activating protein-binding protein 1	G3BP1
BB	2	GAEDDEG	E3 SUMO-protein ligase RanBP2	RANBP2
BB	2	TETVEEPL	Endoplasmic reticulum chaperone	HSP90B1
BB	2	EVTENEQNQL	Band 4.1-like protein 2	EPB41L2

**Table 4.2.32** List of phosphopeptides that were only identified in CHO SEAP cells at 31°C using gallium oxide phosphopeptide enrichment of phosphoprotein-enriched CHO SEAP whole cell lysate samples subject to temperature-shift. Database denotes which CHO proteomic database the phosphopeptide was identified in (N= CHO NCBI) (BB=CHO BB). The table also indicates the number of independent biological replicates that the phosphopeptide was identified in. For this list each phosphopeptide was identified in a minimum of three biological replicates.

Database	Number of times identified	Peptide	Protein Name	Gene symbol
N	3	FAALD <b>S</b> EEEDDEETTEEK	ATP-binding cassette sub-family F member 1	ABCF1
N	3	SLDSDE <b>S</b> EDEDDDYQQK	28 kDa heat- and acid-stable phosphoprotein	PDAP1
N	3	YNLDA <b>S</b> EEEDSNK	Zinc finger Ran-binding domain-containing protein 2	ZRANB2
N	2	ND <b>S</b> WGSFDLR	Nuclear fragile X mental retardation-interacting protein 2	NUFIP2
N	2	RA <b>S</b> VcAEAYNPDEEEDDAESR	cAMP-dependent protein kinase type II-beta regulatory subunit	PRKAR2B
N	2	EVEDKE <b>S</b> EGEDEDEDEDLSK	Zinc finger Ran-binding domain-containing protein 2	ZRANB2

#### **4.2.2.4 Summary**

Following the use of three different phosphopeptide enrichment strategies to profile the phosphoproteome of CHO cells, the qualitative data obtained from LC/MS analysis was collated and 1,307 phosphopeptides and 1,480 phosphosites were identified. Titanium dioxide and iron oxide enriched significantly more phosphopeptides than gallium; however this is likely to be due to a difference in the starting material used prior to phosphopeptide enrichment i.e. phosphoprotein-enriched sample compared to a whole cell lysate. Titanium dioxide and iron oxide, both used to enrich CHO whole cell lysate, were comparable in terms of the numbers of singly and multiply phosphorylated species they identified. The recent availability of two CHO databases has facilitated the use of a Mass spectrometry based approach to analyse the phosphoproteome of CHO cells. This section has shown that both databases increase proteome coverage and the confidence with which phosphopeptides were identified.

#### **4.2.3 Quantitative Phosphoproteomic label-free LC/MS analysis of CHO SEAP cells subject to temperature-shift**

Following the successful phosphopeptide enrichment of CHO SEAP cells, the LC/MS data gathered was analysed for changes in the abundance of the phosphopeptides identified as changes in the abundance of enriched phosphopeptides are potentially indicative of changes in the abundance of phosphorylation on that peptide. Quantitative information on changes in the phosphorylation status of proteins provides significant insights into the active cellular response to reduced culture temperature conditions. This section describes the use of a quantitative label-free LC/MS proteomic profiling approach to quantitatively compare CHO SEAP cells collected from 37°C and 31°C cultures 36hrs post-temperature-shift. Quantitative label-free LC/MS analysis was performed on the following cell lysate preparations:

- Whole cell lysate
- Phosphoprotein enrichment of whole cell lysate
- Gallium phosphopeptide enrichment of phosphoprotein enrichment
- Titanium dioxide phosphopeptide enrichment of whole cell lysate
- Iron oxide phosphopeptide enrichment of whole cell lysate

#### **4.2.4 Quantitative label-free LC/MS analysis of CHO SEAP whole cell lysate**

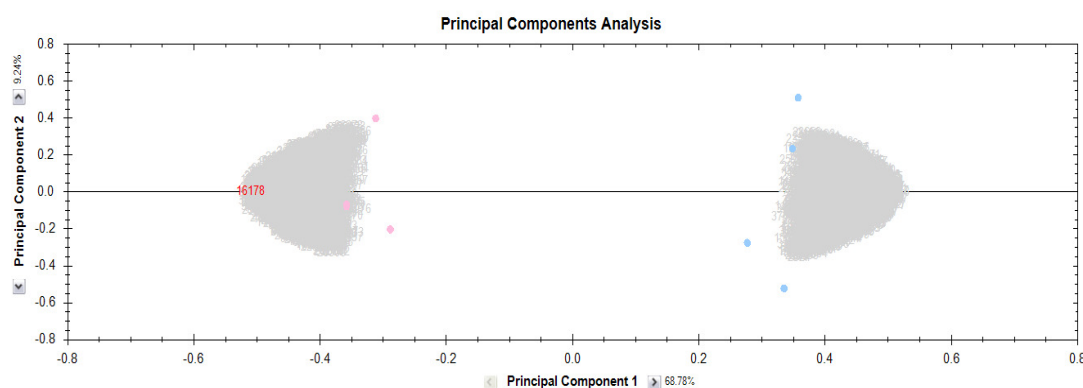
Quantitative label-free LC/MS analysis was conducted on CHO SEAP whole cell lysates for two main reasons:

- LC/MS analysis of the ‘starting material’ provided a good indication of how successful subsequent phosphoprotein / phosphopeptide enrichments were.
- To date, quantitative label-free LC/MS analysis is a technique that has not previously been applied to an investigation of the CHO proteome under reduced culture temperature conditions.

As such, eight CHO SEAP whole cell lysate samples, four from each culture condition (37°C and 31°C) were prepared for LC/MS analysis. Due to the complex nature of the whole cell lysate samples, 5hr reverse phase gradients were used to ensure sufficient peptide separation during mass spectrometry analysis.

After LC/MS analysis of all the samples was complete, inclusion lists were generated and analysis on some of the samples was repeated. This was done to obtain information on additional peptides to further strengthen protein identifications and identify low abundant proteins that were not previously identified in the initial analysis.

Progenesis label-free software was used to analyse the data acquired following LC/MS analysis. Reference run selection was performed on two separate occasions. Initially the software chose what it deemed to be the most suitable reference run and the analysis was performed to completion. A principal component analysis (PCA) plot generated by Progenesis during data analysis showing how the samples from 37°C and 31°C cultures cluster is displayed in Figure 4.2.10. For the second the reference run was carried out manually and the analysis was then carried out to completion as detailed in section 2.5.8.



**Figure 4.2.10** PCA plot for CHO SEAP whole cell lysate sample groups 37°C (pink) and 31°C (blue) generated during label-free LC/MS data analysis using Progenesis software.

The list of differentially-expressed proteins generated from the automatic reference run selection and the list of differentially-expressed proteins generated from the manual reference run selection were then overlapped and any proteins common to both lists were accepted as being differentially-expressed as described in section 2.5.8. In all instances proteins common to both lists also shared common approximate fold-change and fold-change direction. No proteins had to be excluded on the basis of differing direction of fold-change.

A protein was considered to be differentially-regulated if its change in relative abundance was >1.5-fold. Lists of differentially-regulated proteins were generated based on peptides deemed to have a statically significant fold-change (ANOVA <0.05) and an ion score >40 when searched against CHO NCBI and CHO BB proteomic databases.

Using this format for the identification of differentially-expressed proteins, 114 were deemed to have a statistically significant fold-change in relative abundance >1.5 fold. Of these, 42 were down-regulated (Table 4.2.33), while 72 were up-regulated at 31°C (Table 4.2.34).

**Table 4.2.33** List of differentially-expressed proteins, down-regulated at 31°C compared to 37°C, identified in quantitative label-free LC/MS analysis of CHO SEAP whole cell lysates from cells sampled at 36hrs post-temperature-shift.

Gene Symbol	Description	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
HSPG2	basement membrane-specific heparan sulfate proteoglycan core protein	5	5	267	0.0002	2.7
SELENBP1	selenium-binding protein 1	4	4	224	0.0001	1.9
FLOT1	flotillin-1	3	3	234	0.0047	1.7
PLIN4	perilipin-4	3	3	213	0.0017	2.4
HYOU1	Hypoxia up-regulated protein 1	3	3	158	0.0001	1.5
AARS	Alanyl-tRNA synthetase, cytoplasmic	2	2	121	0.0011	1.6
CAMK2D	Calcium/calmodulin-dependent protein kinase type II subunit delta	2	2	116	0.0049	2.0
MET	Hepatocyte growth factor receptor	2	2	109	0.0040	2.3
SERPINH1	Serpin H1	2	2	106	0.0006	2.5
RRM1	Ribonucleoside-diphosphate reductase large subunit	2	2	99	0.0002	2.3
CPNE1	Copine-1	2	2	98	0.0139	1.7
ANXA6	Annexin A6	2	2	95	0.0020	1.8
Cald1	Non-muscle caldesmon	2	1	119	0.0039	1.6
IGBP1	Immunoglobulin-binding protein 1	1	1	90	0.0088	1.8
PALLD	palladin	1	1	78	0.0003	2.1
DDX17	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17, isoform CRA_a	1	1	72	0.0050	1.6
DNMT1	DNA (cytosine-5)-methyltransferase 1	1	1	70	0.0082	1.6
HTRA1	Serine protease HTRA1	1	1	63	0.0025	2.7
PXDN	peroxidasin homolog	1	1	61	0.0001	2.1
PSAT1	Phosphoserine aminotransferase	1	1	61	0.0053	1.6
TPM4	Tropomyosin alpha-4 chain	1	1	60	0.0286	1.6

PFKL	6-phosphofructokinase, liver type	1	1	60	0.0226	1.6
PTGFRN	Prostaglandin F2 receptor negative regulator	1	1	60	0.0086	1.7
TSC22D1	TSC22 domain family protein 1	1	1	58	0.0009	3.1
PCNP	PEST proteolytic signal-containing nuclear protein	1	1	57	0.0065	1.6
AP2A2	AP-2 complex subunit alpha-2	1	1	53	0.0019	1.6
STAT5B	signal transducer and activator of transcription 5B	1	1	53	0.0010	1.9
HDGFRP2	Hepatoma-derived growth factor-related protein 2	1	1	53	0.0162	1.6
SPC25	Kinetochore protein Spc25	1	1	52	0.0064	2.4
MESDC2	LDLR chaperone MESD	1	1	51	0.0006	1.6
SNAP29	Synaptosomal-associated protein 29	1	1	49	0.0001	2.0
NDUFA7	Protein midA homolog, mitochondrial	1	1	48	0.0195	1.7
CATSPER2	Cation channel sperm-associated protein 2	1	1	46	0.0249	1.8
PPID	Peptidyl-prolyl cis-trans isomerase D	1	1	45	0.0022	1.8
PPP1R2	Protein phosphatase inhibitor 2	1	1	45	0.0006	1.9
MAP1B	Microtubule-associated protein 1B	1	1	44	0.0040	2.0
NCAPD2	Condensin complex subunit 1	1	1	44	0.0040	1.6
MYH9	Myosin-9	1	1	43	0.0027	1.6
PAPSS1	Bifunctional 3'-phosphoadenosine 5'-phosphosulfate synthase 1	1	1	43	0.0019	1.7
GPC1	Glypican-1	1	1	41	0.0023	1.6
NID1	Nidogen-1	1	1	40	0.0253	1.7
FABP4	Fatty acid-binding protein, adipocyte	1	1	40	0.0005	1.5



**Table 4.2.34** List of differentially-expressed proteins, up-regulated at 31°C compared to 37°C, identified in quantitative label-free LC/MS analysis of CHO SEAP whole cell lysates from cells sampled at 36hrs post-temperature-shift.

Gene Symbol	Description	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
PLEC	Plectin	13	13	801	0.0001	1.5
DHX9	ATP-dependent RNA helicase A	10	10	555	0.0000	2.4
MYBBP1A	Myb-binding protein 1A	9	9	556	0.0007	1.6
LMNA	Lamin-A/C	7	7	434	0.0002	1.6
MSN	Moesin	7	5	398	0.0000	1.5
EZR	Ezrin	5	3	252	0.0001	2.4
CAST	calpastatin	4	4	247	0.0000	1.7
RPL4	60S ribosomal protein L4	4	4	209	0.0002	2.1
DDX21	Nucleolar RNA helicase 2	4	4	202	0.0026	1.5
NOP56	Nucleolar protein 56	4	4	198	0.0016	1.5
UBAP2L	Ubiquitin-associated protein 2	3	3	212	0.0058	1.6
EIF3I	Eukaryotic translation initiation factor 3 subunit I	3	3	178	0.0028	1.6
RPL26	60S ribosomal protein L26	3	3	173	0.0002	1.5
GLUD1	Glutamate dehydrogenase 1, mitochondrial	3	3	153	0.0047	1.7
LRRC59	leucine-rich repeat-containing protein 59	3	3	138	0.0072	1.6
AKR1A1	Alcohol dehydrogenase (NADP+)	3	3	137	0.0009	1.5
BZW1	basic leucine zipper and W2 domain-containing protein 1	3	3	135	0.0049	1.6
PSMB8	Proteasome subunit beta type-8	2	2	145	0.0001	1.7
RRP12	RRP12 protein	2	2	139	0.0035	1.6
RPS25	40S ribosomal protein S25	2	2	133	0.0303	1.6
Rpl6	60S ribosomal protein L6	2	2	132	0.0029	3.1
CYP51A1	Lanosterol 14-alpha demethylase	2	2	129	0.0003	1.9

RUVBL2	RuvB 2	2	2	120	0.0032	1.5
VIM	Vimentin	2	2	116	0.0001	1.8
RPS9	40S ribosomal protein S9	2	2	111	0.0004	1.8
GTPBP4	Nucleolar GTP-binding protein 1	2	2	110	0.0004	1.8
HEATR1	HEAT repeat-containing protein 1	2	2	107	0.0254	1.6
DDX18	ATP-dependent RNA helicase DDX18	2	2	104	0.0164	1.5
ANXA1	Annexin A1	2	2	99	0.0149	1.8
PDCD11	protein RRP5 homolog	2	2	84	0.0008	1.6
HUWE1	E3 ubiquitin-protein ligase HUWE1	1	1	85	0.0060	1.5
RPL13A	60S ribosomal protein L13a	1	1	79	0.0003	1.7
TSTA3	GDP-L-fucose synthase	1	1	77	0.0000	2.2
EHD4	EH domain-containing protein 4	1	1	75	0.0458	1.7
ATXN10	Ataxin-10	1	1	74	0.0239	1.7
HMOX1	Heme oxygenase 1	1	1	70	0.0026	4.2
RPL28	60S ribosomal protein L28	1	1	68	0.0005	2.1
TFRC	Transferrin receptor protein 1	1	1	67	0.0393	1.6
PUS1	tRNA pseudouridine synthase A, mitochondrial isoform 1	1	1	67	0.0240	1.8
RRM2	Ribonucleoside-diphosphate reductase subunit M2	1	1	65	0.0001	2.0
BAIAP2	Brain-specific angiogenesis inhibitor 1-associated protein 2	1	1	63	0.0053	2.4
SRSF5	Serine/arginine-rich splicing factor 5	1	1	63	0.0166	1.8
HIST1H2AG	Histone H2A type 1	1	1	63	0.0001	2.2
TCEA1	Transcription elongation factor A protein 1	1	1	62	0.0104	1.7
XPO5	Exportin-5	1	1	61	0.0123	1.7
IMPDH1	Inosine-5'-monophosphate dehydrogenase 1	1	1	61	0.0173	2.1
GPRC5A	Retinoic acid-induced protein 3	1	1	60	0.0088	2.1
RPS8	40S ribosomal protein S8	1	1	58	0.0000	1.8

RPS6	40S ribosomal protein S6	1	1	58	0.0259	2.8
SRXN1	Sulfiredoxin-1	1	1	57	0.0006	2.3
CDK6	Cyclin-dependent kinase 6	1	1	55	0.0023	1.7
XRN2	5'-3' exoribonuclease 2	1	1	55	0.0079	2.0
FAT1	Protocadherin Fat 1	1	1	54	0.0006	6.4
RPL14	60S ribosomal protein L14	1	1	53	0.0000	2.5
N/A	uncharacterized protein C7orf50 homolog	1	1	53	0.0000	2.8
APRT	Adenine phosphoribosyltransferase	1	1	51	0.0375	1.6
RPL34	60S ribosomal protein L34	1	1	51	0.0080	2.6
NFKB2	Nuclear factor NF-kappa-B p100 subunit	1	1	50	0.0285	1.9
BCL10	B-cell lymphoma/leukemia 10	1	1	49	0.0019	2.3
HACL1	2-hydroxyacyl-CoA lyase 1	1	1	49	0.0010	2.0
EPS8	Epidermal growth factor receptor kinase substrate 8	1	1	49	0.0037	2.3
LRRC47	Leucine-rich repeat-containing protein 47	1	1	47	0.0303	1.6
PRMT5	Protein arginine N-methyltransferase 5	1	1	46	0.0066	1.6
LXN	Latexin	1	1	46	0.0007	3.3
PCNA	Proliferating cell nuclear antigen	1	1	46	0.0221	1.7
TAP2	Antigen peptide transporter 2	1	1	45	0.0322	1.7
CD44	CD44 antigen	1	1	42	0.0006	1.7
RPL13	60S ribosomal protein L13	1	1	42	0.0001	2.7
PABPC4	Polyadenylate-binding protein 4	1	1	41	0.0432	1.5
IMP4	U3 small nucleolar ribonucleoprotein protein IMP4	1	1	41	0.0380	1.6
EIF3B	Eukaryotic translation initiation factor 3 subunit B	1	1	41	0.0299	1.6
RSL1D1	ribosomal L1 domain-containing protein 1	1	1	41	0.0072	2.1

#### **4.2.4.1 Bioinformatic analysis of differentially-regulated proteins identified in the whole cell lysate of CHO SEAP cells subject to temperature-shift**

While both CHO databases are hugely beneficial, the proteins in the databases lack continuity in the way they are presented - they don't provide a consistent identifier e.g. gi number, gene symbol, UniProt accession number etc. This is not just between databases, but within each database itself. For this reason, each differentially-expressed protein identified was searched manually against UniProtKB (<http://www.uniprot.org/>) and the equivalent gene symbol for the human version of that protein retrieved.

The list of gene symbols were then searched using DAVID functional annotation tool (publically available software) (<http://david.abcc.ncifcrf.gov/>). Gene ontology was then used to classify the differentially-regulated proteins (both up and down at 31°C) and identify enrichment of pathways that may be involved in the cellular response to temperature-shift. Using a p-value of  $\leq 0.05$ , 25 pathways were found to be enriched (Table 4.2.35). This list contains a wide range of cellular processes; however, in order to refine the number of pathways identified, a more conservative approach was adapted and an adjusted p-value (Benjamini) of  $\leq 0.05$  was used (Table 4.2.36). Using an adjusted p-value, 9 pathways of interest were identified. As can be seen in Figure 4.2.11, DAVID analysis has identified a clear enrichment of a number of pathways including translation and RNA processing at the reduced culture temperature.

The list of gene symbols were also searched using DAVID's functional annotation chart function. Of the 114 gene symbols inputted (corresponding to the differentially-expressed proteins identified), 84 of these were identified as phosphoproteins ( $p = 7.7E-16$ ). Interestingly, based on the number of gene symbols, this was the most highly enriched entity despite the analysis being based on a whole cell lysate that has not undergone any form of enrichment prior to LC/MS.

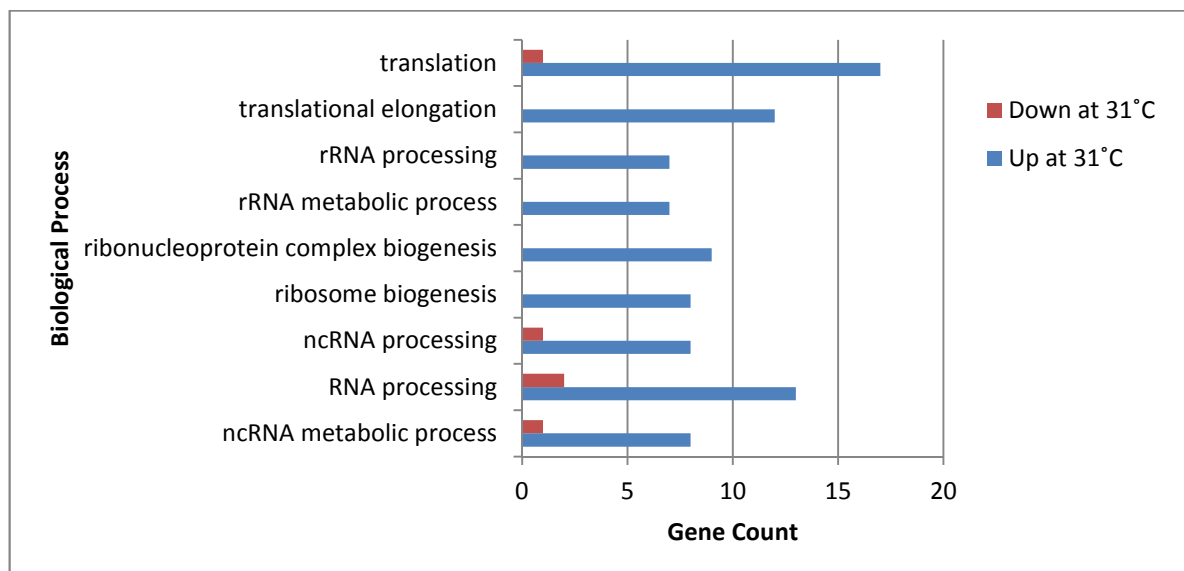
**Table 4.2.35** GO biological process (BP) enrichment for differentially-expressed proteins identified in quantitative label-free LC/MS analysis of whole cell lysate from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of a p-value  $\leq 0.05$ . Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category.

Category	Term	Count	P-Value
GOTERM_BP_FAT	translation	18	2.40E-10
GOTERM_BP_FAT	translational elongation	12	1.70E-10
GOTERM_BP_FAT	rRNA metabolic process	7	0.000072
GOTERM_BP_FAT	rRNA processing	7	0.000057
GOTERM_BP_FAT	ncRNA processing	9	0.000067
GOTERM_BP_FAT	ribosome biogenesis	8	0.000031
GOTERM_BP_FAT	ribonucleoprotein complex biogenesis	9	0.000051
GOTERM_BP_FAT	RNA processing	15	0.000042
GOTERM_BP_FAT	ncRNA metabolic process	9	0.00028
GOTERM_BP_FAT	syncytium formation by plasma membrane fusion	3	0.0033
GOTERM_BP_FAT	syncytium formation	3	0.0046
GOTERM_BP_FAT	regulation of immune effector process	5	0.0064
GOTERM_BP_FAT	response to organic substance	13	0.0063
GOTERM_BP_FAT	leukocyte adhesion	3	0.018
GOTERM_BP_FAT	response to hypoxia	5	0.017
GOTERM_BP_FAT	response to oxygen levels	5	0.020
GOTERM_BP_FAT	nitrogen compound biosynthetic process	7	0.032
GOTERM_BP_FAT	regulation of mast cell cytokine production	2	0.029
GOTERM_BP_FAT	extracellular structure organization	5	0.032
GOTERM_BP_FAT	positive regulation of immune effector process	3	0.041
GOTERM_BP_FAT	response to vitamin A	3	0.038
GOTERM_BP_FAT	membrane to membrane docking	2	0.036
GOTERM_BP_FAT	healing during inflammatory response	2	0.036
GOTERM_BP_FAT	oxidation reduction	10	0.044
GOTERM_BP_FAT	extracellular matrix organization	4	0.041

**Table 4.2.36** GO biological process (BP) enrichment for differentially-expressed proteins identified in quantitative label-free LC/MS analysis of whole cell lysate from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of an adjusted p-value  $\leq 0.05$  (Benjamini). Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category.

Category	Term	Count	P-Value	Benjamini
GOTERM_BP_FAT	translation	18	2.40E-10	1.40E-07
GOTERM_BP_FAT	translational elongation	12	1.70E-10	2.00E-07
GOTERM_BP_FAT	rRNA metabolic process	7	0.000072	0.011
GOTERM_BP_FAT	rRNA processing	7	0.000057	0.011
GOTERM_BP_FAT	ncRNA processing	9	0.000067	0.012
GOTERM_BP_FAT	ribosome biogenesis	8	0.000031	0.012
GOTERM_BP_FAT	ribonucleoprotein complex biogenesis	9	0.000051	0.012
GOTERM_BP_FAT	RNA processing	15	0.000042	0.013
GOTERM_BP_FAT	ncRNA metabolic process	9	0.00028	0.037

**Figure 4.2.11** GO biological process enrichment for differentially-expressed proteins identified in quantitative label-free LC/MS analysis of whole cell lysate from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of an adjusted p-value  $\leq 0.05$  (Benjamini). Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category. This graph shows the breakdown of the gene counts in those enriched pathways.



#### **4.2.4.2 Summary**

Quantitative label-free LC/MS analysis of CHO whole cell lysate identified the differential expression of 114 proteins, 72 of which were up-regulated at 31°C (>1.5 fold-change, ANOVA <0.05). GO analysis of the differentially-expressed proteins identified clearly showed enrichment of pathways involved in translation, RNA processing and ribosomal biogenesis. While a number of phosphoproteins were identified by LC/MS analysis of the whole cell lysate samples, the use of phosphoprotein enrichment methodologies could enable the identification of other phosphoproteins not identified in this section, thus offering further insights into the effects of temperature-shift on the phosphoproteome of CHO cells.

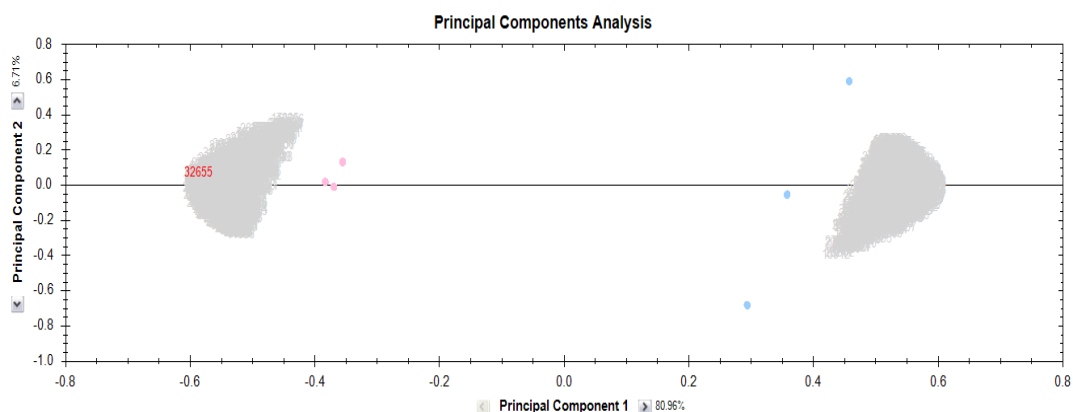
#### **4.2.5 Quantitative label-free LC/MS analysis of phosphoprotein-enriched CHO SEAP lysate**

While it is estimated that nearly a third of all proteins are phosphorylated, many of these proteins are often low in abundance and require enrichment in order to be analysed (Delom and Chevet 2006; Goshe 2006). Although a significant proportion of the differentially-expressed proteins identified in the CHO SEAP whole cell lysate sample were phosphoproteins, the use of a phosphoprotein enrichment strategy could potentially further enhance these results through the identification of an additional cohort of phosphoproteins not previously identified in section 4.2.4. As such, a phosphoprotein enrichment kit from Pierce® was used to enrich for phosphoproteins from the temperature-shift CHO SEAP whole cell lysate samples collected as described in section 4.2.1.

Although eight CHO SEAP whole cell lysate samples were available (four from each culture condition, 37°C and 31°C), six samples were prepared using a Pierce® phosphoprotein enrichment kit as detailed in materials 2.4.1. The resulting comparison was based on three samples from each culture condition. Following phosphoprotein enrichment, all samples were then prepared for label-free LC/MS as detailed materials 2.3.5. After LC/MS analysis of all the samples was complete, inclusion lists were generated and analysis on some of the samples was repeated as described in section 4.2.4.

Progenesis label-free software was used to analyse the data acquired following LC/MS analysis. Reference run selection was performed on two separate occasions. A principal component analysis (PCA) plot generated by Progenesis during data analysis showing how the phosphoprotein-enriched samples from 37°C and 31°C cultures cluster is displayed in Figure 4.2.12. As per section 4.2.4, the lists of differentially-expressed proteins were then overlapped and any proteins common to both lists were accepted as being differentially-expressed. Similarly, all proteins common to both lists also shared common approximate fold-change and fold-change direction. No proteins had to be excluded on the basis of differing direction of fold-change.





**Figure 4.2.12** PCA plot for CHO SEAP phosphoprotein-enriched whole cell lysate sample groups 37°C (pink) and 31°C (blue) generated during quantitative label-free LC/MS data analysis using Progenesis software

Proteins were considered to be differentially-regulated if its change in relative abundance was >1.5 fold. Lists of differentially-regulated proteins were generated based on peptides deemed to have a statically significant fold-change (ANOVA <0.05) and ion scores >40 when searched against CHO NCBI and CHO BB proteomic databases.

Using the DAVID gene ID tool as described in section 3.0.1.1, functional annotation was used to identify all the phosphoproteins in the list obtained from Progenesis LC/MS analysis. From an initial list of 79 differentially-expressed proteins, DAVID classified 53 of these as being phosphoproteins ( $p=5.6E-8$ ).

Because the remaining 26 proteins were not identified as phosphoproteins their presence was deemed to be as a result of nonspecific binding during the phosphoprotein enrichment step. Of the 53 differentially-expressed phosphoproteins identified, 32 were down-regulated (Table 4.2.37) and 21 were up-regulated at 31°C (Table 4.2.38)

When the list of differentially-expressed proteins from the phosphoprotein enrichment and the whole cell lysate analysis were compared it was found that 10 proteins were common to both lists. All 10 proteins were expressed in the same direction in both lists, with 7 down-regulated at 31°C compared to 37°C (Table 4.2.39) and 3 up-regulated at 31°C compared to 37°C (Table 4.2.40).

This possibly indicates that the change in expression observed in the 10 proteins listed is as a result of change in total abundance rather than a change in phosphorylation. It is also possible that a change in both phosphorylation and overall abundance occurred. Further work would be required to determine whether it was a change in abundance, phosphorylation or both that occurred in this instance. In either case, label-free analysis of phosphoprotein-enriched CHO SEAP samples indicates that there are changes in the phosphoproteome as a result of temperature-shift.

**Table 4.2.37** List of phosphoproteins (classified using DAVID as detailed 0), the expression of which decreased at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift. Phosphoproteins were enriched from CHO SEAP whole cell lysates using a Pierce phosphoprotein enrichment kit.

Gene Symbol	Description	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
MAP1B	Microtubule-associated protein 1B	4	4	270	0.01028	5.7
TLN2	Talin-2	4	4	243	0.01471	2.4
GTF2I	General transcription factor II-I	3	3	206	0.00154	3.5
ANXA6	Annexin A6	3	3	194	0.00758	4.8
CKAP5	cytoskeleton-associated protein 5-like	2	2	129	0.00059	1.7
TRIP10	Cdc42-interacting protein 4	2	2	121	0.02647	2.4
FLNB	filamin-B isoform 1	2	2	119	0.01570	1.7
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	2	2	118	0.00916	4.2
SPTBN1	Spectrin beta chain, brain 1	2	2	114	0.01251	1.8
ACTN1	Alpha-actinin-1	1	1	114	0.00610	26.8
WDR5	WD repeat-containing protein 5	1	1	71	0.03418	2.4
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	1	1	68	0.00073	31.4
CPNE1	Copine-1	1	1	59	0.04534	1.8
NAA25	N-alpha-acetyltransferase 25, NatB auxiliary subunit	1	1	56	0.01961	2.2
ST13	Hsc70-interacting protein	1	1	55	0.04640	19.5
MSH2	DNA mismatch repair protein Msh2	1	1	54	0.04275	2.7
HSP90B1	endoplasmic	1	1	52	0.03358	1.6
SMARCC2	SWI/SNF complex subunit SMARCC2	1	1	52	0.01240	7.3
ARRB1	Beta-arrestin-1	1	1	52	0.01873	1.6
TSC22D1	TSC22 domain family protein 1	1	1	50	0.02784	1.9
SF3A3	splicing factor 3A subunit 3	1	1	49	0.03411	1.6

PALLD	Palladin	1	1	49	0.02294	1.6
UGGT1	UDP-glucose:glycoprotein glucosyltransferase 1	1	1	49	0.02278	4.4
ATP6V1B2	V-type proton ATPase subunit B, brain isoform	1	1	48	0.02725	1.9
CBX5	chromobox protein homolog 5	1	1	48	0.00106	2.0
PFKL	6-phosphofructokinase, liver type	1	1	46	0.04257	2.8
HSPA5	78 kDa glucose-regulated protein	1	1	46	0.00485	1.7
FANCI	Fanconi anemia group I protein-like	1	1	45	0.03841	2.8
MVP	Major vault protein	1	1	45	0.00498	2.2
TGM2	protein-glutamine gamma-glutamyltransferase 2	1	1	44	0.04733	1.6
RRBP1	Ribosome-binding protein 1	1	1	43	0.04434	4.3
Sept6	Septin-6	1	1	41	0.01275	2.3

**Table 4.2.38** List of phosphoproteins (classified using DAVID as detailed 0), the expression of which increased at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift. Phosphoproteins were enriched from CHO SEAP whole cell lysates using a Pierce phosphoprotein enrichment kit.

Gene Symbol	Description	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
LMNA	Lamin-A/C	3	3	192	0.0005	3.7
IQGAP1	Ras GTPase-activating-like protein IQGAP1	3	3	139	0.0054	2.1
TPR	Nucleoprotein TPR	2	2	131	0.0300	1.8
MCM6	DNA replication licensing factor MCM6	2	2	97	0.0103	2.4
EZR	Ezrin	2	2	96	0.0068	2.1
EMC8	Neighbor of COX4	1	1	68	0.0170	2.1
FASN	fatty acid synthase	1	1	68	0.0416	5.7
EIF4G1	Eukaryotic translation initiation factor 4 gamma 1	1	1	59	0.0351	2.0
EIF3B	eukaryotic translation initiation factor 3 subunit B	1	1	58	0.0253	2.7
CCT8	T-complex protein 1 subunit theta	1	1	58	0.0433	1.7
BZW2	Basic leucine zipper and W2 domain-containing protein 2	1	1	57	0.0312	1.5
PTGES3	Prostaglandin E synthase 3	1	1	56	0.0487	3.2
EEF1A1	elongation factor 1-alpha 1	1	1	53	0.0284	1.6
KHSRP	Far upstream element-binding protein 2	1	1	52	0.0265	17.8
MYH9	Myosin-9	1	1	52	0.0395	2.5
RPLP0	60S acidic ribosomal protein P0	1	1	48	0.0415	1.9
TPM3	Tropomyosin alpha-3 chain	1	1	47	0.0112	2.7
TK1	thymidine kinase, cytosolic	1	1	45	0.0379	7.5
PABPC1	Polyadenylate-binding protein 1	1	1	45	0.0303	3.9
EIF4A3	Eukaryotic initiation factor 4A-III	1	1	44	0.0034	1.8
AHNAK	Neuroblast differentiation-associated protein AHNAK	1	1	44	0.0448	1.5

**Table 4.2.39** List of proteins common to two lists of differentially-expressed proteins, down-regulated at 31°C compared to 37 °C, identified in quantitative label-free LC/MS analysis of CHO SEAP whole cell lysate and phosphoprotein-enriched CHO SEAP whole cell lysate.

<b>Description</b>	<b>Sample</b>	<b>Total Peptides</b>	<b>Quantitative Peptides</b>	<b>Score</b>	<b>Anova (p)</b>	<b>Fold</b>
Microtubule-associated protein 1B	Whole Cell Lysate	2	2	100	0.0017	2.1
	Enrichment	4	4	273	0.0047	5.6
Annexin A6	Whole Cell Lysate	2	2	96	0.0055	1.7
	Enrichment	3	3	178	0.0018	3.8
Copine-1	Whole Cell Lysate	2	2	98	0.0139	1.7
	Enrichment	1	1	59	0.0453	1.8
6-phosphofructokinase, liver type	Whole Cell Lysate	1	1	60	0.0226	1.6
	Enrichment	1	1	46	0.0426	2.8
Myosin-9	Whole Cell Lysate	1	1	43	0.0027	1.6
	Enrichment	1	1	52	0.0395	2.5
Palladin	Whole Cell Lysate	1	1	78	0.0003	2.1
	Enrichment	1	1	49	0.0229	1.6
TSC22 domain family protein 1	Whole Cell Lysate	1	1	58	0.0009	3.1
	Enrichment	1	1	50	0.0278	1.9

**Table 4.2.40** List of proteins common to two lists of differentially-expressed proteins, up-regulated at 31°C, identified in quantitative label-free LC/MS analysis of CHO SEAP whole cell lysate and phosphoprotein-enriched CHO SEAP whole cell lysate.

Description	Sample	Total Peptides	Quantitative Peptides	Score	Anova (p)	Fold
Lamin-A/C	Whole Cell Lysate	6	6	385	0.0002	1.7
	Enrichment	2	2	141	0.0061	4.2
Ezrin	Whole Cell Lysate	5	3	252	0.0001	2.4
	Enrichment	2	2	96	0.0068	2.1
Eukaryotic translation initiation factor 3 subunit B	Whole Cell Lysate	1	1	41	0.0299	1.6
	Enrichment	1	1	58	0.0253	2.7

#### **4.2.5.1 Bioinformatic analysis of differentially-regulated phosphoproteins**

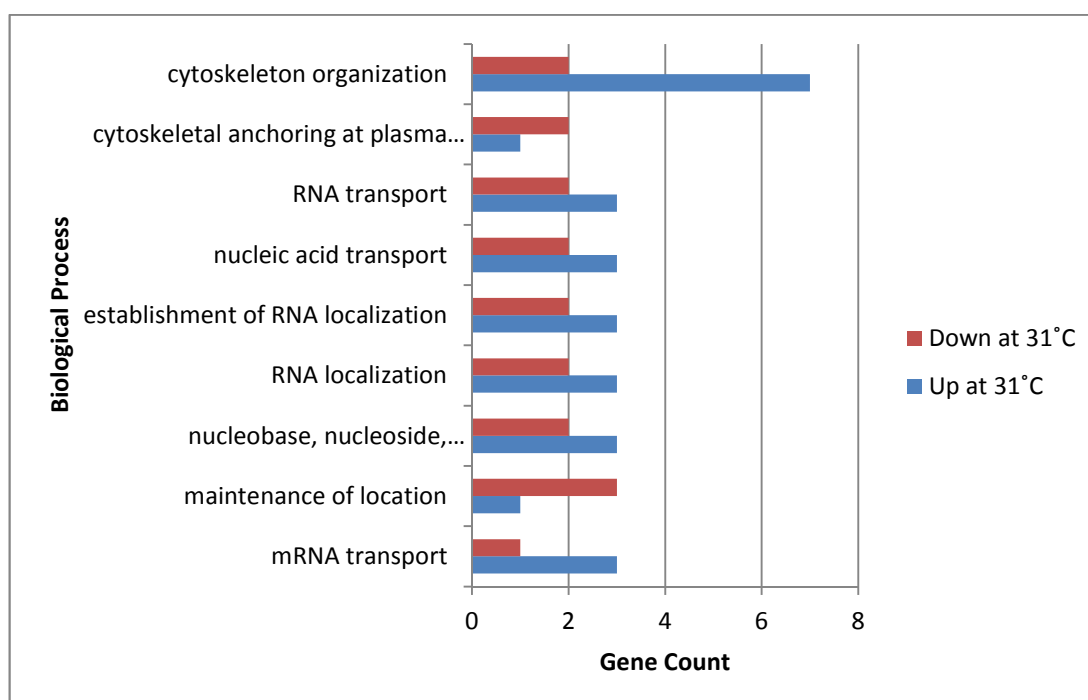
As per section 3.0.1.1, DAVID functional annotation tool was used to analyse the label-free data in order to identify enrichment of different processes that may be involved in the response to reduced culture temperature. Using a  $\leq 0.05$  p-value cut-off 9 pathways were found to be enriched upon temperature-shift (Table 4.2.41). When the number of differentially-expressed proteins associated with each process is graphed (Figure 4.2.13) it is evident that the pathways identified are not as well defined as the results obtained for those from the analysis of the whole cell lysate (section 4.2.4.1). Although it is normally the case that all differentially-expressed genes/proteins would be analysed together (to account for the fact that up or down-regulation may have an inhibitory or stimulating effect on the pathway i.e. down-regulation of a protein may not necessarily correlate with inhibition of a process), it must be remembered that DAVID is a piece of software that merely acts as a guide and that any results obtained must be interpreted as such. As a result, in an effort to identify more a definitive list of pathways, it was decided that the up-regulated proteins from each condition would be searched on DAVID independently of each other. This approach provided a somewhat clearer insight as to the processes the 53 differentially-expressed phosphoproteins might be involved in (Table 4.2.42 and Table 4.2.43). Analysis of the proteins up-regulated at 31°C compared to 37°C indicate that pathways involved in translation and RNA processing may be affected by temperature-shift while processes including cytoskeleton organization and the generation of precursor metabolites and energy are among those predominantly down-regulated at 31°C compared to 37°C (Table 4.2.42 and Table 4.2.43).



**Table 4.2.41** GO biological process enrichment for differentially-expressed phosphoproteins identified in quantitative label-free LC/MS analysis of phosphoprotein-enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of a p-value  $\leq 0.05$ . Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category.

Term	Count	P-Value
cytoskeleton organization	9	0.0001
cytoskeletal anchoring at plasma membrane	3	0.0003
establishment of RNA localization	5	0.0003
nucleic acid transport	5	0.0003
RNA transport	5	0.0003
RNA localization	5	0.0004
nucleobase, nucleoside, nucleotide and nucleic acid transport	5	0.0006
maintenance of location	4	0.0013
mRNA transport	4	0.0032

**Figure 4.2.13** GO biological process enrichment for differentially-expressed phosphoproteins identified in quantitative label-free LC/MS analysis of phosphoprotein-enriched samples from CHO SEAP cells sampled at 36hrs post temperature-shift. Enrichment was considered significant upon observation of an adjusted p-value  $\leq 0.05$  (Benjamini). Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category. This graph shows the breakdown of the gene counts in those enriched pathways.



**Table 4.2.42** GO biological process enrichment for phosphoproteins identified as up-regulated at 31°C compared to 37°C in quantitative label-free LC/MS analysis of phosphoprotein-enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of a p-value  $\leq 0.05$ . Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category.

Term	Count	P-Value
translation	5	0.00065
posttranscriptional regulation of gene expression	4	0.0022
mRNA transport	3	0.0052
nucleic acid transport	3	0.0065
RNA transport	3	0.0065
establishment of RNA localization	3	0.0065
RNA localization	3	0.0068
nucleobase, nucleoside, nucleotide and nucleic acid transport	3	0.0087
regulation of translation	3	0.013
RNA splicing	3	0.049

**Table 4.2.43** GO biological process enrichment for phosphoproteins identified as down-regulated at 31°C compared to 37°C in quantitative label-free LC/MS analysis of phosphoprotein-enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of a p-value  $\leq 0.05$ . Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category.

Term	Count	P-Value
cytoskeleton organization	7	0.00022
maintenance of location	3	0.0077
cytoskeletal anchoring at plasma membrane	2	0.016
generation of precursor metabolites and energy	4	0.026
protein localization	6	0.033
negative regulation of apoptosis	4	0.036
negative regulation of programmed cell death	4	0.037
negative regulation of cell death	4	0.037

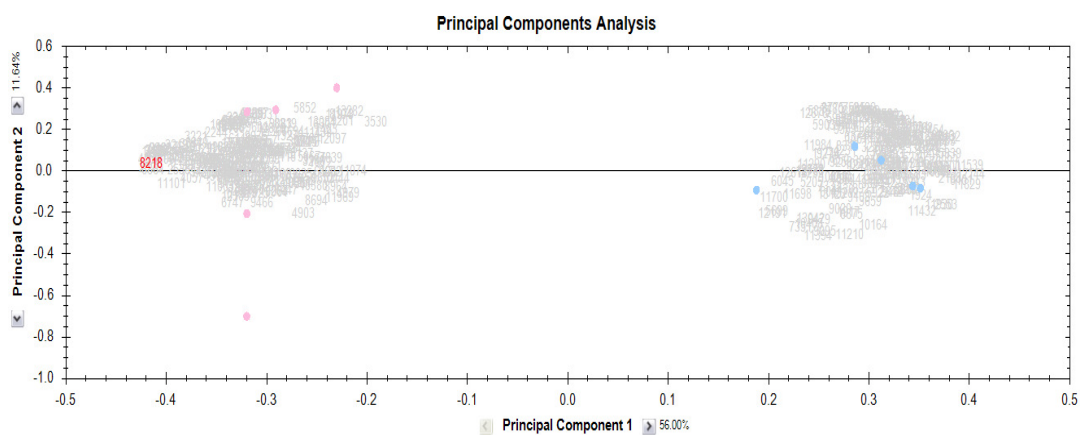
#### **4.2.5.2 Summary**

Although a large number of the differentially-expressed proteins identified in the label-free analysis of the whole cell lysate were phosphoproteins, phosphoprotein enrichment of these samples helped to further increase proteome coverage, identifying 36 differentially-expressed phosphoproteins which were not seen in the whole cell lysate study (53 differentially-expressed phosphoproteins were identified in total, >1.5 fold-change, ANOVA <0.05). The identification of these additional proteins revealed enrichment of cytoskeleton organisation pathways by GO analysis. To enable identification of site-specific phosphorylation however, further enrichment would be required.

#### 4.2.6 Quantitative label-free LC/MS analysis of a Gallium oxide phosphopeptide enrichment of a phosphoprotein-enriched CHO SEAP lysate

Upon completion of phosphoprotein-enrichment of whole cell lysate CHO SEAP samples, 100µg of each sample underwent a phosphopeptide enrichment using Gallium Oxide spin columns. It was hoped that by coupling this second methodology with phosphoprotein enrichment it would greatly improve the number of phosphopeptides identified during LC/MS analysis and allow changes in site-specific phosphorylation to be detected.

Progenesis label-free software was used to analyse the data acquired following LC/MS analysis. A principal component analysis (PCA) plot generated by Progenesis during data analysis showing how the gallium phosphopeptide-enriched samples from 37°C and 31°C cultures cluster and is displayed in Figure 4.2.14. A phosphopeptide was considered to be differentially-regulated if its change in relative abundance was >1.5 fold. Lists of differentially-regulated phosphopeptides were generated based on peptides deemed to have a statically significant fold-change (ANOVA <0.05) and peptide ion scores >40 when searched against CHO NCBI and CHO BB proteomic databases.



**Figure 4.2.14** PCA plot for CHO SEAP gallium phosphopeptide enriched whole cell lysate, phosphoprotein pre-enriched, sample groups 37°C (pink) and 31°C (blue) generated during quantitative label-free LC/MS data analysis using Progenesis software. It should be noted that the five spots from each condition represent the enriched samples that were originally analysed and two subsequent inclusion list runs for each condition used to obtain additional peptide information.

Quantitative label-free analysis of a gallium oxide phosphopeptide enrichment of phosphoprotein-enriched CHO SEAP samples subject to temperature-shift identified 8 phosphopeptides as being differentially-expressed. Of the 8 phosphopeptides, 1 showed decreased expression at 31°C (Table 4.2.44) while the remaining 7 were up-regulated at 31°C compared to 37°C (Table 4.2.45). All differentially-expressed phosphopeptides identified came from unique proteins.

Due to the low number of phosphopeptides identified bioinformatic analysis was not carried out as the sample set is too small to obtain meaningful statistically significant information. Instead each protein was manually searched against UniProtKB and the biological process as described for the human version of the protein was used to infer biological function on the CHO version identified. In addition to this, each protein was also searched against the literature to confirm the biological function involved.

As a result of this search, the proteins COP9 signalosome complex subunit 1 and PITSLRE serine/threonine-protein kinase CDC2L1, also known as Cyclin-dependent kinase 11B, were identified as being involved in cell cycle progression and translation initiation factor eIF-2B subunit epsilon-like is known to have roles in protein synthesis.

In order to determine whether the change in phosphopeptide expression was as a result of an increase in phosphorylated form of the peptide or an increase in total protein abundance, the protein names of the differentially-expressed phosphopeptides identified were searched against the list of differentially-expressed proteins from the whole cell lysate and phosphoprotein enrichment analysis. As a result of this comparison the only protein identified as common to the gallium enrichment and whole cell lysate / phosphoprotein enrichment list was DNA replication licensing factor MCM2 (as denoted with ^ in Table 4.2.45). A comparison between the differentially-expressed protein names and qualitative lists of proteins from whole cell lysate and phosphoprotein enrichment analysis were also compared. Interestingly, the proteins Nucleolar phosphoprotein p130, COP9 signalosome complex subunit 1 and 182 kDa tankyrase-1-binding protein were identified in the qualitative lists. The presence of these proteins in the qualitative list and not in the quantitative proteins lists indicates that the total abundance of these

proteins does not change but the phosphorylated form of the peptide did alter in abundance.

Although the remaining phosphopeptides were identified as changing in abundance, without further analysis it is unclear whether this is as a result of an increase in total protein, phosphorylated peptide or both. It should be noted that further proteomic analysis including further fractionation or enrichment would be required to determine the total abundance of the remaining proteins, especially if the endogenous protein was found to be expressed at low levels within the cell.

**Table 4.2.44** List of differentially-expressed phosphopeptides, down-regulated at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift. Phosphopeptides were enriched from phosphoprotein-enriched CHO SEAP whole cell lysate samples using gallium oxide phosphopeptide enrichment.

Gene Symbol	Sequence	Description	Score	Anova	Fold	Function
CDK11B	DLLSDLQDI <b>S</b> DSER	PITSLRE serine/threonine-protein kinase CDC2L1	60	0.0225	7.4	Cell cycle

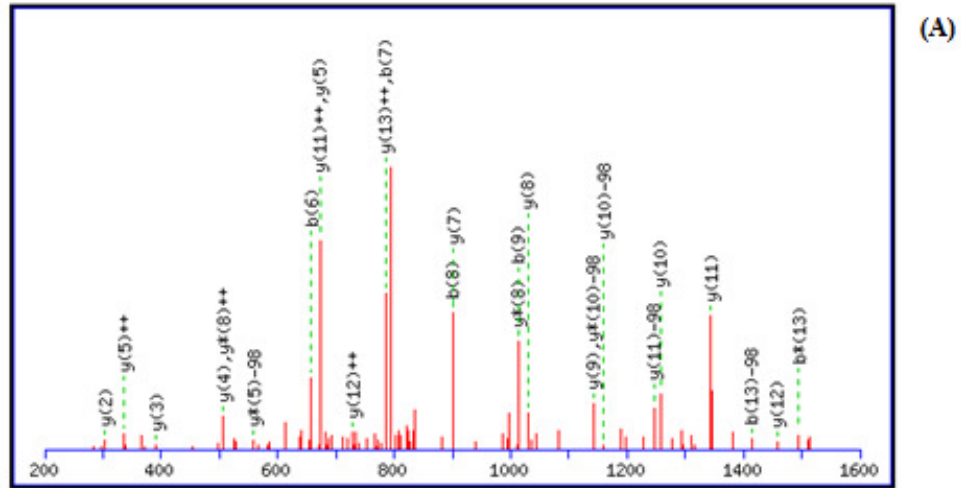
**Table 4.2.45** List of differentially-expressed phosphopeptides, up-regulated at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift. Phosphopeptides were enriched from phosphoprotein-enriched CHO SEAP whole cell lysate samples using gallium oxide phosphopeptide enrichment. (Note: † indicates proteins that don't change in total abundance. ^ denotes proteins that do change in total abundance.)

Gene Symbol	Sequence	Description	Score	Anova	Fold	Function
NOLC1	NETAEEAG <b>T</b> PQAK†	Nucleolar phosphoprotein p130	55	0.0227	2.2	Mitosis
EIF2B5	AV <b>S</b> PQLDDIK	Translation initiation factor eIF-2B subunit epsilon-like	49	0.0429	1.9	Translation
GPS1	EG <b>S</b> QGELTPANSQSR†	COP9 signalosome complex subunit 1	47	0.0069	2.0	Cell cycle
MCM2	ISDPLT <b>S</b> SPGR^	DNA replication licensing factor MCM2	44	0.0188	2.0	Cell cycle
CHAMP1	K <b>S</b> SPASLDFSESQK	Zinc finger protein 828	42	0.0314	2.0	Structural
TNKS1BP1	<b>S</b> AEEGEVTESK†	182 kDa tankyrase-1-binding protein	42	0.0426	2.6	Miscellaneous
SAFB2	APTAS <b>L</b> SPEPQDSK	Scaffold attachment factor B2	41	0.0083	2.5	Transcription

Figure 4.2.15 shows an example of MS/MS fragmentation spectrum output from a MASCOT search resulting in the identification of the phosphopeptide “DLLSDLQDISDSER” belonging to the protein PITSLRE serine/threonine-protein kinase CDC2L1. The file searched on MASCOT contained only those MS features deemed by Progenesis label-free LC/MS software as having a statistically significant differential expression (ANOVA value <0.05). Figure 4.2.16 is an example of the same phosphopeptide, “DLLSDLQDISDSER”, identified following a qualitative search of all LC/MS raw files from mass spectral analysis of gallium oxide Phosphopeptide-enriched CHO SEAP lysate samples, pre-enriched for phosphoproteins, using SEQUEST search algorithm (detailed section 2.4.6). The use of phosphorylation site localisation algorithm PhosphoRS during qualitative analysis of the mass spectral data also identified phosphorylation of the serine, tenth amino acid residue, as being phosphorylated with a >99% level of confidence. When enriched samples from 37°C and 31°C cultures were analysed using Progenesis software, the phosphopeptide “DLLSDLQDISDSER” was identified as being differentially-expressed. The large error bars in both 37°C and 31°C sample clusters are as a result of an outlier in each sample group. Despite this, Progenesis identified “DLLSDLQDISDSER” as being 7.4-fold up-regulated at 37°C (ANOVA 0.0225) (Figure 4.2.17).

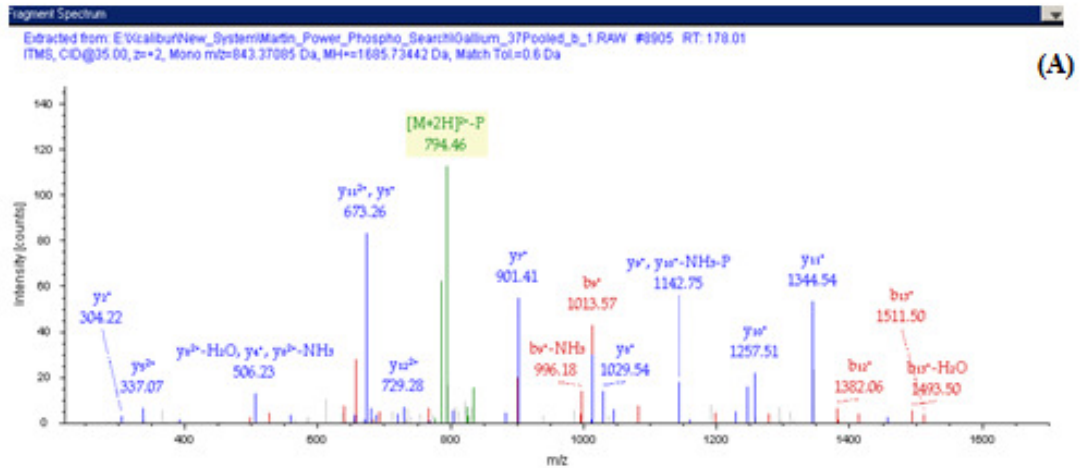
The overall number of differentially-expressed phosphopeptides identified was lower than expected. A qualitative review of the LC/MS data revealed 105 phosphopeptides were identified using Gallium oxide phosphopeptide enrichment. The number of phosphopeptides identified could potentially be due to the reduced protein population present in the sample as a result of phosphoprotein enrichment. This cohort was further reduced when the criteria for differential expression was introduced.





#	b	Seq.	y	#
1	116.0342	D		14
2	229.1183	L	1570.6996	13
3	342.2023	L	1457.6156	12
4	429.2344	S	1344.5315	11
5	544.2613	D	1257.4995	10
6	657.3454	L	1142.4725	9
7	785.404	Q	1029.3885	8
8	900.4309	D	901.3299	7
9	1013.515	I	786.3029	6
10	1180.513	S- Phospho	673.2189	5
11	1295.54	D	506.2205	4
12	1382.572	S	391.1936	3
13	1511.615	E	304.1615	2
14		R	175.119	1

**Figure 4.2.15** (A) MS/MS fragmentation spectrum output from a MASCOT search of MS features deemed by Progenesis label-free LC/MS software have a statistically significant differential expression (ANOVA value  $\leq 0.05$ ) resulting in the identification of the phosphopeptide “DLLSDLQDISDSER” belonging to the protein PITSLRE serine/threonine-protein kinase CDC2L1. (B) The b-and y-ion series used for phosphopeptide identification via MASCOT are indicated.

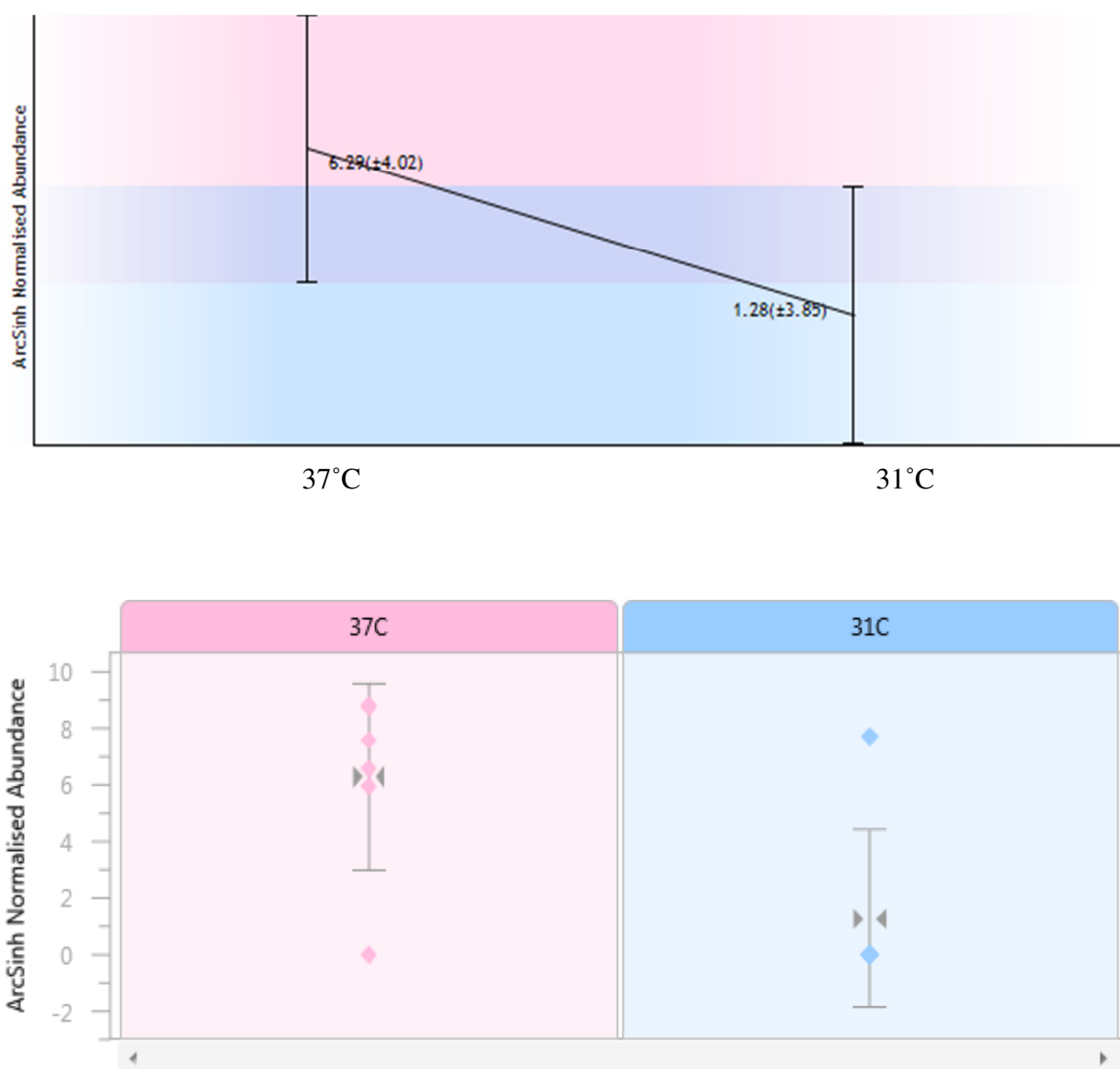


(A)

#1	b <sup>+</sup>	Seq.	y <sup>+</sup>	#2
1	116.03423	D		14
2	229.1183	L	1570.6997	13
3	342.20237	L	1457.6156	12
4	429.2344	S	1344.5316	11
5	544.26135	D	1257.4995	10
6	657.34542	L	1142.4726	9
7	785.404	Q	1029.3885	8
8	900.43095	D	901.32992	7
9	1013.515	I	786.30297	6
10	1180.5134	S-Phospho	673.2189	5
11	1295.5403	D	506.22054	4
12	1382.5724	S	391.19359	3
13	1511.615	E	304.16156	2
14		R	175.11896	1

(B)

**Figure 4.2.16** (A) MS/MS fragmentation spectrum output from a SEQUEST search of raw files from the LC/MS analysis of gallium oxide phosphopeptide enriched CHO SEAP whole cell lysate subject to temperature-shift, pre-enriched for phosphoproteins, resulting in the identification of the phosphopeptide “DLLSDLQDISDSER” belonging to the protein PITSLRE serine/threonine-protein kinase CDC2L1. (B) The b- and y-ion series used for phosphopeptide identification via SEQUEST are indicated.



**Figure 4.2.17** Progenesis label-free LC/MS output showing normalised abundance of the phosphopeptide “DLLSDLQDISDSEK” indicating its 7.39-fold down-regulation at 31°C compared to 37°C (ANOVA 0.0225).

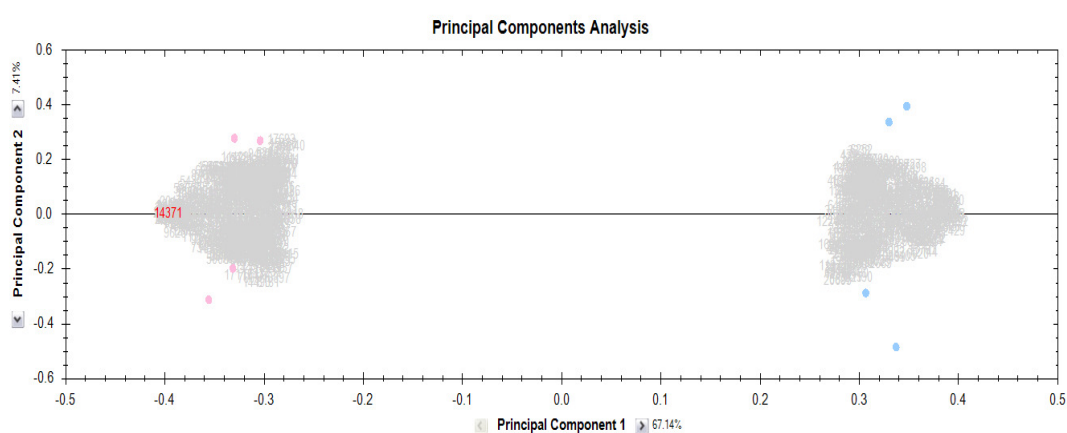
#### **4.2.6.1 Summary**

While Gallium phosphopeptide enrichment of phosphoprotein-enriched whole cell lysate samples facilitated differential expression and site-specific phosphorylation information to be gathered, the total number was lower than expected. However, this is clearly as a result of the number of phosphoproteins identified in the phosphoprotein enrichment sample. Nonetheless, this information was useful in forming an overall picture of changes taking place in the phosphoproteome of temperature-shifted CHO cells. Phosphopeptide enrichment of a whole cell lysate could yield greater numbers of phosphopeptides given that the number of proteins in the sample prior to phosphopeptide enrichment would be significantly higher.

#### 4.2.7 Quantitative label-free LC/MS analysis of titanium dioxide phosphopeptide enriched CHO SEAP whole cell lysate

Due to the low number of differentially-expressed phosphopeptides identified using phosphopeptide enrichment of a sample pre-enriched for phosphopeptides (section 4.2.6), it was decided that a phosphopeptide enrichment of a whole cell lysate may yield better results. As such, 500µg aliquots of the eight CHO SEAP whole cell lysate samples collected from shake flask cultures sampled at 36hrs post-temperature-shift were enzymatically digested and phosphopeptide enriched using titanium dioxide spin columns as detailed in materials 2.4.4. Phosphopeptide enriched samples were then prepared for label-free LC/MS as detailed in section 2.4.6. After mass spectral analysis of all the samples was complete, inclusion lists were generated and analysis on some of the samples was repeated as described in section 4.2.4.

Progenesis label-free software was used to analyse the data acquired following LC/MS analysis. Reference run selection was performed on two separate occasions and phosphopeptides common between the two lists generated were accepted as being differentially-regulated as outlined in materials and methods 2.5.8. A principal component analysis (PCA) plot generated by Progenesis during data analysis showing how the titanium phosphopeptide enriched samples from 37°C and 31°C cultures cluster and is displayed in Figure 4.2.18.



**Figure 4.2.18** PCA plot for CHO SEAP titanium dioxide Phosphopeptide enriched whole cell lysate sample groups 37°C (pink) and 31°C (blue) generated during quantitative label-free LC/MS data analysis using Progenesis software.

The list of differentially-expressed phosphopeptides was then compared to the lists of differentially-expressed proteins from the two separate CHO SEAP whole cell lysate and phosphoprotein-enriched analyses (section 4.2.4 and 4.2.5) and any phosphopeptides identified as belonging to a protein that changed in expression in the whole cell lysate or the phosphoprotein enrichment analysis were noted (indicated by ^ in the tables). Although the phosphopeptides identified changed in relative abundance, changes in the abundance of the total protein may also produce the same effect. While it is possible that an increase in the level of the total protein and the phosphorylated form of the protein may have occurred, these peptides were highlighted to indicate that further work may be required to identify the precise cause of the change in abundance.

Conversely, proteins identified in qualitative lists from whole cell lysate and phosphoprotein enrichment are indicated by † to denote that while the expression of the phosphopeptide does alter, the total protein does not change in abundance as outlined in section 4.2.6.

Phosphopeptides were considered to be differentially-regulated if their change in relative abundance was >1.5 fold. Lists of differentially-regulated proteins were generated based on peptides deemed to have a statically significant fold-change (ANOVA <0.05) and peptide ion scores >40 when searched against CHO NCBI and CHO BB proteomic databases. Based on the phosphopeptides found to be common from the two independent analyses on Progenesis label-free software, this list identified 33 unique phosphopeptides as being differentially-expressed, 18 down-regulated (Table 4.2.46) and 15 up-regulated at 31°C (Table 4.2.47).

**Table 4.2.46** List of differentially-expressed phosphopeptides, down-regulated at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift. Phosphopeptides were enriched from CHO SEAP whole cell lysates using titanium dioxide phosphopeptide enrichment. (Note: \* denotes phosphopeptides with MS<sup>3</sup> information. † indicates proteins that don't change in total abundance. . ^ indicates proteins that do change in total abundance.)

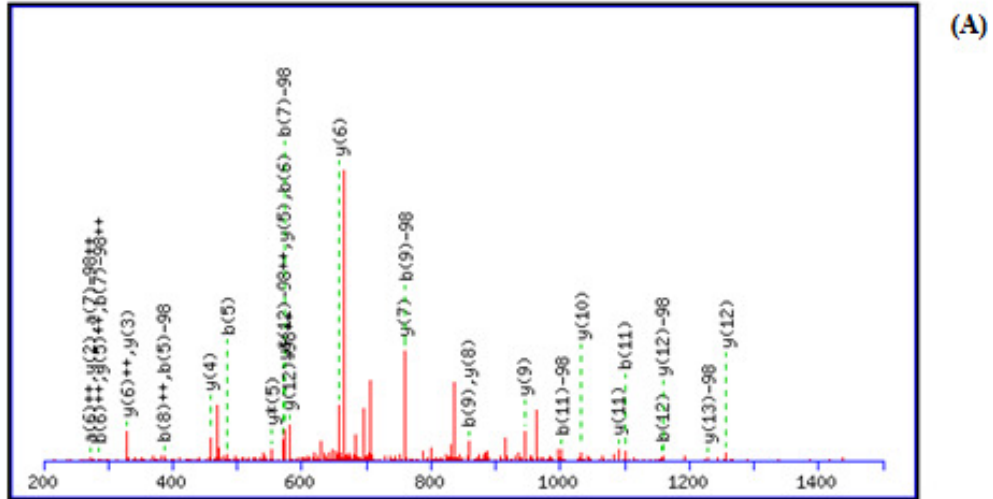
Gene Symbol	Sequence	Description	Score	Anova	Fold	Function
OSBPL11	SFSLASSGNPISQR*	Oxysterol-binding protein-related protein 11	73	0.0041	2.0	Lipid transport
N/A	SSSLDALGPSR*	Uncharacterized protein C1orf198 homolog	70	0.0000	6.2	Unknown
ANAPC1	SPSISNMAALSR*	Anaphase-promoting complex subunit 1	69	0.0258	1.5	Cell division
ARHGEF5	LDSLAEAPGLSSPR	Rho guanine nucleotide exchange factor 5	67	0.0016	2.2	Positive regulation of Rho GTPase activity
GAPVD1	SSDIVSSVR*	GTPase-activating protein and VPS9 domain-containing protein 1	65	0.0255	1.8	Positive regulation of GTPase activity
TLE3	ESSANNSVSPSESLR*	Transducin-like enhancer protein 3-like isoform 1	61	0.0398	1.5	Transcription
RAPH1	TASAGTVSDAEVR*	Ras-associated and pleckstrin homology domains-containing protein 1	59	0.0052	1.6	Regulates membrane signalling
LSM12	TETPPPLASLNVSK†	Protein LSM12 homolog	55	0.0005	2.9	Unknown
SRRM2	SPGMLEPLGSAR*^	Serine/arginine repetitive matrix protein 2	54	0.0027	2.4	mRNA processing
ELMOD3	SSELTEAVETK	ELMO domain-containing protein 3	53	0.0003	2.5	Miscellaneous
FRS2	LITSTSTSDTQININNSAQR	Fibroblast growth factor receptor substrate 2	53	0.0019	8.0	Multiple
DENND4C	VPSSGLFDTNNR*	DENN domain-containing protein 4C	51	0.0026	3.4	Protein transport
SH3PXD2A	NESSLTATDSLRL*	SH3 and PX domain-containing protein 2A	50	0.0169	2.0	Miscellaneous
EIF4G3	SPVAATVVQR*	Eukaryotic translation initiation factor 4 gamma 3	47	0.0060	3.7	Translation
ZYX	SPGGPGPLTLK*†	Zyxin	42	0.0077	1.5	Cell Adhesion
MPDZ	GSLPQVSSPR	multiple PDZ domain protein-like isoform 1	42	0.0167	2.0	Miscellaneous
KIF14	ESSLLGSELGDTTTK	Kinesin-like protein KIF14	42	0.0190	2.5	Cytokinesis
NUP88	EDAEVAE SPLR	Nuclear pore complex protein Nup88	40	0.0070	2.0	mRNA transport

**Table 4.2.47** List of differentially-expressed phosphopeptides, up-regulated at 31°C compared to 37°C, obtained from label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift. Phosphopeptides were enriched from CHO SEAP whole cell lysates using titanium dioxide phosphopeptide enrichment. (Note: \* denotes phosphopeptides with MS<sup>3</sup> information. † indicates proteins that don't change in total abundance. . ^ indicates proteins that do change in total abundance.)

Gene Symbol	Sequence	Description	Score	Anova	Fold	Function
NDRG1	TASGSSVTSLEGPR*	Protein NDRG1	88	4.40E-07	16.5	Miscellaneous
RIN2	QASFLEAEGSAK*	Ras and Rab interactor 2	79	0.0006	2.4	Positive regulation of GTPase activity
EIF5B	VEIFSGSEDDDDSNK*	Eukaryotic translation initiation factor 5B-like	71	0.0333	2.4	Translation
TRIM28	SGEGEVSGLMR*	Transcription intermediary factor 1-beta	69	0.0447	1.6	Transcription
AGPAT9	NSASVGIIQR*	Glycerol-3-phosphate acyltransferase 3	61	0.0105	1.7	Lipid metabolism
VAPB	SLTSSLDDAEVKK*	Vesicle-associated membrane protein-associated protein B	56	0.0048	3.0	Unfolded protein response
Tmpo	QDFESLPPR^	Lamina-associated polypeptide 2, isoforms alpha/zeta	55	0.0013	2.0	Transcription
TSC22D1	VDVESGSSAAGTPPLSR	TSC22 domain family protein 4	54	0.0074	1.5	Transcription
MYBBP1A	SPSLLQSVVK*^	Myb-binding protein 1A	52	0.0005	2.3	Transcription
NDRG1	SRTASGSSVTSLEGPR*	Protein NDRG1	49	0.0001	39.7	Miscellaneous
BAIAP2L1	SISTVDLTEK	Brain-specific angiogenesis inhibitor 1-associated protein 2-like protein 1	46	0.0024	2.0	Regulation of actin cytoskeleton reorganization
SMARCA4	EVDYSDSLTEK	Transcription activator BRG1	44	0.0001	2.8	Transcription
CDC42EP1	LNFDSSPTISSTDGR	Cdc42 effector protein 1	44	0.002	1.7	Multiple
PDCD4	SGVPVPTSPK	Programmed cell death protein 4	42	0.0004	2.1	Apoptosis
HUWE1	HADHSSLTLGSGSSTTR^	E3 ubiquitin-protein ligase HUWE1	41	0.0042	2.7	DNA repair

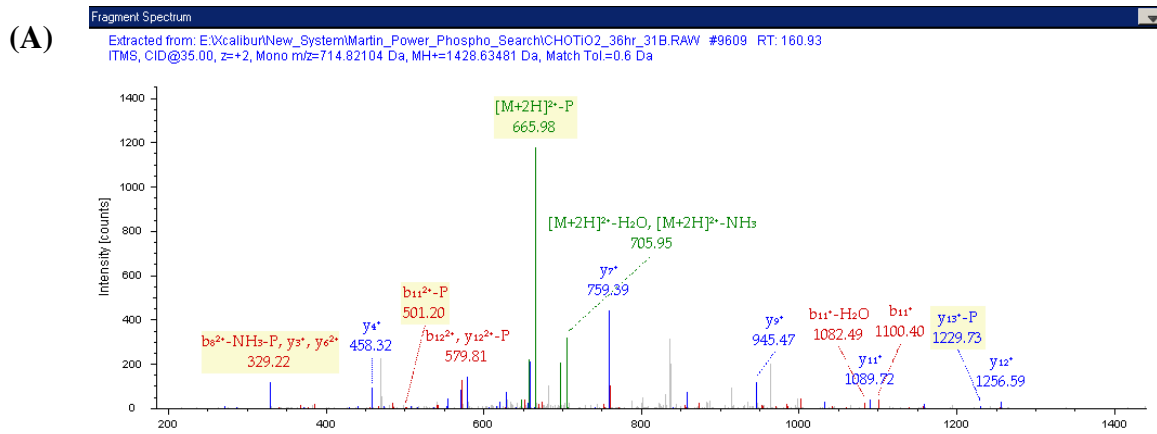


An example of an MS/MS fragmentation spectrum output from a MASCOT search that resulted in the identification of the phosphopeptide “TASGSSVTSLEGPR” from Protein NDRG1 is shown in (Figure 4.2.19). The file used in this MASCOT search contained only MS/MS information from those features deemed as having a statistically significant differential expression (ANOVA value  $\leq 0.05$ ) by Progenesis label-free LC/MS software. Following a qualitative search of all LC/MS raw files from mass spectral analysis of titanium dioxide phosphopeptide enriched CHO SEAP lysates using the SEQUEST search algorithm (detailed section 2.4.6), the same phosphopeptide, “TASGSSVTSLEGPR”, was also identified as evidenced by the MS/MS output in (Figure 4.2.20 A). Phosphorylation of the third amino acid residue, serine, was identified as being phosphorylated with a >99% level of confidence by the phosphorylation site localisation algorithm PhosphoRS. MS/MS/MS information was also obtained on this peptide during mass spectral analysis as can be seen in (Figure 4.2.20 B) providing further proof of the phosphorylation of this phosphopeptide. Figure 4.2.21 shows an output from Progenesis label-free analysis identifying “TASGSSVTSLEGPR” as being 16.5-fold up-regulated at 31°C (ANOVA 4.40E-07).



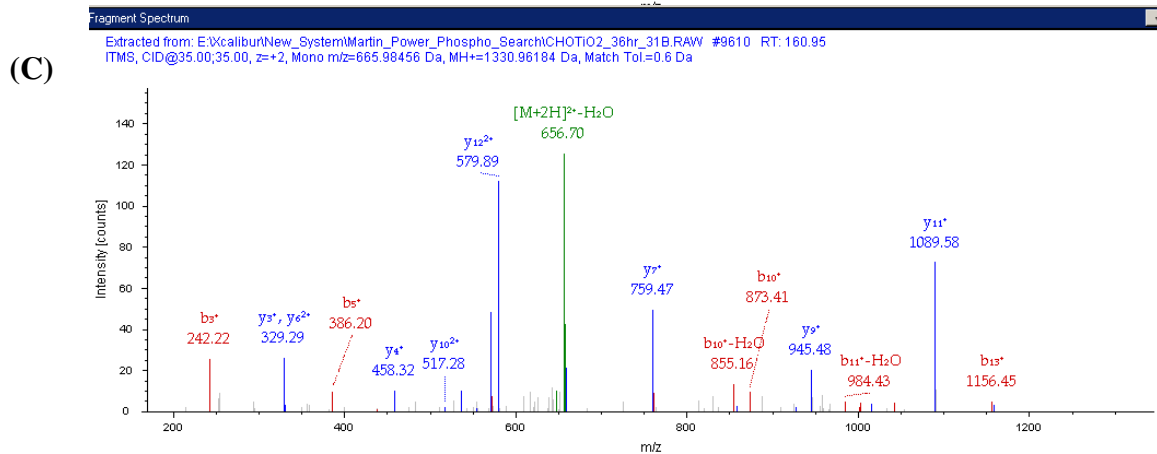
#	b	Seq.	y	#
1	102.055	T		14
2	173.0921	A	1229.612	13
3	242.1135	S	1158.575	12
4	299.135	G	1089.554	11
5	386.167	S-Phospho	1032.532	10
6	473.1991	S	945.5	9
7	572.2675	V	858.468	8
8	673.3151	T	759.3995	7
9	760.3472	S	658.3519	6
10	873.4312	L	571.3198	5
11	1002.474	E	458.2358	4
12	1059.495	G	329.1932	3
13	1156.548	P	272.1717	2
14		R	175.119	1

**Figure 4.2.19** (A) MS/MS fragmentation spectrum output from a MASCOT search of MS features deemed by Progenesis label-free LC/MS software have a statistically significant differential expression (ANOVA value  $\leq 0.05$ ) resulting in the identification of the phosphopeptide “TASSGSSVTSLEGPR” belonging to Protein NDRG1. (B) The b-and y-ion series used for phosphopeptide identification via MASCOT are indicated.



(B)

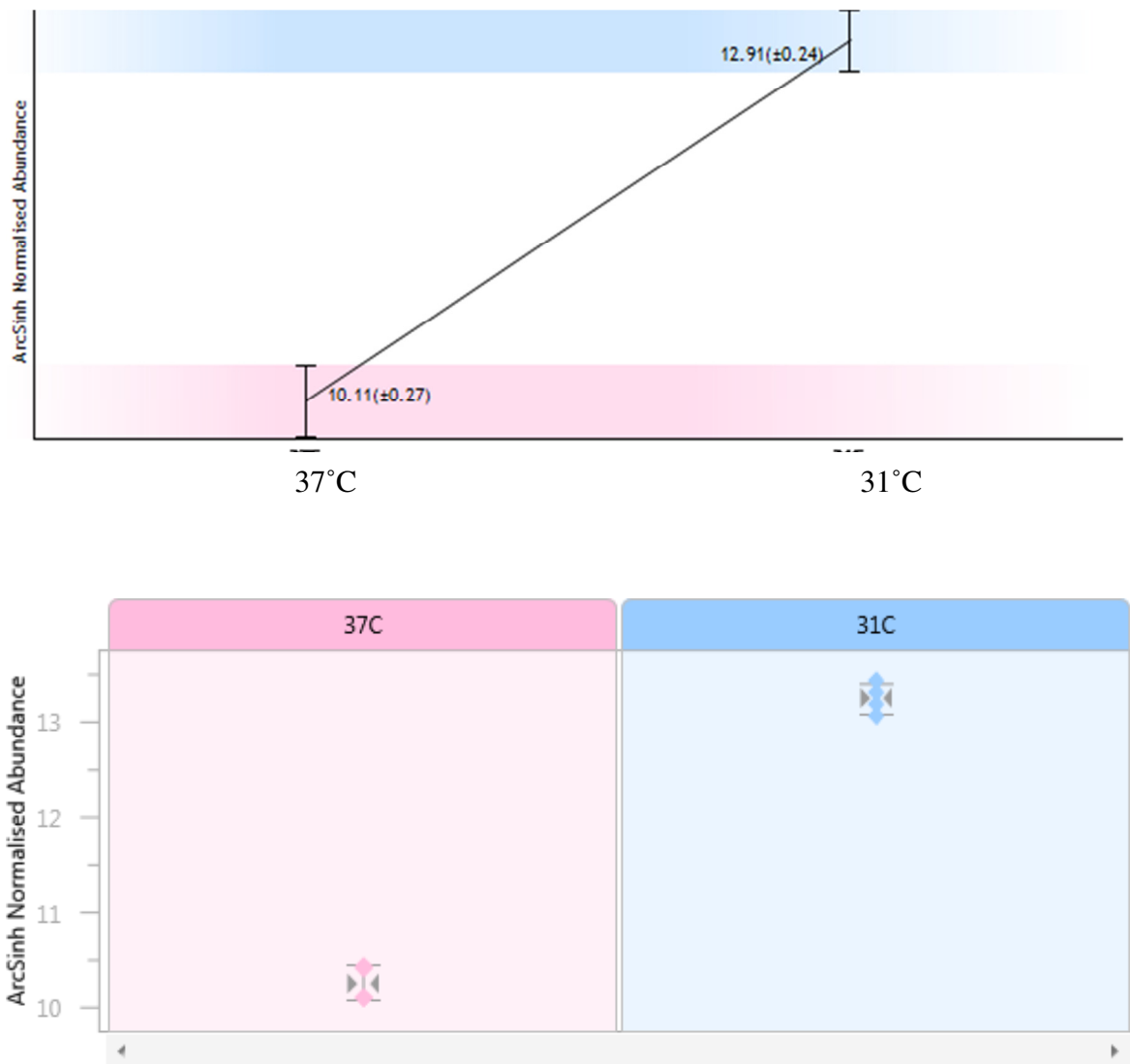
#1	b <sup>+</sup>	Seq.	y <sup>+</sup>	#2
1	102.05496	T		14
2	173.09208	A	1327.589	13
3	340.09044	S-Phospho	1256.5519	12
4	397.11191	G	1089.5535	11
5	484.14394	S	1032.5321	10
6	571.17697	S	945.50003	9
7	670.24439	V	858.468	8
8	771.29207	T	759.39958	7
9	858.3241	S	658.3519	6
10	971.40817	L	571.31987	5
11	1100.4508	E	458.2358	4
12	1157.4722	G	329.1932	3
13	1254.525	P	272.17173	2
14		R	175.11896	1



(D)

#1	b <sup>+</sup>	Seq.	y <sup>+</sup>	#2
1	102.05496	T		14
2	173.09208	A	1229.6121	13
3	242.11354	S-Dehydrated	1158.575	12
4	299.13501	G	1089.5535	11
5	386.16704	S	1032.5321	10
6	473.19907	S	945.50003	9
7	572.26749	V	858.468	8
8	673.31517	T	759.39958	7
9	760.3472	S	658.3519	6
10	873.43127	L	571.31987	5
11	1002.4739	E	458.2358	4
12	1059.4953	G	329.1932	3
13	1156.5481	P	272.17173	2
14		R	175.11896	1

**Figure 4.2.20** (A) MS/MS fragmentation spectrum output from a SEQUEST search of raw files from the LC/MS analysis of titanium dioxide phosphopeptide enriched CHO SEAP whole cell lysate subject to temperature-shift resulting in the identification of the phosphopeptide “TASGSSVTSLEGPR” belonging to Protein NDRG1. (B) The b-and y-ion series used for phosphopeptide identification of the MS/MS are indicated. (C) MS/MS/MS fragmentation spectrum of the peptide “TASGSSVTSLEGPR” following the loss of phosphate (D) The b-and y-ion series used as further evidence of phosphopeptide identification are indicated.



**Figure 4.2.21** Progenesis label-free LC/MS output showing normalised abundance of the phosphopeptide “TASGSSVTSLEGPR” from the Protein NDRG1 indicating its 16.5-fold up-regulation at 31°C compared to 37°C (ANOVA 4.4E-07).

#### **4.2.7.1 Bioinformatic analysis of differentially-expressed phosphopeptides enriched using titanium dioxide**

Differentially-expressed phosphopeptides identified through the label-free LC/MS analysis of titanium dioxide phosphopeptide enrichment of CHO SEAP whole cell lysates subject to temperature-shift were analysed using DAVID. This was carried out in order to determine what pathways were potentially involved in the cellular response to temperature-shift. Unfortunately however, this analysis did not provide any further insight into processes that may be involved (data not shown). This could possibly be due to the relatively small number of proteins being analysed, as DAVID was designed for the analysis of larger data sets involving extensive lists of genes/proteins. In addition to this, pathway activation via phosphorylation events can be complex and may not be possible to fully interpret using functional annotation analysis such as DAVID.

Through manual searching of protein function through a combination of literature mining and using protein databases such as UniProt a number of potentially interesting proteins were identified. Transcription activator BRG1 and Programmed cell death protein 4, both of which have been linked to negative regulation of cell cycle progression, were up-regulated at 31°C compared to 37°C (Table 4.2.47). Protein NDRG1, also up-regulated at 31°C, is known to be involved in cell stress response and growth arrest (Ellen et al. 2008; Wakisaka et al. 2003). Conversely, growth promoting Eukaryotic translation initiation factor 4 gamma 3 was down-regulated at 31°C compared to 37°C (Table 4.2.46).

#### **4.2.7.2 Summary**

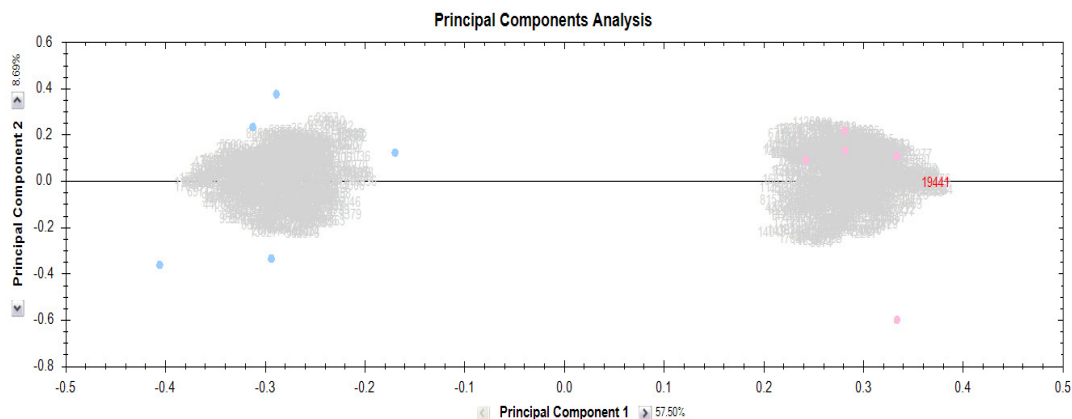
Phosphopeptide enrichment of CHO SEAP whole cell lysate using titanium dioxide significantly improved the number of phosphopeptides identified compared to phosphopeptide enrichment of a phosphoprotein-enriched whole cell lysate sample. Using quantitative label-free LC/MS 33 differentially-expressed phosphopeptides were identified, of which 15 were up-regulated at 31°C. Despite the larger number of phosphopeptides with site-specific information obtained, GO analysis did not reveal any enrichment of process pathways. A number of proteins involved in translation were identified through literature mining however. The utilization of a second phosphopeptide enrichment methodology to complement titanium oxide could increase coverage of the phosphoproteome of CHO cells, providing further information on the effects of temperature-shift.

#### **4.2.8 Quantitative label-free LC/MS analysis of an Iron oxide phosphopeptide enriched CHO SEAP whole cell lysate**

In order to further profile the phosphoproteome of CHO cells a second phosphopeptide enrichment was employed. It was hoped that the use of an Iron oxide based phosphopeptide enrichment would complement the titanium dioxide enrichment carried out in section 4.2.7 and provide greater coverage of the phosphoproteome. Similar to titanium dioxide enrichment described previously, 500µg of all eight CHO SEAP whole cell lysate samples from temperature-shifted and non-temperature-shifted samples were phosphopeptide enriched using Iron oxide spin columns as detailed in materials 2.4.5. Phosphopeptide enriched samples were then prepared for quantitative label-free LC/MS as detailed in section 2.4.6.

As detailed in section 4.2.7, to determine whether the differential expression of the phosphopeptide could be the result of a change in the abundance of the protein as opposed to a change in the level of phosphorylation, the lists of differentially-expressed phosphopeptides was compared to the lists of differentially-expressed proteins from the CHO SEAP whole cell lysate and phosphoprotein-enriched analyses (sections 4.2.4 and 4.2.5). Any proteins identified as being differentially-expressed in both the iron-enriched and whole cell lysate / phosphoprotein-enriched analysis were then indicated by ^ in the iron list. As described in section 4.2.6, the protein names of the phosphopeptides were also searched against qualitative lists from the analysis of whole cell lysate and phosphoprotein enrichment. Proteins identified as being common to the differentially-expressed phosphopeptide lists and qualitative lists are denoted by an †. This is indicative of a change in the abundance of the phosphopeptide as opposed to a change in the abundance of the total protein.

Progenesis label-free software was used to analyse the LC/MS data in duplicate and the same criteria applied to define differentially-expressed phosphopeptides was applied as detailed in section 2.5.8. A principal component analysis (PCA) plot generated by Progenesis during data analysis showing how the Iron phosphopeptide enriched samples from 37°C and 31°C cultures cluster and is displayed in Figure 4.2.22.



**Figure 4.2.22** PCA plot for CHO SEAP Iron phosphopeptide enriched whole cell lysate sample groups 37°C (pink) and 31°C (blue) generated during quantitative label-free LC/MS data analysis using Progenesis software. It should be noted that the five spots from each condition represent the four enriched samples that were originally analysed and an additional run based on an inclusion list for each condition that was used to obtain additional peptide information

As a result of this analysis 60 unique phosphopeptides were identified as differentially-expressed when lists from two independent analyses of LC/MS label-free data were compared. Of these, 38 were down-regulated at 31°C (Table 4.2.48), while the remaining 22 were up-regulated 31°C compared to 37°C (Table 4.2.48).



**Table 4.2.48** List of differentially-expressed phosphopeptides, down-regulated at 31°C compared to 37°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift. Phosphopeptides were enriched from CHO SEAP whole cell lysates using iron oxide phosphopeptide enrichment. (Note: \* denotes phosphopeptides with MS<sup>3</sup> information. † indicates proteins that don't change in total abundance. . ^ indicates proteins that do change in total abundance.)

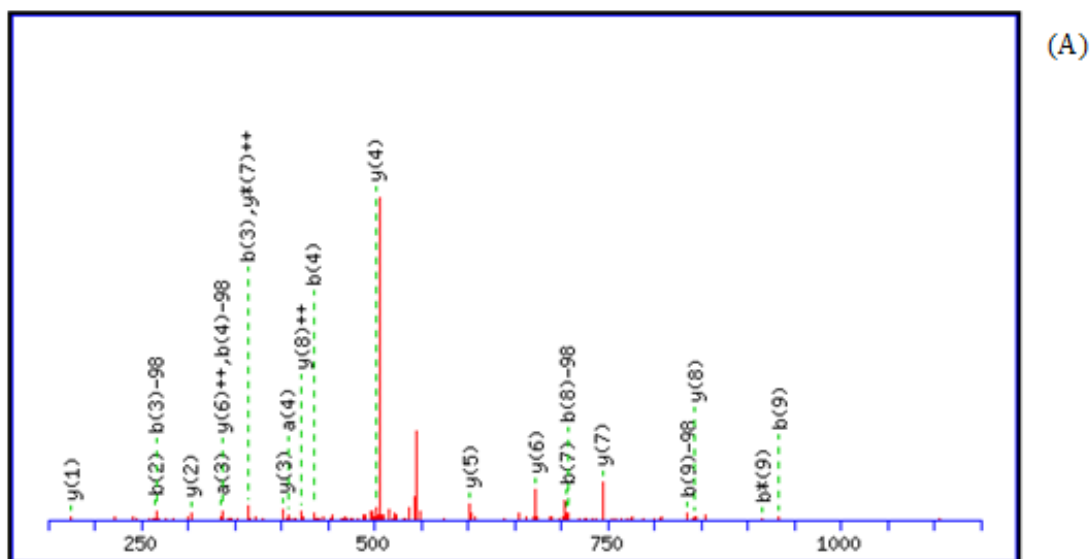
Gene Symbol	Sequence	Description	Score	Anova	Fold	Function
Nr2c2	IQIISTDSAVA <u>S</u> PQR	Nuclear receptor subfamily 2 group C member 2	84	0.0011	2.6	Transcription
SAPCD2	SP <u>S</u> ADASAAVCK*	Protein C9orf140 homolog	79	0.0041	1.8	Unknown
SPAST	AP <u>S</u> CSGLSMVSGGR*	Spastin	78	0.0469	1.8	Cell Division
GIGYF1	<u>S</u> IEEGDGA <u>F</u> GR*	PERQ amino acid-rich with GYF domain-containing protein 1	68	0.0348	15.9	Involved in tyrosine kinase receptor signalling
ANAPC1	SP <u>S</u> ISNMAALSR	Anaphase-promoting complex subunit 1	65	0.0250	1.9	cell division
TBC1D5	VEQALSEVAS <u>S</u> LQSSAPK	Tetratricopeptide repeat protein 7B	62	0.0122	4.7	Unknown
EIF4G3	TS <u>S</u> P TTLAPLAR	Eukaryotic translation initiation factor 4 gamma 3	60	0.0182	3.6	Translation
CCDC88A	<u>S</u> <u>S</u> SQENLLDEVVK †	Girdin	60	0.0276	5.1	Cell proliferation
ARID5B	AV <u>S</u> PLDPAK	AT-rich interactive domain-containing protein 5B	60	0.0461	2.0	Transcription
PHF12	TT <u>S</u> PSSD TLLDR	PHD finger protein 12	59	0.0360	2.0	Transcription
MAP1B	SD <u>S</u> PLTPR^	Microtubule-associated protein 1B	57	0.0003	7.6	Structural
ELMOD3	<u>S</u> SELTEAVETK*	ELMO domain-containing protein 3	57	0.0024	2.7	Miscellaneous
NEK1	SE <u>S</u> MPVQLNK	TBC1 domain family member 5	56	0.0125	2.7	Unknown
GOLGA4	SE <u>S</u> PPQSGDTQTFAQK	Golgin subfamily A member 4	55	0.0384	1.9	Unknown
URI1	<u>S</u> L <u>S</u> CCEASEGR*	Unconventional prefoldin RPB5 interactor	55	0.0001	3.3	Transcription
KMT2D	ASQVEPQ <u>S</u> PGLGLR	Histone-lysine N-methyltransferase MLL2	54	0.0352	3.2	Transcription
EIF4G3	<u>S</u> PVAATVVQR*	Eukaryotic translation initiation factor 4 gamma 3	54	0.0001	3.6	Translation
ARHGAP17	<u>S</u> SGTNFQGLPSK	Rho GTPase-activating protein 17	54	0.0130	2.1	Positive regulation of GTPase activity
DTD1	SASSGAEGDV <u>S</u> EREP	D-tyrosyl-tRNA(Tyr) deacylase 1	53	0.0012	2.1	DNA replication

ZNF687	AVVLPGGNVT <u>S</u> PK	Zinc finger protein 687	53	0.0471	2.1	Transcription
MYO9A	AAST <u>I</u> DLGAGEAVVGK	Myosin-IXa	51	0.0288	2.7	Positive regulation of GTPase activity
UBN1	IC <u>S</u> DEEEDEEKGGR	Ubinuclein-1	51	0.0214	2.8	Miscellaneous
CDCA8	LTAEAIQ <u>I</u> PLK	Borealin	50	0.0421	3.1	Cell cycle
ACAP2	SSPSTG <u>S</u> LDSGNEK	Arf-GAP with coiled-coil, ANK repeat and PH domain-containing protein 2	49	0.0292	2.9	Positive regulation of GTPase activity
ATF2	NDSVIVADQ <u>T</u> PTPTR*	Cyclic AMP-dependent transcription factor ATF-2	49	0.0000	7.4	Transcription
PUM1	RD <u>S</u> LTGSSDLYK*	Pumilio homolog 1	46	0.0184	2.1	Translation
SPAG9	NVSTD <u>S</u> VENEEK	C-Jun-amino-terminal kinase-interacting protein 4	45	0.0471	2.0	Activation of JUN kinase activity
TLN2	LDEG <u>T</u> PPEPK	Talin-2	45	0.0398	1.7	Structural
MAST2	KL <u>S</u> NPDLFSSTGK*	Microtubule-associated serine/threonine-protein kinase 2	45	0.0062	3.0	Miscellaneous
FNBP1L	TI <u>S</u> DGTISASK*	Formin-binding protein 1-like	44	0.0112	2.7	Autophagy
NFIB	DQDMS <u>S</u> PPTMK	Nuclear factor 1 B-type	44	0.0010	3.1	Transcription
AMOTL2	TA <u>S</u> LDSVATTR	Angiomotin-like protein 2	44	0.0020	4.6	Wnt signalling pathway
HEG1	EAIEMHENG <u>S</u> TK*	Protein HEG-like 1	43	0.0436	2.4	Cell membrane
ZKSCAN3	MEDVAPGL <u>S</u> PR	Zinc finger protein with KRAB and SCAN domains 3	42	0.0032	7.1	Transcription
OSBPL3	ALVHQL <u>S</u> NESR*	Oxysterol-binding protein-related protein 3	41	0.0000	5.3	Lipid transport
BOD1L1	RG <u>S</u> ISQEITK*	Biorientation of chromosomes in cell division protein 1-like	41	0.0492	1.9	DNA Binding
EMSY	LTSPVTSI <u>S</u> PIQASEK	Protein EMSY	41	0.0156	2.6	Transcription
NEK1	TC <u>S</u> LPDLSK	Serine/threonine-protein kinase Nek1	40	0.0076	2.3	Cell division

**Table 4.2.49** List of differentially-expressed phosphopeptides, up-regulated at 31°C, obtained from quantitative label-free LC/MS analysis of CHO SEAP cells sampled at 36hrs post-temperature-shift. Phosphopeptides were enriched from CHO SEAP whole cell lysates using iron oxide phosphopeptide enrichment. (Note: \* denotes phosphopeptides with MS<sup>3</sup> information. † indicates proteins that don't change in total abundance. . ^ indicates proteins that do change in total abundance.)

Gene Symbol	Sequence	Description	Score	Anova	Fold	Function
DTL	AGLVMVASSQT <b>I</b> PAK	Denticleless protein homolog	69	0.0221	1.8	Cell division
CDC42EP1	LNFDSSP <b>T</b> SSTDGR	Cdc42 effector protein 1	63	0.0028	1.8	Multiple
KTN1	EIQNGAIHE <b>S</b> DSENVLR* †	Kinectin	61	0.0059	3.3	Microtubule-based movement
KTN1	KNQNGAIHE <b>S</b> DSENVLR†	Kinectin	61	0.0059	3.3	Microtubule-based movement
FAT1	DMPAAGSLG <b>S</b> SSR^	Protocadherin Fat 1	59	0.0268	2.2	Cell adhesion.
DDX21	KSP <b>S</b> EDDMEPPK*^	Nucleolar RNA helicase 2	57	0.0119	1.7	RNA processing
N/A	VLQPSPV <b>S</b> PSDTR	Component of gems 4-like	55	0.0014	3.7	mRNA processing
FAM53B	SS <b>S</b> FSLPAR	Protein FAM53B-like	53	0.0223	1.8	mRNA processing
NDRG1	SRT <b>A</b> S <b>S</b> GSSVTSLEGPR	Protein NDRG1	52	0.0020	230.5	Miscellaneous
PSMA5	GVNTF <b>S</b> PEGR^	Proteasome subunit alpha type-5	51	0.0004	3.0	Cell cycle
TRIP12	DDSLD <b>S</b> PQGR*	Probable E3 ubiquitin-protein ligase TRIP12	51	0.0167	2.0	Protein degradation
NCL	GL <b>S</b> EDTTEETLK*^	Nucleolin	50	0.0036	5.1	Miscellaneous
EFS	LP <b>S</b> TESLSR	Embryonal Fyn-associated substrate	49	0.0364	2.0	Cell adhesion.
ATF7	TD <b>S</b> VIADQTPTR	Cyclic AMP-dependent transcription factor ATF-7	46	0.0098	9.8	Transcription
PSAT1	A <b>S</b> LYNAVTTEDVEK*^	Phosphoserine aminotransferase	46	0.0059	5.0	Miscellaneous
TOP2B	KA <b>S</b> GSENEGDPGR†	DNA topoisomerase 2-beta	45	0.0272	4.4	Miscellaneous
VIM	TY <b>S</b> LGSAALRPSTSR^	Vimentin	44	0.0009	13.2	Miscellaneous
PTPN21	T <b>S</b> ASAADVAPR	Tyrosine-protein phosphatase non-receptor type 21	44	0.0397	1.6	Transcription
VAPB	SLT <b>S</b> LLDDAEVK†	Vesicle-associated membrane protein-associated protein B	42	0.0002	5.8	Unfolded protein response
TTC7B	EVDYSD <b>S</b> LTEK	Transcription activator BRG1-like	41	0.0004	3.0	Transcription
ERCC6	LPGSVNT <b>S</b> FNSR	DNA excision repair protein ERCC-6	41	0.0056	2.0	Transcription
SMA4	KIPDP <b>S</b> DDVSEVDAR	Transcription activator BRG1-like	41	0.0144	2.6	Transcription

The phosphopeptide “SPVAATVVQR” from the protein Eukaryotic translation initiation factor 4 gamma 3 was identified following a search of MS/MS information, through MASCOT (Figure 4.2.23), of those features identified by Progenesis label-free LC/MS software as having a statistically significant differential expression (ANOVA value  $\leq 0.05$ ). “SPVAATVVQR” was also identified following a qualitative search of all LC/MS raw files from mass spectral analysis of iron oxide phosphopeptide enriched CHO SEAP lysates using the SEQUEST search algorithm (detailed section 2.4.6). An example of the MS/MS output from this identification in SEQUEST is given in (Figure 4.2.24 A). PhosphoRS identified the first amino acid residue, serine, as being phosphorylated with a >99% level of confidence. MS/MS/MS information on this peptide was also obtained during mass spectral analysis as can be seen in (Figure 4.2.24 B) thus providing further evidence of phosphorylation of this phosphopeptide. Figure 4.2.25 shows an output from Progenesis label-free analysis identifying “SPVAATVVQR” as being 3.6-fold up-regulated at 37°C (ANOVA 0.0001).

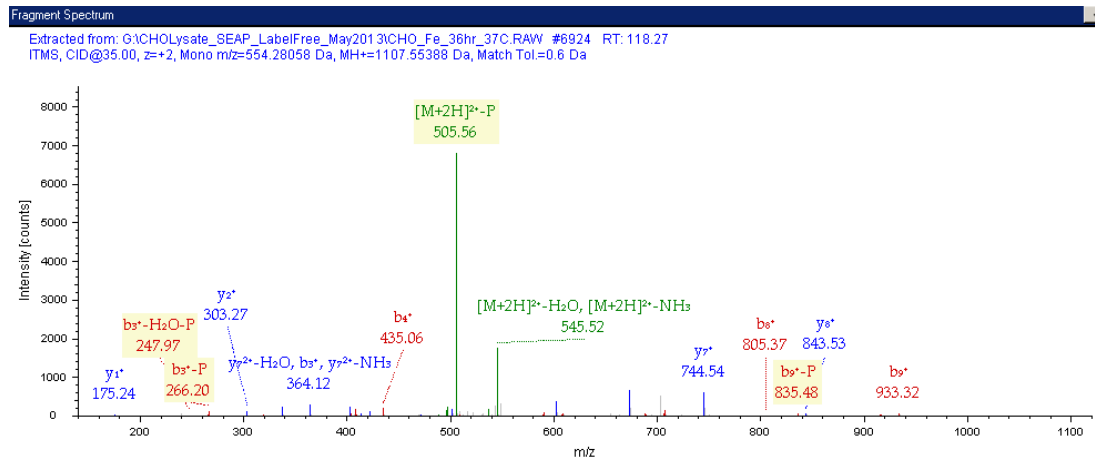


(B)

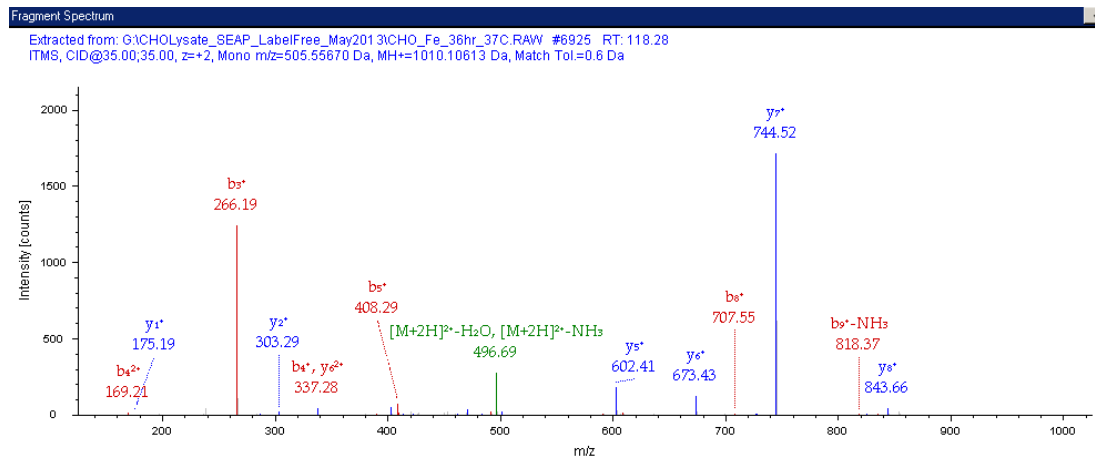
#	b	Seq.	y	#
1	70.0287	S		10
2	167.0815	P	940.5574	9
3	266.1499	V	843.5047	8
4	337.187	A	744.4363	7
5	408.2241	A	673.3992	6
6	509.2718	T	602.362	5
7	608.3402	V	501.3144	4
8	707.4087	V	402.2459	3
9	835.4672	Q	303.1775	2
10		R	175.119	1

**Figure 4.2.23** (A) MS/MS fragmentation spectrum output from a MASCOT search of MS features deemed by Progenesis label-free LC/MS software have a statistically significant differential expression (ANOVA value  $\leq 0.05$ ) resulting in the identification of the phosphopeptide “SPVAAATVVQR” from the protein Eukaryotic translation initiation factor 4 gamma 3. (B) The b-and y-ion series used for phosphopeptide identification via MASCOT are indicated.

(A)



(C)



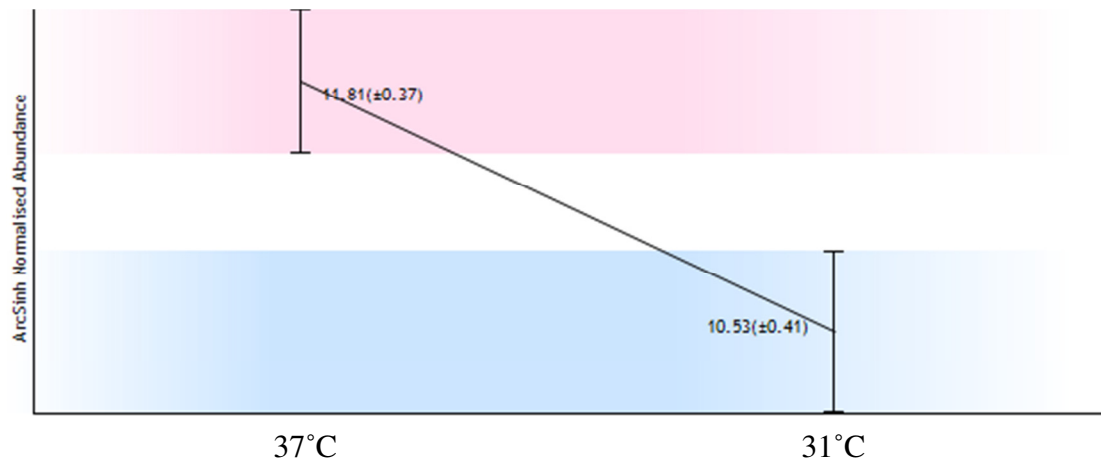
#1	b <sup>+</sup>	Seq.	y <sup>+</sup>	#2
1	168.0056	<b>S-Phospho</b>		10
2	<b>265.0584</b>	P	940.5575	9
3	<b>364.1268</b>	V	<b>843.5047</b>	8
4	<b>435.164</b>	A	<b>744.4363</b>	7
5	506.2011	A	<b>673.3992</b>	6
6	<b>607.2488</b>	T	<b>602.3621</b>	5
7	<b>706.3172</b>	V	<b>501.3144</b>	4
8	<b>805.3856</b>	V	<b>402.246</b>	3
9	<b>933.4442</b>	Q	<b>303.1775</b>	2
10		R	<b>175.119</b>	1

(B)

#1	b <sup>+</sup>	Seq.	y <sup>+</sup>	#2
1	70.02874	<b>S-Dehydrated</b>		10
2	<b>167.0815</b>	P	940.5575	9
3	<b>266.1499</b>	V	<b>843.5047</b>	8
4	<b>337.1871</b>	A	<b>744.4363</b>	7
5	<b>408.2242</b>	A	<b>673.3992</b>	6
6	509.2719	T	<b>602.3621</b>	5
7	<b>608.3403</b>	V	<b>501.3144</b>	4
8	<b>707.4087</b>	V	<b>402.246</b>	3
9	<b>835.4673</b>	Q	<b>303.1775</b>	2
10		R	<b>175.119</b>	1

(D)

**Figure 4.2.24** (A) MS/MS fragmentation spectrum output from a SEQUEST search of raw files from the LC/MS analysis of titanium dioxide phosphopeptide enriched CHO SEAP whole cell lysate subject to temperature-shift resulting in the identification of the phosphopeptide “SPVAATVVQR” belonging to Eukaryotic translation initiation factor 4 gamma 3. (B) The b-and y-ion series used for phosphopeptide identification of the MS/MS are indicated. (C) MS/MS/MS fragmentation spectrum of the peptide “SPVAATVVQR” following the loss of phosphate (D) The b-and y-ion series used as further evidence of phosphopeptide identification are indicated.



**Figure 4.2.25** Progenesis label-free LC/MS output showing normalised abundance of the phosphopeptide “**S**SPVAATVVQR” indicating its 3.6-fold down-regulation at 31°C compared to 37°C (ANOVA 0.0001).

#### **4.2.8.1 Bioinformatic analysis of differentially-expressed phosphopeptides enriched using iron oxide**

Identification of differentially-expressed phosphopeptides through quantitative label-free LC/MS analysis of iron oxide phosphopeptide enriched CHO SEAP whole cell lysates subject to temperature-shift were analysed using DAVID. Similar to DAVID analysis of titanium dioxide enriched phosphopeptides (section 4.2.7), analysis of iron oxide phosphopeptide enriched samples did not identify any pathways of interest that may be involved in the cellular response to temperature-shift (data not shown). Despite the list of differentially-regulated phosphopeptides identified using iron oxide being longer than that obtained using titanium dioxide, the failure to identify any obvious processes involved in the response to temperature-shift through DAVID is most likely for the same reasons as those discussed previously (Section 4.2.7).

Identification of proteins of interest was once again carried out through manual searching of protein function using a combination of literature mining and protein databases such as UniProt. Interestingly, as well as being identified through titanium dioxide enrichment, Transcription activator BRG1, Protein NDRG1 and Eukaryotic translation initiation factor 4 gamma 3 were also identified as differentially-expressed using the iron oxide phosphopeptide enrichment strategy (Table 4.2.50 and Table 4.2.51). In addition to this, phosphopeptides from the proteins Borealin and Serine/threonine-protein kinase Nek1, both of which are known to be involved in cell growth and division, were found to be up-regulated at 37°C.

It is also interesting to note that when the lists of differentially-expressed phosphopeptides from titanium dioxide and iron oxide enrichment methods are compared, it can be seen that a number of phosphopeptides are common to both lists. Closer examination of these phosphopeptides shows that they also share commonality in the condition under which they are up-regulated and the approximate fold-change. Proteins containing these commonly identified phosphopeptides include ELMO domain-containing protein 3, Eukaryotic translation initiation factor 4 gamma 3 (eif4g3), Protein NDRG1 and Vesicle-associated membrane protein-associated protein B (Table 4.2.50 and Table 4.2.51).



**Table 4.2.50** List of proteins and associated differentially-expressed phosphopeptides common to quantitative label-free LC/MS analysis of titanium dioxide phosphopeptide enriched (TiO<sub>2</sub>) and iron oxide phosphopeptide enriched (Fe<sup>+</sup>) CHO SEAP whole cell lysates subject to temperature-shift. This table shows the phosphopeptides that were down-regulated at 31°C.

Gene Symbol	Protein	Enrichment Method	Sequence	Fold	Score	Down	Anova
EIF4G3	Eukaryotic translation initiation factor 4 gamma 3	TiO <sub>2</sub>	<u>S</u> PVAATVVQR	3.7	47	31°C	0.0060
		Fe <sup>+</sup>	<u>S</u> PVAATVVQR	3.6	54	31°C	0.0001
		Fe <sup>+</sup>	T <u>S</u> SPTTLAPLAR	3.6	60	31°C	0.0182
ELMOD3	ELMO domain-containing protein 3	TiO <sub>2</sub>	<u>S</u> SELTEAVETK	2.4	53	31°C	0.0003
		Fe <sup>+</sup>	<u>S</u> SELTEAVETK	2.7	57	31°C	0.0024
ANAPC1	Anaphase-promoting complex subunit 1	TiO <sub>2</sub>	SP <u>S</u> ISNMAALSR	1.5	69	31°C	0.0258
		Fe <sup>+</sup>	SP <u>S</u> ISNMAALSR	1.9	65	31°C	0.0250

**Table 4.2.51** List of proteins and associated differentially-expressed phosphopeptides common to quantitative label-free LC/MS analysis of titanium dioxide phosphopeptide enriched (TiO<sub>2</sub>) and iron oxide phosphopeptide enriched (Fe<sup>+</sup>) CHO SEAP whole cell lysates subject to temperature-shift. This table shows the phosphopeptides that were up-regulated at 31°C.

Gene Symbol	Protein	Enrichment Method	Sequence	Fold	Score	High	Anova
NDRG1	Protein NDRG1	TiO <sub>2</sub>	SRTA <u>S</u> GSSVTSLEGPR	39.7	49	31°C	0.0001
		Fe <sup>+</sup>	SRTA <u>S</u> GSSVTSLEGPR	230.5	52	31°C	0.0020
		TiO <sub>2</sub>	TA <u>S</u> GSSVTSLEGPR	16.5	88	31°C	0.0000
SMARCA4	Transcription activator BRG1	TiO <sub>2</sub>	EVDYSD <u>S</u> LTEK	2.8	44	31°C	0.0001
		Fe <sup>+</sup>	EVDYSD <u>S</u> LTEK	3.0	41	31°C	0.0004
		Fe <sup>+</sup>	KIPDPD <u>S</u> DDVSEVDAR	2.6	41	31°C	0.0144
VAPB	Vesicle-associated membrane protein-associated protein B	TiO <sub>2</sub>	<u>S</u> LTSSLDDAEVKK	3.0	56	31°C	0.0048
		Fe <sup>+</sup>	SLTS <u>S</u> LDDAEVK	5.8	42	31°C	0.0002
CDC42EP1	Cdc42 effector protein 1	TiO <sub>2</sub>	LNFDSSP <u>T</u> SSTDGR	1.7	44	31°C	0.0020
		Fe <sup>+</sup>	LNFDSSP <u>T</u> SSTDGR	1.8	63	31°C	0.0028

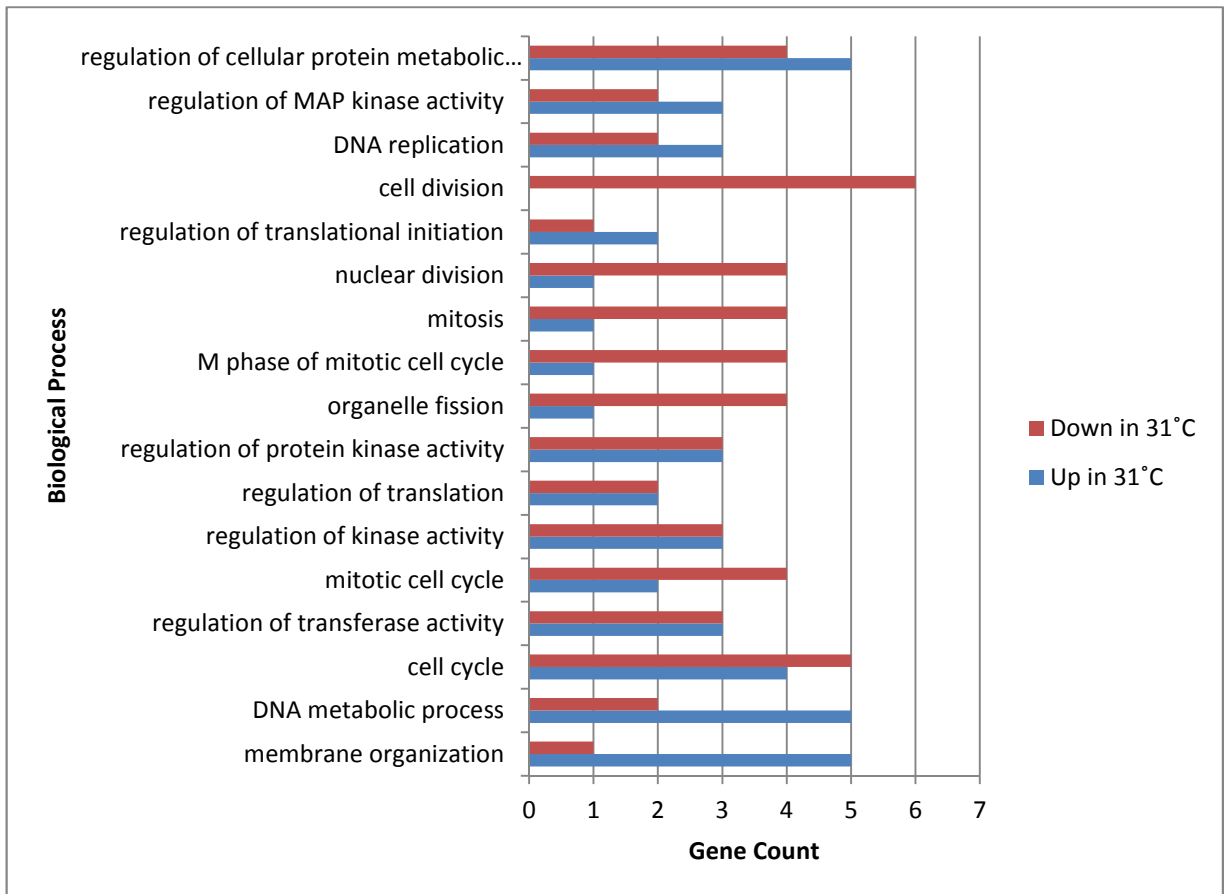
#### **4.2.8.2 Bioinformatic analysis of all differentially-expressed phosphopeptides enriched using gallium oxide, titanium dioxide and iron oxide**

Due to the low number of pathways/processes identified when analysing differentially-expressed phosphopeptides obtained using DAVID, it was decided to perform analysis on all the phosphopeptides pooled together. This would help provide further insight into the biological processes that the identified differentially-expressed phosphopeptides are involved in because although the phosphopeptides were identified using different enrichment methods this is merely as a result of the different affinities of each column, the pathways that these phosphopeptides are involved in are likely to overlap. Table 4.2.52 shows the list of enriched processes identified as a result of using this approach. A breakdown of the number of proteins associated with each process is given in Figure 4.2.26. Although the number of proteins from each condition are evenly divided across many of the processes e.g. regulation of kinase activity and regulation of transferase activity, there are also a number of processes that are obviously enriched for each temperature. Membrane organisation appears increased at 31°C compared to 37°C while processes involved in cell division are down-regulated at 31°C compared to 37°C.

**Table 4.2.52** GO biological process enrichment for differentially-expressed phosphopeptides identified using quantitative label-free LC/MS analysis of gallium oxide, titanium dioxide and iron oxide enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of a p-value  $\leq 0.05$ . Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category.

<b>Term</b>	<b>Count</b>	<b>P-Value</b>
membrane organization	6	0.045
DNA metabolic process	7	0.044
cell cycle	9	0.043
regulation of transferase activity	6	0.041
mitotic cell cycle	6	0.040
regulation of kinase activity	6	0.035
regulation of translation	4	0.033
regulation of protein kinase activity	6	0.031
organelle fission	5	0.029
M phase of mitotic cell cycle	5	0.027
mitosis	5	0.026
nuclear division	5	0.026
regulation of translational initiation	3	0.018
cell division	6	0.017
DNA replication	5	0.016
regulation of MAP kinase activity	5	0.006
regulation of cellular protein metabolic process	9	0.003

**Figure 4.2.26** GO biological process enrichment for differentially-expressed phosphopeptides identified using quantitative label-free LC/MS analysis of gallium oxide, titanium dioxide and iron oxide enriched samples from CHO SEAP cells sampled at 36hrs post-temperature-shift. Enrichment was considered significant upon observation of a p-value  $\leq 0.05$ . Gene Count corresponds to the number of gene symbols from the list of differentially-regulated proteins that are implicated in a particular GO category. This graph shows the breakdown of the gene counts in those enriched pathways.

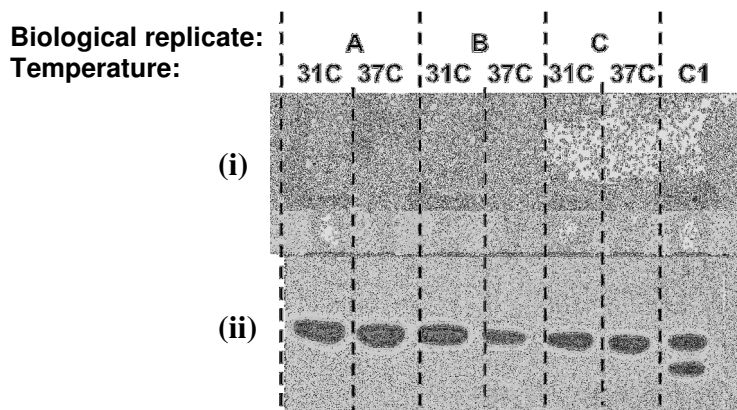


#### **4.2.9 Western blot validation of differentially-expressed phosphopeptides identified by quantitative label-free LC/MS analysis of temperature-shifted CHO SEAP cells**

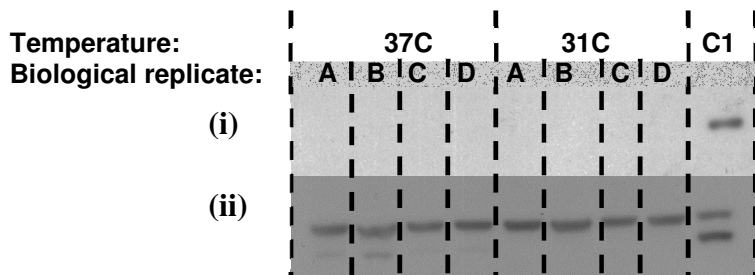
Western blot analysis was used to validate the results obtained in 4.2.7 and 4.2.8. Despite identifying the differential expression of 82 phosphopeptides, only antibodies specific to phosphopeptides from the proteins NDRG1 and ATF2 could be sourced. Antibody selection was based on the ability to detect peptides with homologous amino acid sequences to those identified in this study in species such as Human, mouse and rat. The second requirement was that the antibody was capable of detecting the correct site of phosphorylation on the peptide of interest. Based on these stipulations, the only antibodies available, suitable for the purpose of validation were anti-NDRG1 and anti-ATF2. As such, antibodies raised against the phosphorylated form of the peptide identified and the total protein were purchased for the proteins NDRG1 and ATF2.

Following quantitative label-free LC/MS analysis of titanium dioxide phosphopeptide enriched temperature-shifted CHO SEAP whole cell lysates, the monophosphorylated phosphopeptide TAsGSSVTSLEGPR was found to be up-regulated 16.5-fold at 31°C compared to 37°C (Table 4.2.47 in section 4.2.7). Western blot validation of the phosphorylated form of NDRG1 clearly shows an increase in the phosphorylated form of the protein at 31°C compared to 37°C (Figure 4.2.27). Although total NDRG1 was detected in the control cell line (SKBR3) it was not detected in the CHO samples (Figure 4.2.28); This is possibly as a result of the antibody for total NDRG1 not being able to detect the CHO form of the protein or that the levels of NDRG1 are so low in CHO, that the antibody was unable to detect it. It should be noted that a search of the RAW files from the LC/MS analysis of CHO whole cell lysate (section 4.2.4) failed to identify the NDRG1 protein in the (unenriched) samples, this indicates that NDRG1 is likely to be a low abundance protein and lends further weight to the possibility that it could not be detected by the antibody due to its low abundant nature.

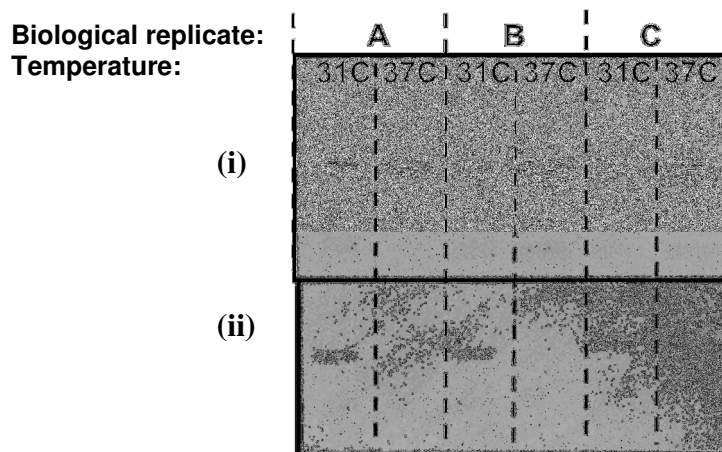
Figure 4.2.29 (i) shows a western blot confirming the down-regulation of the phosphopeptide NDSVIVADQ**I**PTPTR from the protein ATF2. This result is in keeping with the finding that the phosphopeptide NDSVIVADQ**I**PTPTR was 7.4-fold down-regulated at 31°C compared to 37°C as identified by iron oxide phosphopeptide enriched, temperature-shifted CHO SEAP whole cell lysate (shown in Table 4.2.48 in section 4.2.7). Conversely, Figure 4.2.29 (ii) shows that the expression of total ATF2 is slightly increased by temperature-shift.



**Figure 4.2.27** (i) Western blot validation of the differential expression of the phosphopeptide TAS**G**SSVTSLEGPR from the protein NDRG1 identified as being up-regulated at 31°C compared to 37°C in both titanium dioxide enriched and iron oxide enriched phosphopeptide samples following analysis by quantitative label-free LC/MS. (ii) shows western blot of alpha tubulin as a loading control. A breast cancer cell line, SKBR3, acts as a positive control (C1).



**Figure 4.2.28** (i) Western blot showing that no bands for the total protein NDRG1 were detected in CHO despite NDRG1 being detected in the positive control, a breast cancer cell line, SKBR3. (ii) shows western blot of GAPDH as a loading control and SKBR3 acting as a positive control (C1).



**Figure 4.2.29** (i) Western blot validation of the differential expression of the phosphopeptide NDSVIVADQ**T**PTPTR from the protein ATF2 identified as being down-regulated at 31°C compared to 37°C in iron oxide enriched phosphopeptide samples following analysis by quantitative label-free LC/MS. (ii) shows the up-regulation of total ATF2 at 31°C compared to 37°C.



#### **4.2.9.1 Summary**

Quantitative label-free LC/MS analysis of Iron oxide phosphopeptide enriched CHO SEAP whole cell lysates resulted in the identification of the differential expression of 60 phosphopeptides, 22 of which were up-regulated at 31°C. Based on the GO analysis of the differentially-expressed phosphopeptides enriched by Iron oxide, no process pathways were found to be enriched. A similar result was obtained for GO analysis of differentially-expressed phosphopeptides enriched by titanium dioxide. When the lists of proteins identified by iron oxide and titanium dioxide enrichment strategies are combined however, GO analysis revealed alterations in pathways involved in cell division and membrane organisation.

#### **4.2.10 Overall summary of quantitative label-free LC/MS analysis of CHO SEAP cells subjected to temperature-shift.**

GO analysis of the 114 differentially-expressed proteins identified by quantitative label-free LC/MS analysis of CHO SEAP whole cell lysate samples showed enrichment of pathways involved in translation, RNA processing and ribosomal biogenesis while cytoskeleton organisation pathways were found to be enriched following similar analysis of phosphoprotein-enriched CHO samples.

While gallium oxide phosphopeptide enrichment of phosphoprotein-enriched whole cell lysates provided site-specific phosphorylation information on a number of phosphopeptides, by far the greatest contribution came from titanium dioxide and iron oxide phosphopeptide enrichment of whole cell lysates. Together, these strategies identified 95 differentially-expressed phosphopeptides (41 up-regulated and 54 down-regulated at 31°C compared to 37°C). GO analysis of the proteins identified revealed down-regulation of cell cycle and cell division pathways and up-regulation of pathways involved in membrane organisation.

Further protein fractionation or validation by western blot would be required to definitively determine whether the changes observed are as a result of changes in phosphorylation. However, to the best of my knowledge, this is the first time a quantitative label-free LC/MS / phosphopeptide enrichment strategy has been applied to the profiling of temperature-shifted CHO cells.

### **4.3 Overall summary of results from phosphoproteomic and host cell protein analysis of temperature shifted CHO cell lines**

Qualitative label-free LC/MS analysis of temperature-shifted and non-temperature-shifted conditioned media samples from an IgG-producing and non-producing cell line found that the cell line and culture conditions used both impacted on the HCP profile generated. Qualitative analysis of this data identified changes in the abundance of proteins, over time in culture under both temperature-shifted and non-temperature-shifted conditions, that could negatively impact product quality such as proteases (Cathepsin B, Cathepsin D, MMP9) and glycosidases (beta-galactosidase and  $\alpha$ -N-acetylgalactosaminidase). In addition, proteins that could enhance cell growth such as vascular endothelial growth factor isoforms A and C and Hepatoma-derived growth factor were also identified.

In order to decipher some of the cell-signalling events involved in the response of CHO cells to temperature shift, analysis of the phosphoproteome of a SEAP secreting CHO cell line was conducted using a quantitative label-free LC/MS based approach. The use of Gallium, Iron and Titanium phosphopeptide enrichment strategies facilitated the identification of 1,307 unique phosphopeptides (1,480 phosphosites). Enrichment of pathways involved in protein synthesis and cell cycle progression were identified following Gene Ontology analysis of the 92 phosphopeptides determined as being significantly differentially-expressed 36hrs post temperature-shift. Among the proteins identified, many are potentially involved in controlling growth and recombinant protein production in CHO cells including translation initiation factors EIF5B, EIF4G3, Transcription activator BRG1 and the tumour suppressor protein, Protein NDRG1.

# Discussion

## **5.0 Host Cell Protein (HCP) Analysis Overview**

The following section discusses the quantitative label-free LC/MS analysis of host cell proteins in the conditioned media of bioprocess relevant mammalian cell lines. Initially this work focused on the analysis of cell culture supernatant taken at different time points from a non-producing Hybridoma cell line grown under temperature-shifted and non-temperature-shifted culture conditions. Based on the finding that the HCP content changed in abundance over time and under different culture temperatures it was decided that this work would be extended to include analysis of what could perhaps be described as the workhorse of the Biopharmaceutical industry: the CHO cell. Quantitative label-free LC/MS analysis of a non-producing and an IgG secreting CHO cell line also found that the time in culture and culture temperature impacted on the HCP profile of both cell lines. This analysis revealed the differential expression of proteases such as Cathepsin B and MMP-9, and glycosidases including Lysosomal  $\alpha$ -glucosidase and  $\alpha$ -N-acetylgalactosaminidase, both of which have the potential to negatively impact on product quality. In addition to this, the expression of proteins that could affect culture growth and viability such as Clusterin and Hepatoma-derived growth factor were also found to alter over time under temperature-shifted and non-temperature-shifted conditions. Finally, this section looks at the recent release of two CHO specific proteomic databases and the benefits associated with the availability of these resources to this study.

## **5.1 Characterisation of HCP profiles from a non-producing hybridoma cell line using Mass Spectrometry**

### **5.1.1 Quantitative comparison of HCPs identified between early (24hrs) and late (96hrs) in temperature-shifted and non-temperature-shifted cultures**

In a comparison between early and late stage culture conditioned media taken from a non-producing hybridoma cell line, 79 HCPs at 37°C (25 up-regulated, 54 down-regulated at 96hrs) and 65 HCPs at 31°C (31 up-regulated, 34 down-regulated at 96hrs) were identified as differentially-expressed following quantitative label-free LC/MS analysis (Section 3.1.3.1 and 3.1.3.2). Efforts were made to categorise differentially-expressed proteins by biological function e.g. protease, glycosidase, growth factor, apoptotic factor etc., and discuss trends in the up-regulation or down-regulation of these proteins as a group. Where no such trends were observed, individual proteins were selected for discussion based on their potential to impact culture growth or product quality.

Among those identified was the secreted protein Olfactomedin-like protein 3 (OLFML3) which was found to be 2.4-fold up-regulated at 96hrs compared to 24hrs at 31°C and down-regulated at 96hrs under normal temperature conditions (37°C). (Interestingly this protein was also identified as being up-regulated 4.1-fold at 120hrs under temperature-shift and 4.3-fold at 120hrs under non-temperature-shifted conditions in the CHO DP12 cell line). OLFML3 has been shown to inhibit tumour growth by impairing angiogenesis. It does this by binding the growth factor bone morphogenetic protein-4 (BMP4), in turn affecting the SMAD1/5/8 signalling pathway (Miljkovic-Licina et al. 2012). On the other hand the overexpression of OLFML3 in PC12 rat pheochromocytoma cells increased the growth rate of the cells (Lee and Tomarev 2007). Given role OLFML3 plays in regulating growth, it is possible that the secretion of this protein may impact on cell growth and / or viability recombinant protein producing cell lines. Although there is little known about the precise mechanisms by which this particular protein functions, its differential expression in both hybridoma and a CHO cell line under temperature-shifted and non-temperature-shift conditions possibly make it worthy of further follow up in order to define its precise role in recombinant protein producing cell lines.

In addition to OLFML3, a number of proteases were also identified as being differentially-expressed over time in culture. The serine peptidase Tripeptidyl-peptidase 2, known to cleave tripeptides from the N-terminus of oligopeptides (Usukura et al. 2013), was found to be 3.4-fold up-regulated at 96hrs compared to 24hrs in the 31°C culture.

The zinc metallopeptidase Thimet oligopeptidase was 11.5-fold up-regulated at 96hrs post temperature-shift compared to 24hrs at 37°C. Thimet oligopeptidase is ubiquitously expressed and, depending on the cell type, has been located both intracellularly (in the cytoplasm, nucleus, and membrane) and extracellularly (Ray et al. 2004). The primary function of Thimet oligopeptidase is the degradation of peptides released from proteasomes, specifically the metabolism of oligopeptides due to its ability to only cleave short peptides (Russo et al. 2012). It is therefore a protein of interest due to its proteolytic ability and the potential for it to contribute towards the degradation of recombinant therapeutic products.

In the non-temperature-shifted culture Cathepsin D was found to be 732.7-fold up-regulated in the late culture phase (96hrs) compared to the early phase (24hrs). Cathepsin D was a protease also identified in this thesis as being differentially-expressed in the cell culture supernatant of CHO cells and is discussed in greater detail in section 5.2.1.3. Another protease also identified as being differentially-expressed in the conditioned media of both hybridoma and CHO cell lines in this work was Proteasome subunit beta type-8. In the hybridoma cell line, Proteasome subunit beta type-8 increased in expression 3.2-fold at 96hrs compared to 24hrs at 37°C and was 2.9-fold down-regulated in the 31°C culture compared to the 37°C 96hrs post-temperature-shift. Similar to Cathepsin D, this protease is also discussed in further detail in section 5.1.1.3. Taken together, the up-regulation of these proteases at the later time point would indicate that recombinant products would be more susceptible to proteolytic activity.

### **5.1.2 Quantitative comparison of HCPs identified between Temperature-shifted and non-temperature-shifted culture at 24hrs and 96hrs post-temperature shift**

Following the quantitative label-free LC/MS analysis of conditioned media samples collected from a non-producing Hybridoma cell line grown under temperature-shifted and non-temperature-shifted conditions, a comparison between 31°C and 37°C cultures at 24hrs and 96hrs (post-temperature-shift) revealed the differential expression of 27 HCPs at 24hrs (23 down-regulated, 4 up-regulated at 31°C) and 44 HCPs at 96hrs (31 down-regulated, 13 up-regulated at 31°C) (Section 3.1.3.3 and 3.3.1.4).

One example of a protein that was found to be differentially-expressed was Calreticulin, which was down-regulated 2.4-fold at 31°C 24hrs post-temperature-shift. Calreticulin is perhaps best known for its role in the ER where it acts as a molecular chaperone for the proper folding of glycoproteins (Molinari et al. 2004). However, it has also been shown that Calreticulin can be expressed on the surface of apoptotic and pre-apoptotic cells as well as cancer cells (Chao et al. 2010). This is thought to promote the phagocytic uptake of such cells in the body (Raghavan et al. 2013). It is possible that the presence of Calreticulin on the cell surface may also be used for other signalling events, but the lack of a transmembrane domain means that the precise mechanism for signalling remains unclear (Gold et al. 2010). Interestingly, the presence of Calreticulin in the extracellular matrix has also been shown to influence the activity of MMP2 and MMP9 (Wu et al. 2007). More specifically, in Calreticulin-knockout Mouse embryonic fibroblast cells the activity of MMP9 significantly decreased while the activity of MMP2 increased. This was thought to be as a result of activation of the PI3 kinase/AKT dependent pathways and subsequent alteration of Ras/Raf/MEK activities (Wu et al. 2007).

Also of note was the differential expression of proteins that have the potential to impact on product quality, such as proteases. One such protease is Endoplasmic reticulum aminopeptidase 1 (ERAP1). Down-regulated at 31°C compared to 37°C both at 24hrs (4.4-fold) and 96hrs (2.1-fold) post-temperature-shift, ERAP1 is known to be localised to the ER, cytoplasm and cell surface (Cui et al. 2002; Saric et al. 2002). ERAP1 is known to alter the adaptive immune response to both host and

pathogen derived antigens by trimming peptide precursors in the ER to create or destroy antigenic epitopes (Aldhamen et al. 2013).

Another such protease identified was Cathepsin B. Although not identified as being differentially-expressed at 24hrs when temperature-shifted and non-temperature-shifted cultures are compared, it is however 27.8-fold down-regulated at 31°C when the two culture temperatures are compared 96hrs post-temperature-shift (Cathepsin B is discussed in further detail in section 5.1.1.3). The down-regulation of these proteases at 31°C would suggest that recombinant therapeutics are less likely to experience proteolytic activity in a temperature-shifted culture compared to a non-temperature-shifted culture. Reduced protease activity is therefore likely to contribute toward the increase in overall product yield often seen following the use of the temperature-shift strategy.

### **5.1.3 Qualitative assessment of the HCP profile of a non-producing hybridoma cell line.**

Overall the number of proteins identified in the cell culture supernatant of the 4/2D cells is in keeping with the numbers one would typically expect to see when profiling the secretome of a eukaryotic cell line (Piersma et al. 2010). When a qualitative comparison was made between the proteins identified under temperature-shift and non-temperature-shifted conditions it could clearly be seen that there were proteins specific to the two culture conditions. It should be pointed out that the viability of the 37°C culture was significantly lower than that of the 31°C and that some of the proteins identified as unique to 37°C may be present as a result of cell lysis.

However this makes the proteins identified as unique to 31°C even more interesting, because one would expect that the list of proteins identified in the low viability 37°C culture would also include all of those released from the more highly viable cells maintained at 31°C and so the only unique proteins found would be in the 37°C culture.



However despite the release of a whole myriad of intracellular proteins from the dead cells, there is still a significant number only identified in the 31°C culture. Although the use of monolayer flasks for the growth of hybridoma cells is not representative of what would be used in industry (Gong et al. 2006), the absence of any shear forces (further discussed in section 5.2.2) aids in the identification of changes in the HCP profile as a result of the physiological effect of the reduction of culture temperature on the cell itself.

## **5.2 Characterisation of HCP profiles from CHO-K1 and DP12 cell lines using Mass Spectrometry**

Proteomic profiling of CHO cells has been underway since the late 90's in an effort to identify the cellular components involved in the makeup of industrially desirable characteristics such as increased productivity and extended viability (Champion et al. 1999; Kaufmann et al. 1999). This is most likely due to the initial emphasis being placed on increasing product yield, which has happened, albeit largely from process and media optimisation (De Jesus and Wurm 2011; Hacker, De Jesus, Wurm 2009). As a result of this increase in product titre, coupled with increases in cell density, a bottleneck has been created in downstream processing with purification technologies struggling to keep pace with larger volumes of product (Thiel 2004; Thömmes and Etzel 2007). It has taken some time, however, for the same global profiling approaches to be applied to the characterisation of host cell protein components from cell culture supernatant (Grzeskowiak et al. 2009; Jin et al. 2010; Krawitz et al. 2006). This is not to say that the secretome of CHO cells was never investigated. Older studies, however, followed a much more targeted approach to identify specific proteases and glycosidases present in the media (Elliott, Hohmann, Spanos 2003; Gramer and Goochee 1993).

Global profiling of cell culture supernatant can provide a wealth of information on the impact of process changes on the HCP profile as well as identification of specific HCPs of interest such as proteases and growth factors (Baycin-Hizal et al. 2012; Grzeskowiak et al. 2009; Jin et al. 2010; Lim et al. 2013; Slade et al. 2012).

The HCP profiling presented in this work used quantitative label-free mass spectrometry to analyse the cell culture supernatant from a non-producing CHO-K1 and an IgG producing DP12 cell line, over different phases of the culture cycle from both temperature-shifted and non-temperature-shifted cultures. Of the 1,181 host cell proteins identified in this study, quantitative information was obtained on 790-fold. The quantitative information obtained was based on the analysis of both CHO-K1 and DP12 cell lines over time in culture at 37°C and 31°C, and between the two cell culture conditions in both early and late phases of the cell growth cycle. Among those identified were 138 known secreted proteins, 23 glycosidases, 58 proteases and 16 proteins involved in the regulation of apoptosis.

Normalisation by protein concentration is typically used during quantitative label-free LC/MS analysis to ensure accurate quantitation of differentially-expressed proteins identified. Given that this study is investigating the secretome of CHO cells another approach would be to normalise by cell number. This approach takes into account the significant differences in cell number between experimental groups and the variation in HCP concentration produced by those cells over the course of the culture. Applying a correction factor to the data did not result in a significant reduction in the number of proteins identified as differentially-expressed in seven of the eight comparisons (protein lists were reduced by roughly 5-10%). In one comparison however (DP12, 37°C vs 31°C at 144hrs), the protein list was reduced from 222 to 105. This was due to the significant number of proteins with relatively 'small' fold-changes (<2.09). While this correction factor was not applied to the data generated in this study, this result demonstrates a means by which further confidence could be added to lists of proteins identified as significantly differentially-expressed when analysing the secretome.

Qualitative analysis of this data revealed unique cohorts of HCPs were expressed in both cell lines under temperature-shifted and non-temperature-shifted conditions. Taken together this provides evidence that the choice of cell line and culture conditions used can have an impact on the HCP profile generated.

Many of the profiling studies published to date have used gel based technology (Grzeskowiak et al. 2009; Jin et al. 2010; Krawitz et al. 2006; Tait et al. 2012). It is only more recently that mass spectrometry based techniques have been used to characterise HCPs produced by CHO cells during culture (Doneanu et al. 2012; Lim et al. 2013; Slade et al. 2012). Quantitative label-free LC/MS has a number of advantages over gel based methods including facilitating the analysis of proteins with a broader range of characteristics, such as extremes in *pI*, size and hydrophobicity (Tscheliessnig et al. 2013). It is believed that the benefits inherent to quantitative label-free LC/MS enabled this study to characterise a wide spectrum of HCPs.

It should be noted that samples used for LC/MS analysis in this study were obtained from shake flask cultures. Although some studies have also used this means of culturing cells, many have utilised stirred tank bioreactors (Grzeskowiak et al. 2009; Jin et al. 2010; Lim et al. 2013; Slade et al. 2012; Tait et al. 2012).

While a significant proportion of the changes in the HCP profile has been attributed to reduction in culture viability, the use of different cell culture platforms may also contribute toward any differences seen given that hydro-mechanical forces in a stirred tank bioreactor are approximately 10 times higher than in a shake flask (Peter, Suzuki, Büchs 2006; Tait et al. 2012). The use of stirred tank bioreactors would potentially negate any differences in the HCP profile of cells under temperature-shift conditions due to cellular lysis through mechanical agitation experienced in the bioreactor (Tait et al. 2012; Velez-Suberbie et al. 2012). The low shear stress nature of shake flasks allowed this investigation to identify changes in the HCP profile of two CHO cell lines as they grew over time in culture under both temperature-shifted and non-temperature-shifted conditions, leading to the identification of both secreted and intracellular proteins in the cell culture supernatant.

In a proof of concept investigation, metabolic labelling was used to enable the enrichment of secreted proteins from two CHO cell lines (Slade et al. 2012). While this application is interesting from the point of view of identifying only those proteins secreted by CHO cells under different conditions, it does not consider the release of intracellular proteins into the media.

Such proteins could have a significant impact on product quality given that the intracellular proteome comprises of a large number of proteases and glycosidases (Gramer and Goochee 1993; Sandberg et al. 2006). While information on the HCP profile of CHO cells gathered from shake flask cultures may not be of direct relevance for stirred tank cultures, it is likely to be more directly comparable with culture conditions seen in disposable culture systems such as wave bags due to the lower shear stress environment they provide (Öncül et al. 2010). This investigation has identified differences in the HCP profile of two CHO cell lines and the impact of reduced culture temperature conditions on the HCP population. With the use of wave bag technologies set to increase, proteomic profiling of HCPs produced by CHO cells under different conditions provides valuable information that may be used in the selection of CHO cell lines, culture conditions and when to harvest (Shukla and Gottschalk 2012; Tait et al. 2012).

### **5.2.1 Quantitative Label-free LC/MS analysis of conditioned media from CHO-K1 and DP12 cell lines**

The use of a label-free LC/MS approach enabled quantitative information on the HCPs identified to be obtained. This section discusses some of the proteins identified using this approach.

#### **5.2.1.1 Growth factors**

Although Biopharmaceutical companies have turned away from the use of fetal bovine serum (FBS) due to regulatory concerns for quite some time, chemically defined protein free media has typically relied on amino acids, lipids and L-glutamine as nutritional sources for cell survival (Even, Sandusky, Barnard 2006).

While there have been several studies in hybridoma, NS0 and insect cell lines to identify autocrine factors that may support growth, this has only recently been successfully conducted in CHO cells (Dutton, Scharer, Moo-Young 1999; Eriksson et al. 2005; Lim et al. 2013; Spens and Häggström 2005; Spens and Häggström 2009).

Given the complex nature of extracellular signalling, this section looks at growth factors identified in this study on an individual basis. A summary table (Table 5.2.1) provides an overview of some of the growth factors identified as differentially-expressed in the cell culture media of CHO-K1 and DP12 cell lines. Vascular endothelial growth factor C (VEFGC), Vascular endothelial growth factor A (VEFGA) and Hepatoma-derived growth factor (HGDF) were identified as differentially-expressed in the conditioned media of the IgG producing DP12 cell line (summary Table 5.2.1). Although not identified as being differentially-expressed, qualitative searches of LC/MS data revealed these proteins to be present in the culture supernatant of the non-producing CHO-K1 cell line. The incorporation of growth factors, including VEFGC, into culture media has been shown to increase cell growth in CHO-K1 cells (Lim et al. 2013).

It is worth noting that VEGFA and VEGFC become down-regulated at 31°C when temperature-shift and non-temperature-shift cultures are compared at 144hrs in the DP12 cell line. This down-regulation is interesting given that cell growth was almost completely ceased in the 31°C culture. The lack of differential information on VEGFA and VEGFC in the CHO-K1 cell line indicates that these proteins continued to be secreted by the cells. This could have led to the continued growth of these CHO-K1 cells under temperature-shift conditions albeit at a reduced rate. While it has been shown that the cell cycle arrest of CHO cells under temperature-shift occurs via the ATR-p53-p21 pathway (Roobol et al. 2011), it does raise the question; is the reduced secretion of extracellular proteins (such as VEGFA and VEGFC) influencing cell growth under temperature-shift conditions? Further investigations would be required to definitively answer this question, future proteomic studies could focus on such growth factor signalling pathways as an alternative means of regulating cell growth in the bioreactor e.g. through the addition of these proteins to the media.

**Table 5.2.1** Summary table of differentially-expressed growth factors identified following quantitative label-free LC/MS analysis of conditioned media from CHO-K1 and DP12 cells.

Growth factor	Cell line	Comparison	Fold-change	Direction of fold-change	Comments
Vascular endothelial growth factor A	DP12	24hrs vs 144hrs at 37°C	3.7	Up at 144hrs	VEGFA binds Vascular endothelial growth factor receptor (VEGFR)-1 and VEGFR2 while VEGFC binds VEGFR2 and VEGFR3 (Tammela et al. 2005). In colorectal cancer cells signalling via VEGF1 involves activation of ERK1/2 and SAP/c-JNK and the Src family kinases (Fan et al. 2005; Lesslie et al. 2006). Although little is known of the downstream elements involved in VEGFR2 and VEGFR3 signalling pathways, it does appear that Src and AKT play a central role and that site-specific autophosphorylation of the receptors is essential for their activation (Kowanetz and Ferrara 2006).
	DP12	37°C vs 31°C at 24hrs	10.8	Down at 31°C	
	DP12	37°C vs 31°C at 144hrs	4.8	Down at 31°C	
Vascular endothelial growth factor C	DP12	24hrs vs 144hrs at 37°C	3.1	Down at 144hrs	
	DP12	37°C vs 31°C at 24hrs	2.4	Down at 31°C	
Hepatoma-derived growth factor	DP12	24hrs vs 144hrs at 31°C	229.7	Up at 144hrs	HDGF is an acidic heparin-binding growth factor that is ubiquitously expressed in mammalian tissue and has been shown to stimulate the growth of a variety of cells including vascular smooth muscle, endothelial and hepatoma cells (Li et al. 2013). The means by which HDGF increases proliferation varies depending on the cell line. It has been proposed that HDGF may regulate tumorigenicity in breast cancer cells through the modulation of E-cadherin and vimentin expression (Chen et al. 2012b). This is unlikely to be the case in CHO cells however as they do not express E-cadherin (Suriano et al. 2003). In gastric cancer AGS cells, HDGF was found to activate Erk1/2 via a Ras-independent pathway (Mao et al. 2008). Treatment of Keloid fibroblasts (KF) with recombinant HDGF was found to increase their proliferation. It was reported that the addition of HDGF to the media of KFs caused activation of the ERK pathway which resulted in the up-regulation of the secretion of vascular endothelial growth factor (VEGF) (Ooi et al. 2010).

Table 5.2.1 continued.

Growth factor	Cell line	Comparison	Fold-change	Direction of fold-change	Comments
Transforming Growth Factor $\beta$ 1 (TGF $\beta$ 1)	CHO-K1	24hrs vs 120hrs at 37°C	16.5	Up at 120hrs	TGF $\beta$ 1 is highly conserved across mammalian species and its secretion by CHO-K1 cells has previously been reported, however this report merely confirmed its existence in CHO and did not infer any specific biological function (Beatson et al. 2011; Lu et al. 2012). TGF $\beta$ 1 has been implicated in a wide range of biological processes including growth, differentiation, apoptosis and inhibition of cell proliferation (Kubiczkova et al. 2012). It is only through dissociation from its binding partner latent TGF $\beta$ -binding protein (LTBP) that TGF $\beta$ 1 can bind to TGF $\beta$ 1 receptor type II (TGF $\beta$ RII) (Kubiczkova et al. 2012; Verrecchia and Mauviel 2007). This causes the phosphorylation of TGF $\beta$ 1 receptor type I, which results in the activation of smad1, smad5 and smad8 or smad2 and smad3 and their subsequent translocation to the nucleus to regulate transcription (Klass, Grobbelaar, Rolfe 2009). TGF $\beta$ 1 activation may occur through activation with plasmin, thrombin or the matrix metalloproteinases MMP2 and MMP9 (Annes, Munger, Rifkin 2003). While MMP9 was identified as being up-regulated at 37°C compared to 31°C after 120hrs post temperature-shift, it should be pointed out that the integrin $\alpha$ <sub>v</sub> $\beta$ <sub>3</sub> is also required for activate latent TGF $\beta$ 1 (Mu et al. 2002).
	CHO-K1	24hrs vs 120hrs at 31°C	1.9	Down at 120hrs	
	CHO-K1	37°C vs 31°C at 120hrs	4.5	Down at 31°C	
TGF $\beta$ -binding protein (LTBP)	CHO-K1	37°C vs 31°C at 120hrs	43	Down at 31°C	



### 5.2.1.2 Apoptosis

A number of proteins involved in the regulation of apoptosis were also identified as differentially regulated in the conditioned media of CHO-K1 and DP12 cell lines.

In both cell lines, these proteins were largely identified as being up-regulated at the later time points compared to the earlier time points in cultures grown under temperature-shift and non-temperature-shift conditions. This result is somewhat unsurprising as culture viability begins to decline at the later time points, whether through nutrient deprivation or a build-up of waste by-products in the media, the expression of many of these proteins become increased thereby inducing apoptotic signalling pathways (Wei et al. 2011). Intracellular in nature, the presence of apoptotic proteins in the cell culture media is almost certainly as a result of their release from dead cells. However this many not entirely account for all the proteins as some are also secreted e.g., Clusterin. In addition, biological function of each protein identified in this study was inferred through manual searches of the Uniprot protein database. However this database does not necessarily provide information on whether the protein is pro or anti-apoptotic. This is understandable as many of the proteins identified are multi-function proteins and have been found to play either pro or anti-apoptotic roles depending on the means by which apoptosis is induced in the original study.

For example Guanine nucleotide-binding protein subunit beta-2-like 1 (GNB2L1), also commonly referred to in the literature as Receptor for activated C kinase (RACK1), is a scaffold protein that is known to interact with over 100 proteins, as such it is involved in a diverse range of functions within the cell such as apoptosis, translation and cellular organisation (Gandin et al. 2013). In terms of apoptosis, there are a number of mechanisms by which GNB2L1 has been shown to operate. There have been contradicting reports as to whether GNB2L1 promotes or inhibits apoptosis through its interaction with Bax, Bcl-2 and Bcl-XL. In HeLa and HT-29 cell lines, as well as disassociating the interaction of Bax with Bcl-XL, GNB2L1 promotes Bax oligomerization with Bcl-XL thus inducing apoptosis (Wu et al. 2010a). In human colon cells GNB2L1 inhibits the expression of anti-apoptotic Bcl-2 and Bcl-XL and translocates pro-apoptotic proteins Bax and Bim to the mitochondria (Mamidipudi and Cartwright 2009), thereby causing the mitochondrial membrane to becoming porous allowing the release of Cytochrome c, and the

activation of Caspase 3 and 9 resulting in cell death. The overexpression of GNB2L1 in human hepatocellular carcinoma cell lines resulted in the upregulation of Bcl-2, as well as other growth and survival promoting proteins such as upregulation of cyclin D1, MYC and Survivin (Ruan et al. 2012). While the role of GNB2L1 in apoptosis appears to be cell line specific, its involvement in the regulation of Bcl-2 is interesting given that the overexpression of Bcl-2 has been used to generate ‘apoptosis-resistant’ CHO cell lines (Kim et al. 2012; Reynolds et al. 1996). Another protein identified in this work that is also known to have both pro and anti-apoptotic functionality is Clusterin. The function and mechanisms of Clusterin are further discussed in section 5.2.2.2.

At the later time point in CHO-K1 and DP12 cell lines, when 37°C and 31°C cultures were directly compared, the number of up-regulated and down-regulated apoptotic proteins were evenly divided between the two culture conditions. As such, no firm conclusion can be drawn as to the impact of the presence of these proteins in the culture media. Whether the secretion of anti-apoptotic proteins helped extend culture viability at 31°C or the secretion of pro-apoptotic proteins lead to the widespread induction of apoptosis in the 37°C culture is unclear from this analysis. However a panel of proteins known to be involved in the regulation of apoptosis has been identified through this work. Future work could focus on these proteins to determine if their secretion impacts on culture viability and whether these should be considered for future cell engineering through the knockdown or overexpression of such proteins.

### **5.2.1.3 Proteases**

In addition to the potential for HCPs positively or negatively impacting on cell growth and viability, HCPs have also been reported to affect both product yield and product quality (Eon-Duval, Broly, Gleixner 2012). The existence of proteases in the cell culture media of a number of bioprocess relevant cell lines has including CHO, hybridoma and NS0 has been known for some time (Elliott, Hohmann, Spanos 2003; Gao et al. 2011a; Karl, Donovan, Flickinger 1990; Sandberg et al. 2006; Spens and Häggström 2005; Van Erp et al. 1991).

It is only in more recently published data however that the activity of these enzymes negatively impacting on product quality has been reported; this includes the degradation of an Fc-fusion recombinant protein by a protease homologous to Cathepsin D (Robert et al. 2009), N-terminal clipping of a fusion protein by a serine-threonine class of protease (Dorai et al. 2011), C-terminal lysine variants on a monoclonal antibody as a result of carboxypeptidase B activity (Luo et al. 2012) and the degradation of a highly purified monoclonal antibody product due to residual protease activity caused by an acidic protease (Gao et al. 2011a). This investigation identified 58 differentially-expressed proteases in the cell culture supernatant of a CHO-K1 and DP12 cell line as shown in Table 5.2.2.

In both cell lines proteases were typically found to increase in abundance at the later stage in culture relative to the early stage culture in both temperature-shift and non-temperature-shift cultures. While this may be partly due to the release of intracellular proteins from dead cells as cell viability reduces towards the end of the culture. In addition, many of the proteases identified are secreted. Therefore, such proteases would not only be released from the intracellular components of dead cells, but they would also accumulate in the media as a result of their secretion by viable cells. This would indicate that recombinant therapeutics would be vulnerable to proteolytic activity during cell culture operations, even where culture viability remains high. Among the secreted proteases identified were the endopeptidases Cathepsin D, Cathepsin B, MMP9 and MMP19 (summarised in Table 5.2.3).

Gene Symbol	Protease	CHO-K1	DP12
ADAMTS1	A disintegrin and metalloproteinase with thrombospondin motifs 1	x	x
ADAMTS7	A disintegrin and metalloproteinase with thrombospondin motifs 7-like	x	
BMP1	Bone morphogenetic protein 1	x	x
PRSS22	brain-specific serine protease 4-like	x	x
P15156	Calcium-dependent serine proteinase	x	x
P15156	calcium-dependent serine proteinase-like		x
CTSB	Cathepsin B		x
CTSB	cathepsin B-like	x	x
CTSD	Cathepsin D	x	x
CTSH	Cathepsin H	x	
CTSL1	Cathepsin L1		x
CTSL1	cathepsin L1-like	x	x
CTSZ	Cathepsin Z	x	x
C1ra	Complement C1r-A subcomponent	x	x
LAP3	Cytosol aminopeptidase		x
CNDP2	Cytosolic non-specific dipeptidase	x	x
DPP7	Dipeptidyl peptidase 2	x	x
DPP7	dipeptidyl peptidase 2-like		x
DPP3	Dipeptidyl peptidase 3	x	
ADAM10	Disintegrin and metalloproteinase domain-containing protein 10	x	x
ADAM15	Disintegrin and metalloproteinase domain-containing protein 15	x	
ADAM17	Disintegrin and metalloproteinase domain-containing protein 17	x	x
ADAM19	Disintegrin and metalloproteinase domain-containing protein 19	x	
ADAM9	disintegrin and metalloproteinase domain-containing protein 9		x
EIF3F	Eukaryotic translation initiation factor 3 subunit F	x	
LTA4H	leukotriene A-4 hydrolase	x	
CTSA	Lysosomal protective protein	x	x
PRCP	lysosomal Pro-X carboxypeptidase-like		x
MMP19	Matrix metalloproteinase-19	x	x
MMP9	matrix metalloproteinase-9	x	x
ACIA	N(4)-(beta-N-acetylglucosaminyl)-L-asparaginase		x
CTSH	pro-cathepsin H-like		x
PRBP	Prolyl endopeptidase	x	
PSMA1	Proteasome subunit alpha type-1		x
PSMA3	Proteasome subunit alpha type-3	x	x
PSMA4	Proteasome subunit alpha type-4	x	
PSMA5	Proteasome subunit alpha type-5		x
PSMA6	Proteasome subunit alpha type-6		x
PSMA7	Proteasome subunit alpha type-7		x
PSMB1	Proteasome subunit beta type-1	x	x
PSMB2	Proteasome subunit beta type-2		x
PSMB3	Proteasome subunit beta type-3		x
PSMB4	Proteasome subunit beta type-4	x	
PSMB6	Proteasome subunit beta type-6	x	x
PSMB6	proteasome subunit beta type-6-like	x	
PSMB8	Proteasome subunit beta type-8		x
PARK7	Protein DJ-1		x
SCPPP1	Retinoid-inducible serine carboxypeptidase		x
HTRA1	Serine protease HTRA1	x	x
HTRA1	serine protease HTRA1-like		x
THOP1	Thimet oligopeptidase		x
THOP1	thimet oligopeptidase-like	x	x
PLAT	Tissue-type plasminogen activator	x	x
TPP1	Tripeptidyl-peptidase 1	x	x
TPP1	tripeptidyl-peptidase 1-like		x
UCHL3	Ubiquitin carboxyl-terminal hydrolase isozyme L3	x	
Otub1	Ubiquitin fibronectase OTUB1	x	
XPMPHP1	zinc-Pro aminopeptidase 1	x	

**Table 5.2.2** List of differentially-expressed proteases identified in the cell culture supernatant of a CHO-K1 and DP12 cell line. (Protein function was obtained following a search of the protein names against the Uniprot protein database.)

**Table 5.2.3** Summary table of differentially-expressed proteases identified following quantitative label-free LC/MS analysis of conditioned media from CHO-K1 and DP12 cells.

Protease	Cell line	Comparison	Fold-change	Direction of fold-change	Comments
BMP1	CHO-K1	24hrs vs 144hrs at 37°C	12.5	Up at 120hrs	Unlike all other BMPs, BMP-1 is not a TGFβ-like protein (Hopkins, Keles, Greenspan 2007). Instead, it possesses the ability to hydrolyse a wide range of ECM substrates including chondrin, protosyl oxidase, biglycan, osteoglycin, procollagen I-III, V, VII, XI, laminin-5 (γ2 chain) and several members of the TGFβ superfamily (von Marschall and Fisher 2010). Pro-BMP-1 comprises of a signal peptide, an amino-terminal prodomain, a conserved catalytic domain, similar to that of the astacin M12A family of metzincin metalloproteinases, three CUB (complement-negf-BMP1) domains, and an epidermal growth factor (EGF)-like domain (von Marschall and Fisher 2010).
	CHO-K1	37°C vs 31°C at 144hrs	2.3	Down at 31°C	
	DP12	24hrs vs 144hrs at 37°C	2.7	Up at 144hrs	
	DP12	37°C vs 31°C at 144hrs	3.0	Down at 31°C	
MMP9	CHO-K1	24hrs vs 120hrs at 37°C	3.3	Down at 120hrs	MMP-9 has a wide range of substrates, the three main groups include; (1) a Pro-X-X-Hy-(Ser/Thr) consensus sequence (where X represents any residue and Hy is a hydrophobic residue). (2) A Gly-Leu-(Lys/Arg) motif and (3) two subsequent Arg residues (Arg-Arg) (Kidel et al. 2001).
	CHO-K1	37°C vs 31°C at 120hrs	1520	Down at 31°C	
	DP12	24hrs vs 144hrs at 37°C	7.6	Down at 144hrs	
	DP12	24hrs vs 144hrs at 31°C	4.1	Down at 144hrs	
	DP12	37°C vs 31°C at 144hrs	1.7	Down at 31°C	
MMP19	CHO-K1	37°C vs 31°C at 120hrs	4.0	Down at 31°C	MMP-19 is involved in the degradation of various components of the ECM including collagen IV, gelatin, laminin-1, nidogen-1 and fibronectin (Mysliwy et al. 2006). While MMP-19 differentiates itself structurally from other subfamilies of MMPs with the presence of five Glu residues within the linker region and an additional Cys residue in the catalytic domain it maintains its ability to hydrolyse general MMP substrates (Stracke et al. 2000). This indicates that, given its activity, MMP-19 has the potential to adversely affect product quality and/or yield in a similar manner to MMP-9.
	DP12	24hrs vs 144hrs at 31°C	3.3	Down at 144hrs	
	DP12	37°C vs 31°C at 144hrs	1.8	Down at 31°C	

Table 5.2.3 continued.

Protease	Cell line	Comparison	Fold-change	Direction of-fold-change	Comments
Cathepsin D	CHO-K1	24hrs vs 120hrs at 31°C	3.1	Up at 120hrs	A homolog of Cathepsin D has previously been reported as being responsible for the clipping of a recombinant Fc fusion protein (Robert et al. 2009). In this instance, the source of the protease was traced back to viable CHO cells. Interestingly, the 31°C culture in which this was identified as being differentially-expressed was also highly viable. Under normal biological conditions Cathepsin D is typically found intracellularly, though it has been shown that it is secreted by many cancer cell lines (Masson et al. 2010). One possible reason for its secretion is the defective acidification of the endosomal compartments disrupts normal lysosomal sorting by reducing the amount of the Golgi associated sorting receptor, CI-MPR, available for Cathepsin D processing (Kokkonen et al. 2004). This is entirely plausible given that the intracellular pH of temperature-shifted CHO cells is known to reduce as a result of a change in the permeability of the plasma membrane (Stapalioinis, Koli, Deutscher 1997).
	CHO-K1	37°C vs 31°C at 120hrs	3.1	Up at 31°C	
	DP12	37°C vs 31°C at 24hrs	2.7	Down at 31°C	
Cathepsin B	CHO-K1	24hrs vs 120hrs at 31°C	2.3	Up at 120hrs	One of eleven cysteine Cathepsins, Cathepsin B functions as an exopeptidase and an endopeptidase with dual specificity for phenylalanine or glycine residues (Comas and Simon 2010). While many Cathepsins only become active under low pH conditions, Cathepsin B can maintain its activity at near neutral pH levels. This would suggest that it would be capable of clipping recombinant protein products under normal cell culture conditions (Strojnik et al. 2005; Zavašnik-Bergant and Turk 2006). Interestingly the proteolytic activity of Cathepsin B has been linked to the activation of the Matrix metalloproteinase-9 (MMP9) (Gondi and Rao 2013).
	DP12	24hrs vs 144hrs at 37°C	1.6	Up at 144hrs	
	DP12	24hrs vs 144hrs at 31°C	4.3	Up at 144hrs	
	DP12	37°C vs 31°C at 24hrs	3.7	Down at 31°C	

Cell engineering strategies could provide a means of alleviating proteolytic degradation of the drug substance. The presence of MMP9 in the cell culture supernatant of a CHO-K1 cell line has been previously reported and it has been noted that its presence can significantly reduce the overall product yield of a recombinant therapeutic susceptible to its action (Elliott, Hohmann, Spanos 2003).

Interestingly, an inhibitor of MMP9, Extracellular matrix protein 1 (ECM1) was also identified in the media of CHO-K1 and DP12 cell lines in my study.. The proteolytic activity of MMP9 is at least partially inhibited through the interaction of the metalloproteinase with the C-terminal second tandem repeat of ECM1. Complete inhibition may be achieved through the expression of PTMs on the peptide fragments of ECM1 (Fujimoto et al. 2006). Nonetheless, the presence of this endogenously expressed MMP inhibitor could be investigated further to ascertain whether it has the same protease inhibiting properties in CHO and, if so, could it be used to alleviate product degradation caused by MMP9 e.g. possibly through overexpression of the ECM1 or its C-terminal second tandem repeat domain coupled with the 19 amino acid signal peptide to facilitate its secretion by the endoplasmic/Golgi-dependent pathway (Merregaert et al. 2010). However care must be taken when selecting a strategy to protect the protein product from degradation. Although engineering of the protein product to make it resistant to cleavage may be an attractive alternative (Kinder et al. 2013), this may not always be feasible as altering the protein structure may impair its activity (Strome, Sausville, Mann 2007).

Given that many proteases also operate intracellularly (Wiera et al. 2012), simply knocking out the protein may have a negative impact on cell function. In some circumstances however, future strategies could consider engineering of the protease itself to alter its activity. For example, the zinc dependant metalloproteinase, Bone-morphogenetic-protein 1 (BMP-1) (summary Table 5.2.3) has CUB and EGF-like domains that regulates the specificity with which it cleaves proteins and/or cleavage sites (Hintze and AUFENNEC 2007). For example, the absence of these domains resulted in complete degradation of a truncated form of procollagen VII by the BMP-1 catalytic domain whereas the full length enzyme produced only two fragments of the substrate.

It has also been reported that if the CUB domain lacks two of its N-linked glycans (Asn<sup>332</sup> and Asn<sup>599</sup>, but not Asn<sup>363</sup>) BMP-1 appears to exhibit a slower rate of cleavage (Garrigue-Antar, Hartigan, Kadler 2002), while the removal of all three glycans results in the translocation of BMP-1 to the proteasome for degradation.

While the inhibition of proteases in the culture media may appear to be an obvious alternative, it should be borne in mind that many growth factors require some form of cleavage in order for them to function correctly (Khatib and Geraldine 2006). Therefore the effect of protease inhibition on cell growth and viability may also need to be assessed prior to its implementation as a means of preventing product degradation. In addition, the use of protease inhibitors would be costly for large scale operations and their removal from the final drug substance must also be considered.

When temperature-shift and non-temperature-shift cultures were compared in late stage culture of the CHO-K1 cell line it was observed that the majority of proteases were down-regulated at 31°C. However, when the same comparison is made in the DP12 cell line 4 proteases were up-regulated and 5 were down-regulated at 31°C. Interestingly, despite the majority of proteases being down-regulated, Cathepsin D (known to negatively impact product quality as discussed previously) was up-regulated at 31°C when 37°C and 31°C cultures were compared at 120hrs in CHO-K1 cells but was not found to be differentially-expressed under the same conditions in the DP12 cell line. This shows that the effect of culture conditions on protease expression can be cell line dependent and that identifying a cell line and culture conditions suitable for the production of a product sensitive to protease activity would need to be made on a case-by-case basis.

#### **5.2.1.4 Reductases**

Disulfide bonds are critical for imparting stability and structure to a recombinant protein molecule (Liu and May 2012). Reduction of these bonds as a result of the introduction of intracellular components to the harvested cell culture liquid from cells lysed during the harvest step has recently been reported (Trexler-Schmidt et al. 2010).



Investigations to identify the component responsible for the reduction of antibody disulfide bonds in CHO cell culture revealed TXN1 to be the terminal enzyme involved (Koterba, Borgschulte, Laird 2012). In the DP12 cell line, TXN1 was 2.9-fold up-regulated at 144hrs compared to 24hrs at 31°C and 5.6-fold up-regulated at 144hrs compared to 24hrs at 37°C. The greater upregulation of TXN1 37°C compared to 31°C is likely to be as a result of the reduction in culture viability in the non-temperature-shift culture compared to the temperature-shift culture at 144hrs. This result indicates that a recombinant therapeutic would be more likely experience disulfide bond reduction over a prolonged period in the cell culture supernatant. Strategies to prevent antibody reduction such as the addition of CuSO<sub>4</sub>, EDTA, lowering pH and air sparging of the harvested cell culture liquid have all been proposed (Trexler-Schmidt et al. 2010). Although many of these strategies could not be employed during cell culture, information on the differential-regulation of reductases could help manufacturers make informed decisions when selecting culture conditions and harvest times.

#### **5.2.1.5 Glycosidases**

The presence of a sugar moiety on a protein, called glycosylation, is among one of the most common PTMs observed in proteomics (Segu and Mechref 2010). The glycosylation pattern of a recombinant therapeutic can affect the protein molecule in a variety of ways (summarised Table 5.2.4); for example, one particularly important glycan group is situated at arginine 297 in the Fc region of a recombinant therapeutic antibody (Jefferis 2005), the presence of this oligosaccharide is required for Fc-receptor mediated functions such as antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (Kim et al. 2010). The removal of carbohydrate side chains may not only impact on product functionality, but may also leave regions of the protein molecule exposed to further proteolytic degradation (Waszkiewicz et al. 2013). In addition, the engineering of specific protein glycans has been shown to enable targeted delivery of the therapeutic to the desired tissue in the body, thus increasing product efficacy (Solá and Griebenow 2010).

Role/effect	Comment
<b>Protein folding</b>	Glycosylation can affect local protein secondary structure and help direct folding of the polypeptide chain.
<b>Protein targeting/trafficking</b>	The glycocomponent can participate in the sorting/directing of a protein to its final destination.
<b>Ligand recognition/binding</b>	The carbohydrate content of antibodies, for example, function in antibody binding to monocyte Fc receptors and interaction with complement component C1 <sub>q</sub> .
<b>Biological activity</b>	The carbohydrate side chain of gonadotrophins (specifically the $\alpha$ -subunit N <sup>S2</sup> side chain) is essential to the activation of gonadotrophin signal transduction.
<b>Stability</b>	Sugar side chains can potentially stabilize a glycoprotein in a number of ways including enhancing its solubility, shielding hydrophobic patches on its surface, protecting from proteolysis and directing participation in intrachain stabilizing interactions.
<b>Regulates protein half-life</b>	Large amounts of sialic acid can increase a glycoprotein's plasma half-life. Exposure of galactose residues can decrease plasma half-life by promoting uptake through hepatic galactose residues.

**Table 5.2.4** Table summarising the potential role and effect of the glycocomponent of glycoproteins (adapted from (Walsh and Jefferis 2006)).

While the cleavage of recombinant protein as a result of proteolytic activity has been widely reported in literature, little has been presented in the way of the potential degradation of important PTMs, such as glycosylation or its associated groups, during the culture process. The study presented in this thesis has identified the differential expression of a number of glycosidases in the cell culture media of a non-producing CHO-K1 and an IgG producing DP12 cell line. In this section a brief overview of the mechanisms of action of some of the glycosidases is provided. A table of the 23 differentially-expressed glycosidases identified in the cell culture supernatant of a CHO-K1 and DP12 cell line are displayed in Table 5.2.5 and a more detailed overview of a selection of these is provided in Table 5.2.6. In general, glycosidases were identified as being up-regulated at the later time points compared to the earlier time points in cultures grown at 37°C and 31°C in both cell lines. Given that many of the glycosidases are not known to be secreted, their presence is almost certainly as a result of their release from dead cells. Therefore, reduced culture viability and extended periods of time in culture are two scenarios most likely to negatively impact the quality of recombinant therapeutics susceptible to glycosidase activity.

Gene Symbol	Glycosidases	CHO-K1	DP12
GLA	Alpha-galactosidase A		X
IDUA	Alpha-L-iduronidase	X	X
MAN2A1	Alpha-mannosidase 2		X
NAGA	Alpha-N-acetylgalactosaminidase	X	X
NAGA	alpha-N-acetylgalactosaminidase-like	X	X
NAGLU	alpha-N-acetylglucosaminidase-like	X	X
GLB1	Beta-galactosidase	X	X
GUSB	Beta-glucuronidase		X
HEXA	beta-hexosaminidase subunit alpha	X	X
HEXB	Beta-hexosaminidase subunit beta		X
CTBS	Di-N-acetylchitobiase	X	
MAN2B2	Epididymis-specific alpha-mannosidase	X	
GBA	glucosylceramidase	X	X
GAA	Lysosomal alpha-glucosidase	X	X
GAA	lysosomal alpha-glucosidase-like	X	X
MAN2B1	lysosomal alpha-mannosidase		X
MAN1A1	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA		X
MAN1A1	mannosyl-oligosaccharide 1,2-alpha-mannosidase IA-like,		X
GANAB	Neutral alpha-glucosidase AB	X	X
GANAB	neutral alpha-glucosidase AB isoform 1	X	X
NEU1	Sialidase-I	X	X
SMPD1	Sphingomyelin phosphodiesterase	X	X
FUCA1	tissue alpha-L-fucosidase	X	X

**Table 5.2.5** List of 23 differentially-expressed glycosidases identified in the cell culture supernatant of a CHO-K1 and DP12 cell line. (Protein function was obtained following a search of the protein names against the Uniprot protein database.)

Interestingly, when temperature-shift and non-temperature-shift cultures were compared at the later time point in the respective cell lines, the numbers of up-regulated and down-regulated glycosidases are evenly split between the two culture conditions i.e. 3 up-regulated, 4 down-regulated at 31°C in CHO-K1, 5 up-regulated, 4 down-regulated at 31°C in DP12. As such, the selection of culture conditions for the production of a glycosidase sensitive product should be made on a case-by-case basis.

While the bioprocessing community has primarily focused on the presence of N-linked glycan's, the importance of O-glycosidic linkages on future recombinant products should not be discounted. Although currently not readily predicted, the presence of such O-glycans can modulate important therapeutic protein characteristics such as aggregation, protein stability and protease resistance (Zhong and Somers 2012). Given their ability to cleave such O-glycosidic linkages, glycosidases that may negatively impact on the quality of such future products include  $\alpha$ -N-acetylgalactosaminidase, Hex A and Hex B.

The potential for Hex A and Hex B to enzymatically degrade recombinant glycoproteins has previously been reported in a number of insect cell lines, however our study is the first report of their presence in the culture media of CHO cells (Licari, Jarvis, Bailey 1993). Cell engineering strategies have been previously used to modify glycosidases to improve the recombinant protein characteristics (Solá and Griebenow 2010). Similar engineering approaches could be used to down regulate or eliminate the expression of glycosidases that diminish product quality.

Similar to proteases, the presence of glycosidases should not only be considered during cell culture but also elsewhere during production. For example, the lysosomal  $\alpha$ -glucosidase (summary Table 5.2.6) is capable of enzymatic activity under acidic conditions. This indicates the potential for this glycosidase to negatively impact on product quality during downstream purification of the therapeutic protein where it may be exposed to buffers of low pH (e.g. Protein A elution) if it is not removed prior to such steps.

The information obtained from this study on the glycosidases identified provides important information for the potential degradation of the protein product. The presence of such glycosidases in the cell culture media should be considered in future cell engineering and product recovery campaigns.

**Table 5.2.6** Summary of differentially-expressed glycosidases identified following quantitative label-free LC/MS analysis of conditioned media from CHO-K1 and DP12 cells.

Glycosidase	Cell line	Comparison	Fold-change	Direction of fold-change	Comments
<b><math>\alpha</math>-glucosidase</b>	CHO-K1	24hrs vs 120hrs at 31°C	2.2	Up at 120hrs	Essential for the degradation of glycogen (Hoeftloot et al. 1988), $\alpha$ -glucosidase cleaves it's substrate at both $\alpha$ -1,4 and $\alpha$ -1,6 glycosidic linkages (Douillard-Guilhoux et al. 2009). The substrate specificity of $\alpha$ -glucosidase is known to depend on the species of origin e.g. microorganism, plant or animal (Chiba 1997). While glycogen is not a PTM on any recombinant therapeutic protein, the effect of its enzymatic activity on such products is worthy of further investigation to determine if it is capable of negatively impacting on product quality.
	DP12	24hrs vs 144hrs at 37°C	2.9	Up at 144hrs	
	DP12	24hrs vs 144hrs at 31°C	2.4	Up at 144hrs	
	DP12	37°C vs 31°C at 24hrs	2.8	Down at 31°C	
	DP12	37°C vs 31°C at 144hrs	2.4	Down at 31°C	
<b><math>\alpha</math>-N-acetyl-galactosaminidase</b>	CHO-K1	24hrs vs 120hrs at 37°C	2.4	Down at 120hrs	Lysosomal glycosidase which hydrolyses the O-glycosidic linkage between terminal alpha-N-acetylgalactosamine moiety and serine or threonine residues (Mohamad et al. 2002).
	CHO-K1	24hrs vs 120hrs at 31°C	595.8	Up at 120hrs	
	CHO-K1	37°C vs 31°C at 120hrs	2.8	Up at 31°C	
	DP12	24hrs vs 120hrs at 31°C	1.8	Up at 144hrs	
<b>beta-hexo-saminidase subunit alpha</b>	DP12	24hrs vs 144hrs at 37°C	1.6	Up at 144hrs	Hex A is a heterodimer consisting of an $\alpha$ - and a $\beta$ -subunit while Hex B is a homodimer of $\beta$ -subunits (Zarghooni et al. 2004). This dimerization is required for the activity of Hex A and Hex B as a number of residues from each subunit are used to structurally complete and stabilize active-site residues of the opposing subunit (Tropak and Mahuran 2007). Both enzymes are involved in the degradation of glycans through the hydrolysis of the O-glycosidic bond of N-acetylglucosamine residues (Anniller, Hollister, Jarvis 2006). The suboptimal incorporation of the inactive precursor $\alpha$ and $\beta$ subunits into the lysosome results in their secretion by the cell, release of the active, fully mature protein complex into the cell culture media only occurs upon cell death (Mahuran 1999).
	DP12	37°C vs 31°C at 144hrs	2.7	Down at 31°C	
<b>beta-hexo-saminidase subunit beta</b>	DP12	24hrs vs 144hrs at 37°C	2.6	Down at 144hrs	
	DP12	37°C vs 31°C at 144hrs	2.1	Up at 31°C	

In summary, through the use of a quantitative label-free LC/MS based approach to analyse the extracellular proteome of CHO cells, one of the main aims of this thesis has been achieved; it has provided information on the differential-expression of proteins that could impact on cell growth and viability such as VEGFA and VEGFC and Hepatoma-derived growth factor, and the apoptotic proteins GNB2L1 (pro-apoptotic) and Clusterin (anti-apoptotic). In addition, this analysis has identified an overall increase in the abundance of a number of proteases and glycosidases at late phase culture compared to early phase culture (under temperature-shift and non-temperature-shift conditions) and an overall decrease in the abundance of proteases and glycosidases under temperature-shift compared to non-temperature-shift conditions. The proteases Cathepsin B, MMP19, BMP-1 and the glycosidases,  $\alpha$ -N-acetylgalactosaminidase, Hex A and Hex B have been identified in this study, all of which have the potential to adversely affect product quality.

### **5.2.2 Qualitative analysis of CHO-K1 and DP12 cell lines under temperature-shifted conditions**

In addition to quantitative proteomic analysis, a qualitative study of mass spectral data gathered from the profiling of a CHO-K1 and DP12 cell line was also carried out. It was hoped that by comparing the lists of proteins identified in each cell line, any similarities and / or differences in the HCP profile of each cell line would be revealed. In total, 1,181 proteins (with a minimum of 1 peptide) were identified between both cell lines. Of these, 661 had two or more peptides. In both instances a significant cohort of proteins were unique to each cell line. 304 proteins with a minimum of one peptide, of which, 154 proteins had two or more peptides, were uniquely identified in the non-producing CHO-K1 cell line. In the IgG producing DP12 cell line, 215 proteins with a minimum of one peptide were uniquely identified. This includes 145 proteins that had two or more peptides (Section 3.3.7). Given the unfractionated methodology used in this investigation, this figure compares favourably with other studies.

For example, a recent study used 2D LC prior to LC/MS analysis to fractionate the proteome and secretome of a CHO-K1 cell line. Analysing 12 conditioned media fractions taken from a CHO-K1 cell line that had been serum starved for 12 hours after being grown in media supplemented with 10% FCS to 80% confluence, 1,977 proteins were identified (Baycin-Hizal et al. 2012).

It should be pointed out however, that of the proteins identified, the precise number of secreted proteins was never specified. A similar fractionation based approach was used to identify autocrine factors secreted by CHO cells grown in a 5L bioreactor (Lim et al., 2013). Two different mass spectrometry technologies - LTQ Orbitrap Velos, a hybrid linear ion trap-orbitrap instrument, and a Synapt G2, quadrupole-time-of-flight (Q-TOF) instrument were also used to aid in the identification of 2,512 proteins of which 290 were identified as secreted proteins. It was thought that many of the remaining proteins were present as a result of cell lysis occurring during cell culture or harvest of the conditioned media (Lim et al. 2013).

In this investigation both cell lines were also found to maintain unique HCP subpopulations independent of temperature. A recent investigation into the secretome of DG44 and CHO-S cell lines using a metabolic labelling strategy found that there was 'significant overlap' in the proteins secreted by the two cell lines (Slade et al. 2012). While shake flask cultures were used in the generation of conditioned media samples analysed in the report, making it comparable with the results discussed in this thesis, it should be pointed out that the investigation only compared samples taken from the supernatant day 3 cultures, compared with samples taken from multiple time points up to day 7 (168hrs CHO-K1, 120hrs post-temperature-shift) and day 9 (216hrs DP12, 144hrs post-temperature-shift) as reported in this investigation. The direct comparison of samples taken from such early stages in the culture may therefore account for the similarity of the proteins secreted by DG44 and CHO-S cell lines.

A 2D PAGE analysis of three DHFR negative CHO cell lines (FR4, DP12 and DP7) found that there was very little difference in their HCP profiles (Krawitz et al. 2006). Similar findings were also reported in other studies that utilized 2D DIGE to examine the HCP profile of different clones, under different culture conditions, with different levels of viability (Grzeskowiak et al. 2009; Jin et al. 2010).

However all three studies analysed conditioned media samples taken from stirred tank bioreactors which are known to cause cell damage as a result of the use of mechanical agitation (Tait et al. 2012; Velez-Suberbie et al. 2012). This could lead to a situation where, although cell viability is high, the culture media contains a large population of the CHO proteome as a result of cell lysis.

This conclusion appears to be backed up in two of the studies by their findings on the impact of cell viability; although they both found that cell viability was the main factor in determining HCP profile, dropping from 92% to 10% and from >90% to 40% respectively, this decrease mostly appeared to affect fold-change of the proteins identified as opposed to the actual HCP content (Grzeskowiak et al. 2009; Jin et al. 2010). One would have expected that if cells were maintaining membrane integrity while in a state of high viability that a different HCP profile would be observed upon reduction in cell viability and resulting lysing of cells.

Due to the low shear nature of shake flasks, this study identified HCPs deposited into the culture media through secretion or cell leakage as a result of an accumulation of dead or damaged cells over the natural progression of the cell growth cycle as opposed to the lysis of cells due to shear forces. The findings discussed here are therefore most applicable to other low shear environments such as disposable wave bag bioreactors (Öncül et al. 2010). It should be pointed out however that this may not be true for all disposable technologies as disposable stirred tank bioreactors are likely to provide an environment similar to their stainless steel equivalents (Fernald et al. 2009; Ullah et al. 2008).

It should be pointed out that the reduction in cell viability is likely to be as a result of a combination of apoptosis and necrosis, despite the relatively low shear environment of the shake flask. While the trypan blue staining method employed in this study is a well-established technique for the detection of non-viable cells, it would not take into account necrotic cells that have lost all membrane integrity resulting in the cell becoming fragmented. This may have led to an over-estimation of cell viability and would account for the significant proportion of intracellular proteins in the culture media.



Future studies should consider the use of a cell necrosis assay to facilitate the detection of this form of cell death by this means. This would provide a clearer understanding of the source of intracellular proteins in the culture media i.e. cell death or secretion by viable cells.

While the differences in the HCP profiles observed in quite striking, it should be remembered that a Protein A affinity chromatography column was used to reduce the levels of IgG in the DP12 cell culture supernatant samples prior to analysis while the CHO-K1 supernatant did not undergo such treatment.

HCPs have been reported to bind to Protein A columns (Hogwood et al. 2013), although such binding has been shown to be low in abundance and by a small cohort of proteins (Follman and Fahrner 2004; Shukla and Hinckley 2008), particularly for resins with an agarose backbone, such as that used in MabSelect, as used in this thesis (Tarrant et al. 2012), especially when compared to resins whose backbones are based on alternative materials such as controlled pore glass (Tarrant et al. 2012). However, the propensity for HCPs to interact with such Protein A columns is known to increase where the overall IgG titre is lower than the maximum binding capacity of the column, as was the case in this instance (B. Mullen 2014, personal communication). Therefore, the binding of a cohort of the HCPs from the DP12 culture to the Protein A column may account for the identification of some proteins in the CHO-K1 culture that were not identified in the DP12 supernatant. However, since CHO-K1 samples were not processed using a Protein A column, this does not account for the HCPs identified as being unique to the DP12 culture. This raises important questions about the suitability of selecting a non-producing cell line to generate antigenic material against which antibodies are raised in host animals for the production of anti-HCP ELISAs (further discussed in section 5.2.2.1).

### **5.2.2.1 Effect of temperature-shift on the HCP profile**

Although DP12 cells were grown for 48hrs longer than the non-producing CHO-K1 cells, both cell lines maintained high cell viability over the selected culture period. This is particularly true for the 31°C culture where both cell lines had viabilities >95%. CHO-K1 and DP12 37°C cultures were also comparable however, with cell viabilities of 95% and 85% respectively. This makes the identification of a sub-population of HCPs unique to both cell lines quite an interesting result.

One may have expected that the DP12 cell line, with its longer time in culture and slightly reduced viability at 37°C compared to that of the K1 cells, would have produced a HCP profile more closely overlapping that of the non-producing cell line. Culture temperature-shift also appears to have an effect on the HCP profile, with unique HCPs identified in both cell lines. It is known that temperature-shift can have cell line specific effects; these results indicate that this also extends to the HCP profile (Kumar, Gammell, Clynes 2007).

This however does not explain the differences seen in the non-temperature-shifted culture. While other studies have found temperature-shift does not significantly impact on the HCP profile, these investigations were carried out using stirred tank bioreactors as discussed previously (Jin et al. 2010). Aside from extended time in culture and a slightly reduced viability, another reason for the distinct DP12 HCP profile could be due to the secretion of an IgG from the cell. Identification of intracellular proteins in the culture supernatant of IgG secreting cell line, particularly those associated with the production of recombinant protein, has been reported previously (Hogwood et al. 2013).

One reason for the difference in the HCP profiles could be due the production of the IgG molecule by the DP12 cell line. The presence of expression machinery in addition to cellular stress and alterations in metabolism brought about by burden of producing the recombinant therapeutic could result in the expression of proteins not seen in the CHO-K1 culture (Dinnis et al. 2006; Tait et al. 2012).

Although DP12 cells are derived from the CHO-K1 lineage, genetic mutations in the IgG producer may create elements of the proteome that are distinct from its CHO-K1 ancestor (Xu et al. 2011).

However since we are using a proteomic database derived from the CHO-K1 cell line to profile the HCPs from both cell lines, genetic difference between the non-producing and IgG producing cell line cannot entirely account for the differences seen. In fact, with further iterations of the CHO genome and associated proteomic database, it may be possible to identify greater distinctions between the HCP profile of the CHO-K1 and DP12 cell lines.

When the lists of proteins unique to temperature-shifted and non-temperature-shifted cultures in CHO-K1 and DP12 cell lines are compared it can be seen that the vast majority of the proteins present are intracellular proteins. This appears to indicate that the way in which the cell lines lyse, either as a result of shear forces or loss of viability, at the respective temperature has the greatest impact on HCP profile rather than deliberate secretion by the cell at a given temperature.

#### **5.2.2.2 Qualitative analysis of secreted proteins**

The protein Clusterin was found in both non-producing and IgG producing cell line at reduced and non-reduced culture temperatures. In its mature form, Clusterin is a secreted glycoprotein of 76-80 kDa that is known to play an anti-apoptotic role in a wide variety of cells (Carlage et al. 2012).

Interestingly, a truncated form of the protein has been implicated in both apoptotic (nucleus) and anti-apoptotic (cytoplasm) signalling events depending on their cellular translocation (Nuutinen et al. 2009). These localisation dependent roles provide an insight into the complex mechanisms through which this protein regulates cellular function. Although the precise mechanism by which secreted clusterin enhances cell survival is largely unknown it has been shown that overexpression of the stress chaperone prevents TNF $\alpha$  mediated induction of p21 (Flanagan et al. 2010). This may occur either through clusterin binding to TNF $\alpha$  in the media or as a result of a truncated form of clusterin being translocated to inhibit Bax activation in the mitochondria which in turn prevents the release of cytochrome and subsequent caspase activation (Zhang et al. 2005). Clusterin-induced apoptosis has been reported to occur through binding of a truncated, non-glycosylated form of the protein binding to Ku70/80 in the nucleus inhibiting DNA repair (Leskov et al. 2003). From the data obtained in this study it is not possible to definitively determine which form

of Clusterin has been identified, although given that it is cell culture supernatant that is being analysed it is most likely the secreted protein.

It would be potentially interesting to perform further studies to investigate expression levels of Clusterin in the nucleus and culture media of the CHO-K1 and DP12 cell lines under both temperature-shifted and non-temperature-shifted conditions over time in culture to identify the role that Clusterin plays in the cell culture cycle of CHO cells.

Another interesting finding was the identification of CIRBP (Cold-inducible RNA-binding protein) and RBM3 (Putative RNA-binding protein 3) in the conditioned media of both temperature-shifted cell lines. CIRBP and RBM3 are both highly conserved proteins that belong to a glycine rich RNA-binding protein family (Al-Fageeh and Smales 2013). CIRBP expression may be induced in response to oxidative stress and mild hypothermia (Kumar et al. 2008a). Known to stabilize mRNA by binding to the 3'untranslated region of specific transcripts, CIRBP also facilitates the transport of mRNAs to the ribosome for translation (Brochu et al. 2013). It has been hypothesized that CIRBP may play a role in preventing apoptosis under mildly hypothermic conditions after showing that the knockdown of CIRBP significantly increased apoptosis in temperature-shifted MEB5 stem cells (Saito et al. 2010). Interestingly, it has also been postulated that CIRBP may have a neuroprotective effect on brains that have suffered ischemia under hypothermic conditions (Liu et al. 2010). Further work is required however to determine whether this is as a result of active gene modulation or activation of the extracellular signal regulated protein kinase pathway. Interestingly, while CHO cells overexpressing CIRBP showed increased productivity of INF-gamma, neither the overexpression nor the down-regulation of CIRBP has been shown to effect growth (Hong et al. 2007; Tan et al. 2008).

Similar to CIRBP, RBM3 stabilises RNA under reduced culture temperature conditions (Fujita 1999). This facilitates the translation of specific proteins despite the overall inhibition of protein synthesis via a cap-independent mechanism (Chappell, Owens, Mauro 2001). Given the dampening effect of miRNAs on protein synthesis, it has been suggested that RBM3 may inhibit miRNA activity either through the blockage of miRNA binding sites on mRNAs, interactions with

ribonucleoprotein complexes containing microRNAs or through the regulation of miRNA synthesis/degradation pathways (Barron et al. 2011; Dresios et al. 2005).

More recently however it has been shown that RBM3 actually promotes pre-miRNA processing by directly binding to these ~70 nt intermediates and facilitates the interaction of ribonucleoproteins with the active Dicer protein complex in neuronal cells under mildly hypothermic conditions (Pilotte, Dupont-Versteegden, Vanderklish 2011). It is perhaps then interesting to note that the increased expression of RBM3 as a result of temperature-shift has also been shown to play an important role in the hypothermia-induced neuroprotection of neuronal cells; however the precise mechanisms by which this occurs remain unclear (Chip et al. 2011).

A protease common to both cell lines but unique to 31°C was also identified; i.e., Proteasome subunit beta type-8, also known as Low Molecular Mass Protein 7 (LMP7). The 20S (or constitutive) proteasome contains 28 subunits, of which three are proteolytically active:  $\beta$ 1,  $\beta$ 2 and  $\beta$ 5 (Arima et al. 2011; Muchamuel et al. 2009). LMP7 ( $\beta$ 5i) replaces the  $\beta$ 5 subunit upon the expression of the cytokine IFN- $\gamma$  by activated T helper 1 lymphocytes, CD8<sup>+</sup> CTLs and natural killer (NK) cells (Kloetzel and Ossendorp 2004).

As such,  $\beta$ 5i (along with  $\beta$ 1i and  $\beta$ 2i which are also induced, replacing their  $\beta$ 1 and  $\beta$ 2 counterparts) are often referred to as 'immunoproteasome' subunits. The replacement subunits display enhanced cleaving efficiency after basic or hydrophobic residues, however their ability to cleave after acidic residues appears somewhat diminished (Rock and Goldberg 1999). Specifically, the  $\beta$ 5i is said to have chymotrypsin-like activities (Borissenko and Groll 2007). This affinity for cleaving basic or hydrophobic residues could significantly impact on the product yield of recombinant proteins rich in such residues.

Engineering the product such that these residues would be replaced by neutral, hydrophilic ones, would not only make the product more resistant to cleaving by this protease, but would also have the added advantage of making the product more soluble and less susceptible to aggregation (Wu et al. 2010b).

The identification of unique HCP subpopulations has important implications for generating antigen material for the preparation of an ELISA. Non-producing cell lines are typically used in the creation of anti-HCP antibodies as they do not contain the product of interest and therefore will not cross-react with the drug substance (Savino et al. 2011a). Any antigenic material generated must be representative of the cell line and process that the ELISA will be used for (Savino et al. 2011b).

One of the aims at the outset of this thesis was to identify whether there were differences between (i) temperature-shifted and non-temperature-shifted cultures (ii) IgG producing and non-producing cell lines.

Qualitative analysis of LC/MS data has revealed differences in the HCP profiles of non-producing CHO-K1 and IgG producing Dp12 cells under both temperature-shifted and non-temperature-shifted conditions. Furthermore, comparisons of the HCPs profiles of both cell lines resulted in the identification of cohorts of proteins that appeared to be unique to each cell line.

The results discussed show that the choice of cell line and culture temperature can have an impact on the HCP profile generated. Given that a significant proportion of a HCP profile is made up of intracellular proteins (Hogwood et al. 2013; Tait et al. 2012), it makes sense that lysates from non-producing host cell lines are used as antigenic material for the immunisation of animals to generate anti-HCP antibodies for use in ELISAs (Krawitz et al. 2006). However if lysates are obtained from cultures grown under conditions different to those in which the ELISA is expected to perform e.g., temperature-shift, then the ELISA may not contain antibodies specific to proteins that are only expressed under the certain conditions. This also extends to the use of cell culture supernatant as anti-genic material for the immunisation of animals to generate an ELISA.

For example, anti-HCP antibodies generated from the cell culture supernatant from the CHO-K1 cell line at 37°C had been used in the been chosen as antigenic material for the immunisation of animals to generate anti-HCP antibodies, this would not be able to detect a large number of HCP contaminants in the DP12 cell line in either temperature-shifted or non-temperature-shifted cultures. It would also be incapable of detecting certain HCPs found in the CHO-K1 31°C culture.

Failure to detect such HCPs present in the final drug substance could potentially illicit an adverse immune reaction in the patient receiving the contaminated product (Savino et al. 2011a).

Placing such a high emphasis on the ELISA for the identification of HCP impurities in the final drug product ignores the fact that a proportion of the HCP population may not result in an immune response from the animal being immunised, as such, no anti-HCP antibodies will be generated against these proteins (Zhu-Shimoni et al. 2014). This could result in the presence of a HCP impurity that could go undetected in the final drug product. The application of a mass spectrometry based approach, while not as sensitive as an ELISA, provides important information on proteins that may go undetected by immunological methods e.g., identification and, in the presence of a known standard, abundance. Information such as this advances our knowledge of the process and changes the focus from monitoring the presence of process related contaminants to being able to define and control the risk factors associated with the presence of HCP impurities.

### **5.2.3 Utilisation of CHO specific proteomic databases in the analysis of LC/MS data in a HCP study**

The recent release of two CHO specific databases has filled in a significant gap in suite of tools available to study CHO cells. A report by Meleady et al. (2012) showed that use of the Bielefeld-BOKU-CHO (BB CHO) expressed protein database complemented the use of the CHO NCBI database generated from the CHO-K1 genomic data as published by Xu et al. (Meleady et al. 2012; Xu et al. 2011). Although the CHO-NCBI database accounted for a large proportion of the HCPs identified, the BB-CHO database enabled the identification of a number of proteins not identified by the CHO-NCBI database. The fact that the BB-CHO database was curated from the cDNA of a number of different CHO cell lines grown under different processing conditions made this database particularly well suited for the proteomic analysis of HCPs generated by two different CHO cell lines grown over time in culture under temperature-shifted and non-temperature-shifted conditions. The use of both databases has enabled the identification of additional proteins that would not otherwise have been identified by one database alone.

The resulting increased proteome coverage has provided a wealth of information on the make-up of the HCP profile under different processing conditions as well as at multiple time points over the culture period. The availability of the two CHO databases enabled the identification of a large number of proteins with criteria that gave a high level of confidence in the identification.

Previously proteomics studies relied on sequence homology between the CHO and another well annotated species such as human or mouse for the identification of a protein (Baik et al. 2006; Baik et al. 2008; Baik et al. 2011; Kantardjieff et al. 2010; Kumar et al. 2008a). Relying on sequence homology for the study presented here would inevitably have led to the loss of a number of identifications. This may have included proteins of interest because of their potential ability to impact on product quality or culture viability, such as the protease Cathepsin D or the growth factor Hepatoma-derived growth factor.

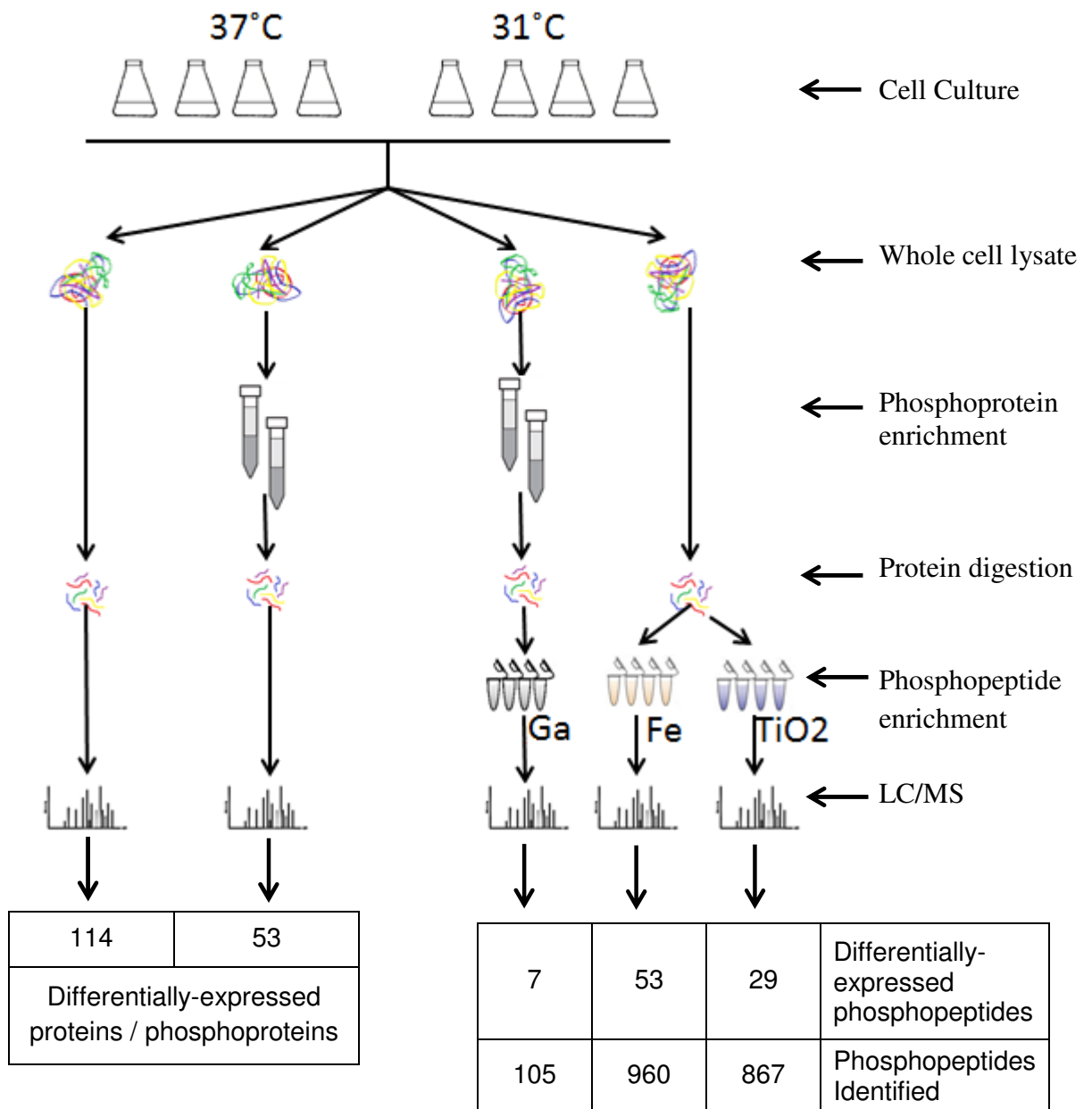
#### **5.2.4 Overall Summary**

The quantitative LC/MS analysis of a non-producing CHO-K1 and an IgG secreting DP12 cell line identified changes in the relative abundance of proteins that could negatively impact on product quality such as Cathepsin D, Thioredoxin 1 and  $\alpha$ -galactosidase A through an increase in their abundance at late stage culture under both temperature-shift and non-temperature-shift conditions. In addition, proteins that could affect cell growth (HGDF and VEGFA) both over time in culture and under temperature-shifted and non-temperature-shifted conditions were also identified. This study also found that culture temperature and cell line selection can both have an effect on the HCP profile. A change in the HCP profile has important implications for the ability of an ELISA to detect contaminating HCPs in the final drug substance. This investigation shows that LC/MS analysis is a useful tool for providing information on HCPs produced by these cells that affect product integrity and culture viability.



## **6.0 Phosphoproteomic Analysis of CHO Cells Overview**

This section discusses the findings of the investigation into the phosphoproteome of temperature-shifted CHO cells. This investigation initially sought to identify changes in the phosphoproteome by using a MAP kinase array to detect changes in phosphorylation of 26 kinases. Changes in the phosphorylation of Erk, AKT and Jnk are discussed in this section as is the utility of protein arrays for profiling CHO cells. The recent availability of two CHO specific proteomic databases made the use of mass spectrometry a viable option for profiling the phosphoproteome of CHO cells under reduced culture temperature conditions. As a result, LC/MS was utilized to obtain a more global overview of the cell signalling events occurring 36hrs post temperature-shift. As such, whole cell lysate, phosphoprotein-enriched whole cell lysate, gallium oxide phosphopeptide enrichment of phosphoprotein-enriched sample and titanium and iron oxide phosphopeptide enrichment of whole cell lysate samples were analysed (Figure 6.0.1). The use of different phosphopeptide enrichment methods, the identification of signalling events effecting translation and cell structure as well as the contribution made by the availability of two CHO specific proteomic databases are discussed, as well as the partialities of each method and the biological insights gained from their use.



**Figure 6.0.1** Schematic of the different approaches used to profile the phosphoproteome of CHO cells under temperature-shift conditions using mass spectrometry. This schematic also includes tables of the number of differentially-expressed proteins, phosphoproteins and phosphopeptides identified following quantitative label-free LC/MS analysis. The table also includes the number of phosphopeptides identified from each phosphopeptide enrichment following qualitative analysis of the LC/MS data.

## **6.1 Phospho-Kinase Array of SEAP cells subject to temperature-shift**

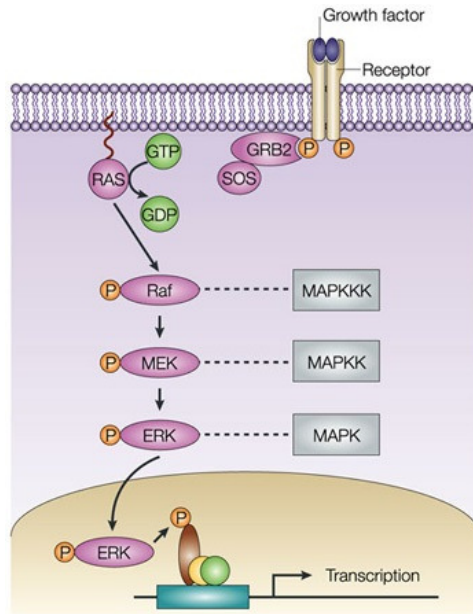
Although a number of proteomic profiling studies have been carried out (Gammell et al. 2007; Kumar et al. 2008a; Roobol et al. 2009; Underhill et al. 2005; Underhill and Smales 2007) the mechanisms behind the cellular response to temperature-shift by mammalian cells are poorly understood (Kumar, Gammell, Clynes 2007; Sunley and Butler 2010). Given the dynamic response of CHO cells to reduced culture temperature it makes sense that phosphorylation would have a role to play as has been detailed in a number of studies (Kaufmann et al. 1999; Roobol et al. 2011; Underhill et al. 2005). Despite this, there still remains a gap in our knowledge as to the precise signalling pathways through which this response occurs. Protein Kinase Arrays are one means of profiling such signalling cascades and have been successfully employed in a variety of studies involving human cell model systems (Lin et al. 2012). They have however yet to be utilised for the profiling of the CHO phosphoproteome. In this experiment a kinase array was used to profile the relative levels of phosphorylation on 26 kinases in CHO cells subject to temperature-shift.

MAP Kinase array analysis identified changes in the phosphorylation of 12 kinases. This included P53 and three isoforms of P38. Phosphorylation of P53 by P38 whose activation was initiated by ATR (ataxia telangiectasia mutated- and Rad3-related kinase) had previously been identified as a key signalling pathway involved in cell cycle arrest of temperature-shifted cells (Roobol et al. 2011). Agreement between the array and findings by Roobol et. al. provided a partial validation of the ability of the array to successfully detect alterations in the phosphorylation of different kinases in CHO cells. Actual validation was carried out by western blot investigating not only the 24hr period post temperature-shift profiled by the array, but also 48, 72 and 96hrs after culture temperature reduction. This provided an indication of how phosphorylation of the respective kinase behaved at both the time point investigated by the array as well as over time in culture.

### **6.1.1 Phospho-ERK1/2 in Temperature-shifted CHO cells**

Immunoblotting of ERK1/2 confirmed the increase in phosphorylation of the kinase in the 31°C culture 24hr post temperature-shift as found by the array. Furthermore it provided evidence of a sustained increase in phosphorylation of ERK1/2 over the 96hr period compared to the 37°C culture. Follow up western blot studies using lysates taken from DP12 cells grown over the same 96hr period post temperature-shift used in the CHO SEAP culture revealed phosphorylated ERK1/2 to be slightly reduced at 31°C and for this decrease to become more pronounced over time in culture. One potential reason for this is that growth of the DP12 cell line is greatly reduced by culture temperature reduction. This growth remains stifled with little increase in cell number over the growth cycle compared to the CHO SEAP cell which continues to grow albeit at a reduced rate. Such a dramatic repression, possibly caused by a near complete cell cycle arrest involving other signalling pathways, could negatively impact on growth signalling pathways involving ERK (Mendoza, Er, Blenis 2011). Reduced culture temperature is also known to affect cells lines differently (Yee, Gerdtzen, Hu 2009); this too could account for the different phosphorylation profiles of ERK1/2.

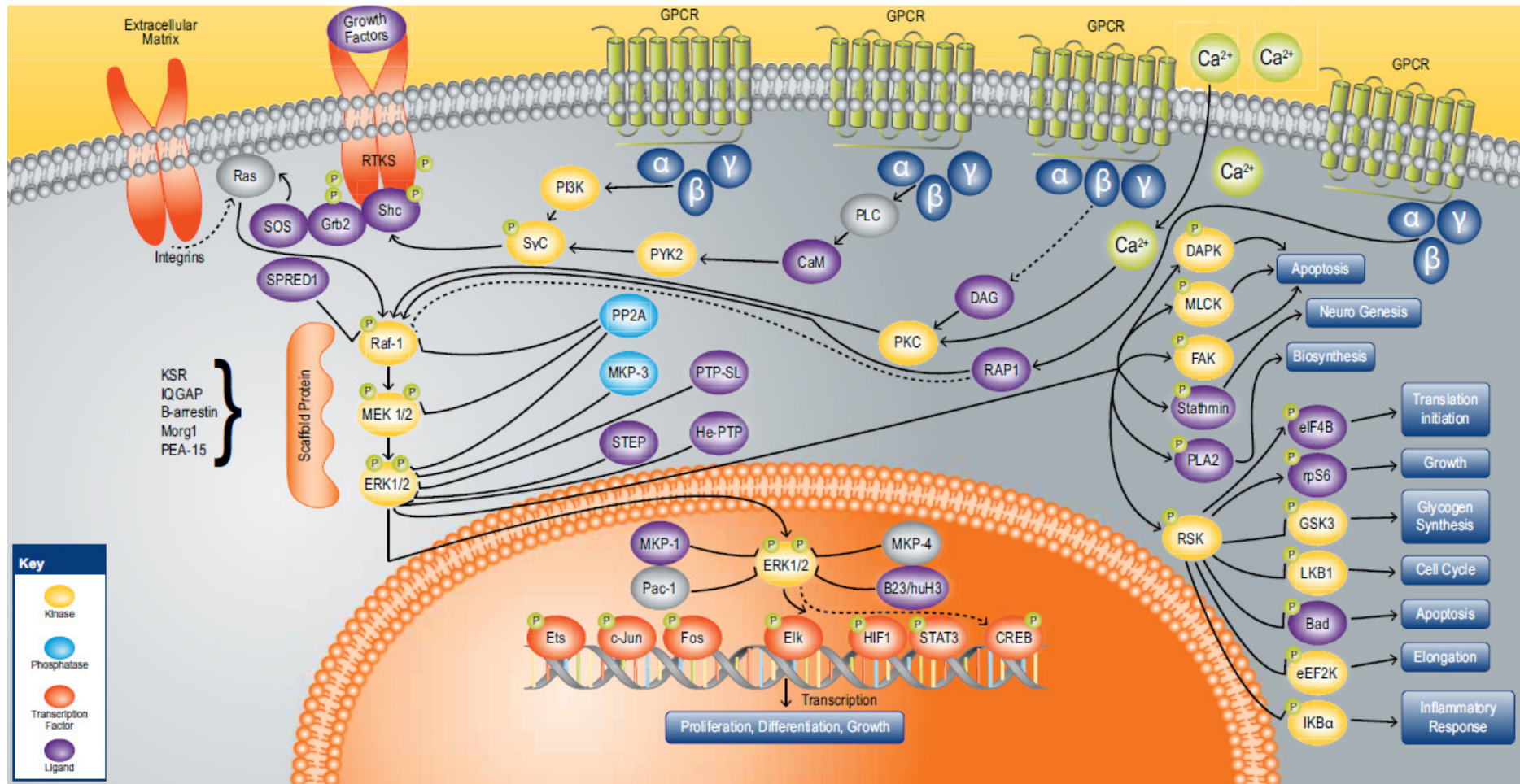
To gain a better understanding of the signalling taking place via the ERK1/2 pathway western blots were conducted of its well-known upstream regulator; MEK1/2 (Figure 6.1.1 and Figure 6.1.2) (Lee Jr and McCubrey 2002). Although total ERK abundance remain constant over the course of the time points observed, abundance of total MEK decreased upon temperature-shift, and remained consistently low throughout the 31°C culture. Unfortunately no phosphorylated MEK could be detected in the CHO samples.



**Figure 6.1.1** Images shows simplified schematic of MEK – ERK pathway adapted from image by (Kim and Bar-Sagi 2004).

Without direct evidence of the phosphorylation status of MEK no firm conclusion can be drawn as to whether MEK is directly involved in the activation of ERK, although reduction in the overall expression of MEK would suggest that this is not the case.

This leads to the possibility that ERK activation is taking place via some MEK independent mechanism (Jiang et al. 2011). Characterising the role that ERK plays in temperature-shifted CHO cells would require further work to determine the activating kinase. Further profiling may also be necessary in order to identify subsequent signalling pathways involved given ERKs ability to transduce multiple signals from various sources to specific cellular responses (Ebisuya, Kondoh, Nishida 2005).



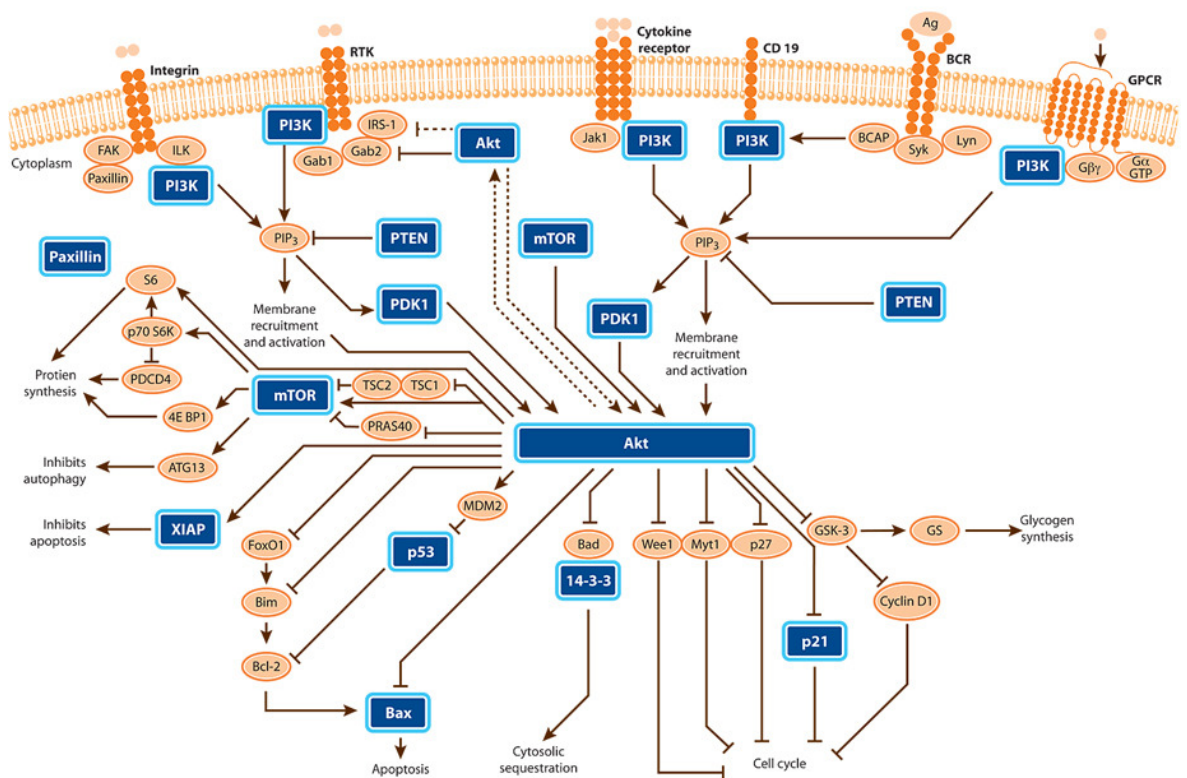
**Figure 6.1.2** Image showing the variety of cellular processes that Erk is involved in controlling and the complexity of the related pathways. Image adapted from abcam (<http://www.abcam.com/index.html?pageconfig=resource&rid=15825> 02Dec2013).

### **6.1.2 Phospho-AKT in Temperature-shifted CHO cells**

Kinase array profiling indicated that there was no change in the expression of phosphorylated AKT1. Validation by western blot revealed that total AKT1 was in fact down-regulated in CHO SEAP and the phosphorylation of AKT1 was down-regulated in DP12 at 31°C. While this result is at odds with that of the initial profiling, it is possible that non-specific binding on the array could account for this as discussed later in this chapter (Section 6.2.3). It has been shown that overexpression of constitutively active AKT causes delay in the onset of apoptosis and extends culture viability while inhibition of AKT activity causes induction of cell cycle arrest and apoptosis during the early G<sub>1</sub> phase in CHO cells (Figure 6.1.3) (Opstal et al. 2012). CHO cells experiencing culture temperature reduction are also known to enter cell cycle arrest at the G<sub>1</sub> phase (Chen et al. 2004; Moore et al. 1997). This suggests that AKT1 may play some role in causing cell cycle arrest in temperature-shifted CHO cells. Western blot analysis indicated a reduction but not total inhibition of phosphorylated AKT1. Tempering of signalling via AKT1 may allow the cells to continue to proliferate, albeit at a reduced rate, but stopping short of inducing apoptosis. Although the increase in phosphorylation of AKT2 was not validated, this result could potentially point to independent signalling mechanisms between AKT1 and AKT2.

Both AKT isoforms AKT1 and AKT2 have been implicated in biologically distinct functions and are known to operate independently of each other (Manning and Cantley 2007). Further studies would however be required to confirm such roles in temperature-shifted CHO cells.





**Figure 6.1.3** Schematic of pathways and cellular processes regulated by AKT. Image adapted from horizon discovery (<http://www.horizondiscovery.com/images/pathways/p13k-akt.jpg> 03Dec2013)

### 6.1.3 Phospho-JNK and MEK6 in Temperature-shifted CHO cells

Lack of antibody specificity proved problematic not only for the detection of MEK, but also for the accurate identification of JNK and MEK6 during western blot validation. Due to the poor availability of anti-CHO antibodies immunoblotting of CHO samples very often requires the use of anti-human or anti-mouse antibodies for the detection of the protein of interest. Successful detection relies on homology between specific amino acid sequences in the human or mouse version of the protein and its CHO counterpart. In the absence of antibody reactivity between species the target protein may not be detected (Farady et al. 2009).

Conversely inter-species cross reactivity can also result in the detection of proteins other than the protein of interest. Non-specific binding is a problem known to be associated with protein-protein interaction based technologies such as arrays and



other assays (Farajollahi et al. 2012). Results such as those observed in the JNK and MEK6 immunoblots raise the possibility that non-specific binding to the array may lead to erroneous results. To help increase specificity many arrays utilise a sandwich based approach requiring a second labelled detection antibody to bind to a recognition site on the kinase that does not overlap with that of the capture antibody (Simpson 2007). Therefore, anti-human capture and detection antibodies used in this array require a homologous binding site on the CHO kinase without which the likelihood of non-specific binding increases. Antibodies used for western blot validation are less likely to suffer from non-specific binding as they only require one binding site on the target protein since the secondary binds to the primary antibody and not the protein of interest. This was potentially in evidence where phosphorylation of AKT1 was not deemed to change significantly in the array analysis but was found to be down-regulated at 31 °C by western blot.

#### **6.1.4 Summary**

While protein array technology has been around for some time (Tse-Wen Chang 1983), to the best of our knowledge, this is the first instance of a protein kinase array being applied to CHO cells. Although mass spectrometry may be viewed as a more comprehensive means of acquiring a global overview of the phosphoproteome of a cell, protein array based techniques provide a more targeted approach to investigating specific kinases of interest. Future applications would benefit from the use of capture and detection antibodies that are specific to CHO as anti-human antibodies were found to possibly reduce specificity. Nevertheless, its utilisation identified ERK, JNK and AKT as being potential targets of interest in temperature-shifted CHO cells.

## **6.2 Quantitative Phosphoproteomic label-free LC/MS analysis of CHO SEAP cells subject to temperature-shift**

### **6.2.1 Quantitative label-free LC/MS analysis of CHO SEAP cells subject to temperature-shift**

A number of proteomic studies have been conducted on CHO cells under temperature-shifted conditions in an effort to gain insights into changes that take place in the proteome which may contribute to increased productivity, altered growth rate and extended cell viability (Baik et al. 2006; Kaufmann et al. 1999; Kumar, Gammell, Clynes 2007). To date however, these profiling studies have utilized gel based technologies. To the best of my knowledge, this is the first such proteomic investigation to use a quantitative label-free mass spectrometry based approach. Using this technology to analyse the proteome of temperature-shifted CHO SEAP cells 114 proteins were identified as having  $\geq 1.5$  fold-change (72 up-regulated, 42 down-regulated at 31°C).

#### **6.2.1.1 Differential expression of proteins involved in translation**

GO analysis on the differentially regulated proteins identified revealed significant enrichment of processes relating to translation and RNA processing at 31°C. This is partially due to the high number of ribosomal proteins identified as up-regulated at 31°C including 60S ribosomal proteins L4, L6, L13, L13a, L14, L26, L28, L34 and 40S ribosomal proteins S6, S8, S9 and S25. Although reduced culture temperature is known to limit global translation initiation (Roobol et al. 2009), it is evident from the increase in specific productivity of cells exposed to such conditions that this tightly controlled process must still allow relatively high levels of the product's gene coding sequence be translated. It has been reported that the existence of ribosomal subpopulation with specific translation properties may allow such an event to occur (Xue and Barna 2012).

In addition to this, cold shock proteins such as CIRP and Rbm3 may play a role as RNA chaperones, facilitating transcription and translation, under low culture temperature conditions (Ferry, Vanderklish, Dupont-Versteegden 2011; Tong et al. 2013).

While neither of these proteins were identified in this study, it has been shown that Rbm3 associates with a number of 40S and 60S ribosomal proteins including 60S ribosomal protein L4 (Dresios et al. 2005).

Two eukaryotic translation initiation factors, eIF3i and eIF3b, were both found to be 1.6-fold up-regulated upon temperature-shift. It has been reported that eIF3i in CHO-K1 cells is down-regulated following temperature-shift and is subsequently up-regulated upon restoration of normal culture temperature conditions (Roobol et al. 2009). While overexpression of eIF3i has been shown to increase cell proliferation, cell size and cell cycle progression its upregulation has also been associated with cell lines that maintain high levels of specific productivity over time in culture (Ahlemann et al. 2006; Meleady et al. 2011).

Although eIF3i is known to be involved in mRNA translation and transcription studies have found gene expression of eIF3i sufficiently stable to be considered a house keeping gene in CHO cells (Bahr et al. 2009). Although eIF3i evolutionarily highly conserved between human and yeast, deletion and knockdown studies have shown it to be dispensable for active eIF3 complex formation (Bahr et al. 2009; Masutani et al. 2007). This would suggest that eIF3i may be involved in translational control of specific mRNAs or required under particular cellular conditions (Dong et al. 2013; Masutani et al. 2007). Like eIF3i, eIF3b is also part of the eIF3 complex, both of which bind concurrently to the spectrin domain of eIF3a. Unlike eIF3i however, eIF3b, along with eIF3a and eIF3c, make up the three subunits essential for the formation of the active mammalian eIF3 complex, of which, eIF3b is a scaffolding subunit (Dong et al. 2013; Liang et al. 2012; Masutani et al. 2007). Upregulation of eIF3b has been associated with increased proliferation; conversely knockdown of eIF3b has been shown to inhibit proliferation and increase apoptosis (Liang et al. 2012; Wang et al. 2012).

While this is somewhat at odds with the observed phenotype in this instance, it is possible that, given the complex crosstalk between different cellular pathway and multiple layers of regulation within the cell, eIF3b and eIF3i are in some way contributing to the altered cellular state we see in temperature-shifted cells (extended viability/increased productivity), however proliferation enhancement is being blocked downstream of the eIF3 complex. Although there is not always a direct

correlation between transcription and protein synthesis (de Nobel et al. 2001; Gygi et al. 1999), a number of studies investigating the transcriptome of temperature-shifted CHO cells have shown changes in the gene expression of various cellular processes including those related to production of the protein product (Baik et al. 2006; Kantardjieff et al. 2010; Yoon, Song, Lee 2003). As a consequence, this could give rise to changes in the abundance of RNA levels requiring upregulation of RNA processing machinery. Interestingly, the upregulation of proteins involved in mRNA translation in high producing cell lines when comparing NS0 cells with varying levels of specific productivity has also been previously shown (Dinnis et al. 2006).

#### **6.2.1.2 Comparison of this quantitative label-free LC/MS analysis of temperature-shifted CHO cells with other proteomic profiling of temperature-shifted CHO cells in the literature**

When the list of differentially-expressed proteins obtained in this study is compared to other proteomic studies of temperature-shifted CHO cells (Baik et al. 2006; Kumar et al. 2008a), it is clear that using a quantitative label-free LC/MS methodology has provided new insights into the effects of temperature-shift on the proteome of CHO cells. Although there are some common proteins identified such as Vimentin and Annexin A1, both of which are up-regulated at 31°C, displaying a similar trend to previous reports (Baik et al. 2006; Kumar et al. 2008a); many of the proteins identified are unique to this study. This is in contrast to the comparison of results that Kumar et al. (2008) made with those of Baik et al. (2006) which found a number of overlapping proteins. It should however be pointed out that while some common proteins were identified, only half of these had a fold-change in the same direction. It should also be pointed out that there was also significant difference between the two lists of proteins obtained.

There are a number of potential reasons why there may be such little overlap between the proteins identified in this study and those identified in the two studies previously mentioned, these are summarised in Table 6.2.1.

**Table 6.2.1** Summary table comparing quantitative label-free LC/MS analysis of temperature-shifted CHO cells conducted in this thesis with other proteomic profiling of temperature-shifted CHO cells in the literature.

	Experimental conditions presented in this study	Kumar et al. (2008)	Baik et al. (2006)	Comment
Sampling time point	36hrs post temperature-shift (after a total of 84hrs in culture)	72hrs post temperature-shift (after a total of 144hrs in culture)	Cell culture samples were obtained from the 37°C culture ~110hrs after inoculation, while 31°C culture samples were obtained ~160hrs after inoculation.	Both Baik et al. (2006) and Kumar et al. (2008) took temperature-shifted cell samples after cells had entered the stationary phase, ≥144hrs in culture. It has been shown that the proteome of CHO cells can alter over time in culture (Carlage et al. 2012; Wei et al. 2011). It was hoped that by using an earlier time point the effect of temperature-shift would be more pronounced e.g. alterations in pathways involved in growth would be more easily identified in a temperature-shifted CHO cell compared to that of a rapidly growing cell (at 37°C) than if both temperature-shifted and non-temperature-shifted cells were in stationary phase.
Cell culture media	Serum free media	DMEM/F-12 Ham basal media supplemented with 10% FCS	DMEM basal media supplemented with 10% FCS	Both studies used media supplemented with serum whereas this investigation utilised serum-free media. Although the proteome of cells grown in serum-containing media compared to those growing in serum-free media are relatively similar, some differences do exist (Baik et al. 2011). The growth of cells in serum-containing and serum-free media could also partially explain the differences observed between the studies.
Cell line	SEAP-producing CHO-K1	Non-producing CHO-K1	EPO-producing DUKX-B11 (CHO cell line)	Cellular response to temperature-shift has also been shown to be cell line specific. This could also be a contributing factor to the different lists obtained given that an EPO producer and a non-producing CHO-K1 cell line were used in the two studies described (Baik et al. 2006; Kumar, Gammell, Clynes 2007; Kumar et al. 2008); Yoon, Song, Lee 2003). In this investigation a CHO cell line producing SEAP was selected for use as a model recombinant protein product to replicate a cell line producing a glycosylated protein product in industry (Lipcomb et al. 2005).

## **6.2.2 Label-free LC/MS analysis of phosphoprotein-enriched lysates from CHO SEAP cells subject to temperature-shift**

Fractionation of the proteome is often required to obtain further insights into the mechanisms involved in the cellular response to a given condition (Lee, Tan, Chung 2010). This becomes particularly important when trying to study the phosphoproteome of a cell due to the low abundant nature of many phosphoproteins (Beltran and Cutillas 2012). In order to investigate the phosphoproteome of CHO cells subject to temperature-shift, a phosphoprotein enrichment strategy was employed to aid in the identification of differentially-expressed proteins that otherwise may not be seen due to sample complexity.

Label-free LC/MS analysis of phosphoprotein-enriched whole cell lysates from temperature-shifted CHO SEAP cells identified 53 differentially-expressed phosphoproteins, 21 of which were up-regulated upon reduction of culture temperature. It is perhaps interesting to note that a significant number of differentially-expressed proteins identified in the non-enriched whole cell lysate were phosphoproteins (84 out of 114, as classified using Gene Ontology). Although there was some overlap between the lists obtained from the two experiments, phosphoprotein enrichment provided differential information on a cohort of proteins that were not identified in the previous whole cell lysate study.

### **6.2.2.1 Differential expression of phosphoproteins involved in cell structure**

GO analysis, performed using DAVID, on the differentially-expressed phosphoproteins identified significant enrichment of the cytoskeletal organisation pathway. The effects of low culture temperature on cell structure have also been observed in other studies (Baik et al. 2006; Kumar et al. 2008a). Disassembly of cytoskeletal components are known to occur upon exposure to severely reduced culture temperatures. While this does not happen under conditions of mild hypothermia, it could partially explain the down-regulation of structural proteins upon temperature-shift (Al-Fageeh and Smales 2006; Fujita 1999; Roobol et al. 2009).

Although structural proteins such as Vimentin and Annexin A1 were identified using whole cell lysate samples, it is clear that without phosphoprotein enrichment this cohort of cytoskeletal proteins would not have been identified. Knowledge of these proteins has contributed to the overall picture of the effect of temperature-shift on CHO SEAP cells. Other structural phosphoproteins identified in this experiment are further discussed in sections 6.2.2.1.1 to 6.2.2.1.3.

The upregulation of Vimentin and Annexin A1 (from analysis of unenriched whole cell lysate) is in keeping with transcriptome studies which identified an overall upregulation of genes involved in cytoskeletal organisation (Yee, Gerdtzen, Hu 2009). However the phosphoproteins identified in this thesis are predominantly down-regulated (2 up-regulated, 7 down-regulated). The observed down-regulation could be due to a reduction in phosphorylation as opposed to a decrease in overall protein abundance (as explained section 6.2.2.1.3). Also, changes in the transcriptome may not necessarily materialise in the proteome due to the non-linear relationship between the genome and the proteome of a cell as a result of events such as alternative splicing and the presence of non-coding RNAs (e.g. miRNAs) (Altelaar, Munoz, Heck 2013). Nonetheless, changes in the cytoskeleton could be indicative of remodelling of vesicle trafficking systems in the cell to facilitate increased recombinant protein productivity as reported following transcriptome and proteomic studies of temperature-shift CHO cells (Baik et al. 2006; Kantardjieff et al. 2010). Given the co-localisation of mRNAs and initiation factors with the cytoskeleton (Stapulionis, Kolli, Deutscher 1997), changes in the cytoskeleton are also likely to have impacted on other processes such as mRNA transport and translation as evidenced by the enrichment of phosphoproteins involved in mRNA transport (from phosphoprotein enrichment analysis, table 4.3.41) and proteins involved in translation (from unenriched whole cell lysate analysis, table 4.3.35) through GO analysis. Therefore, this analysis provides evidence of signalling events within the cytoskeleton and how this could result in alterations protein synthesis as a result of a reduction in cell culture temperature. Sections 6.2.2.1.1 to 6.2.2.1.3 discuss some of the specific phosphoproteins identified in further detail.

#### **6.2.2.1.1 Ezrin**

Ezrin was identified as being up-regulated at 31°C following quantitative label-free LC/MS analysis of whole cell lysate and phosphoprotein-enriched samples by 2.4 and 2.1-fold respectively, most likely indicating a change in the total abundance of the protein rather than a change in phosphorylation. Despite Ezrin being well documented as a structural protein, the specific mechanisms by which it alters cellular behaviour are largely unknown (Xie et al. 2009). It has however been hypothesised that the membrane cytoskeletal linker protein is activated via phosphorylation of Thr<sup>567</sup> enabling binding of F-actin and the membrane thus regulating cell shape and signal transduction (Zhu et al. 2007).

#### **6.2.2.1.2 Annexin A6**

The down-regulation of Annexin A6 at 31°C in whole cell lysate (1.8-fold) and phosphoprotein-enriched samples (3.4-fold) indicates that there could be a reduction in both the overall abundance and phosphorylation of this protein as the fold-changes appear to be significantly different. Interestingly, the upregulation of Annexin A6 has been linked to increased growth and metastasis in a number of different types of cancer (Oda et al. 2013). Similar to Ezrin, the precise function of this calcium-dependent, membrane-binding protein is largely unknown, although there is some evidence to suggest that the down-regulation of this protein in cancer cell lines leads to increased proliferation and invasiveness (de Muga et al. 2008; Wang et al. 2013). This appears to occur, in part, via the EGFR/Ras signalling pathway (de Muga et al. 2008).

It is possible however that Annexin A6 may not have the same function in CHO cells given that they do not endogenously express EGFR and the overexpression of constitutively active Ras in CHO has not been shown to significantly impact on cell growth, viability or productivity despite the upregulation of Ras being associated with many cancer types (Kim et al. 2011).

#### **6.2.2.1.3 GAPDH**

It was noticed that the 4.2-fold down-regulation of GAPDH in this study was at odds with previous findings in the literature that reported this protein to be up-regulated upon temperature-shift (Baik et al. 2006; Kumar et al. 2008a).



Therefore, the list of differentially-expressed proteins observed would be due to changes in the abundance of the phosphorylated species rather than overall changes in the expression of the total protein itself. If levels of total GAPDH had been differentially-expressed then this would most likely have been detected in the analysis of the whole cell lysate. Given that this was not the case, a change in the abundance of the phosphorylation status of GAPDH is most likely to account for the result obtained.

Existing as a single copy gene, multiple phosphorylation sites on the enzyme most likely enable GAPDH to perform such a diverse range of functions DNA repair and replication, cytoskeletal organisation and apoptosis (Sirover 2005; Tisdale and Artalejo 2007). One such example is the phosphorylation of GAPDH by PKC $\iota/\lambda$  resulting in the cytoskeletal remodeling of vesicular tubular clusters between the endoplasmic reticulum and the Golgi complex. Specifically, phospho-GAPDH is required for microtubule binding to the vesicular tubular cluster sub-compartments that contains Rab2 (Tisdale 2002).

In addition to this, GAPDH also has the ability to auto-phosphorylate (Yogalingam et al. 2013). While it has been established that GAPDH will auto-phosphorylate in the presence of Mg-ATP, the recipient of this phosphate remains unknown (Yogalingam et al. 2013).

#### **6.2.2.2 Differential expression of phosphoproteins involved in translation**

Given the number of proteins involved in translation described in the label-free LC/MS analysis of the whole cell lysate (section 4.3.4), it is perhaps interesting to note that the organisation of the cytoskeleton has been implicated in regulating transcription (Kim and Coulombe 2010). Phosphoprotein enrichment also aided in the identification of several other translation factors including eIF4g1, 2-fold up-regulated upon temperature-shift.

A subunit of the scaffolding protein eIF4g in the eIF4f protein complex, eIF4g1 is involved in the selective translation of mRNAs involved in mitochondrial function (Fujioka et al. 2013).

The translation factor eIF4a3, a DEAD-box RNA helicase that forms part of the exon junction complex, facilitates the stable binding of spliced RNA enabling its modulation, was also identified as being 1.8-fold up-regulated at 31°C using this phosphoprotein enrichment strategy (Budiman et al. 2009).

### **6.2.3 Label-free LC/MS analysis of a Gallium oxide phosphopeptide enrichment of phosphoprotein-enriched CHO SEAP cells subject to temperature-shift**

While phosphoprotein enrichment contributed toward the identification of a significant number of phosphoproteins that would not otherwise have been observed, in order to obtain site-specific information on the phosphorylation status of the proteins identified a second enrichment step was required. This second enrichment was conducted at the peptide level to enrich for phosphopeptides using a gallium oxide column.

Overall the number of differentially-expressed phosphopeptides was lower than expected. This was almost certainly as a consequence of using a phosphoprotein enrichment step prior to phosphopeptide enrichment. This is evident from the qualitative analysis showing the significantly reduced number of phosphopeptides obtained from phosphoprotein-enriched samples compared to phosphopeptide enrichment of whole cell lysate samples using titanium or iron. In addition, it should be remembered that 120µg of phosphoprotein-enriched sample was analysed compared to 500µg of starting material for the titanium dioxide and iron phosphopeptide enrichment experiments. It is therefore possible that the quantity of starting material may also have impacted on the number of phosphopeptides obtained.

The reduced number of phosphopeptides is unlikely to be as a result of poor affinity for gallium as it has been shown to compare favourably with titanium, iron, aluminium and zirconium (Aryal and Ross 2010; Posewitz and Tempst 1999). Nonetheless this second enrichment strategy provided additional information on the abundance of phosphopeptides that had not been previously identified from analysis of the whole cell lysate or phosphoprotein-enriched fractions.

Gene Ontology (GO) of large profiling datasets using publically available software such as DAVID enables the identification of biological processes that have been enriched within the comparison undertaken (Gunaratne et al. 2010; Kumar and Mann 2009). However, this provides limited information on the specific pathways directly involved in the alteration of the biological processes observed. Network databases such as STRING or Ingenuity<sup>®</sup> Pathway Analysis can be used to help bridge this gap by mapping known and predicted protein interactions (Kumar and Mann 2009). Where smaller datasets are obtained, a ‘protein of interest’ may only be identified through searches of the literature. Whether the phosphoproteomic dataset is large or small however, our ability to interpret them still suffers from the sheer number of phosphorylation sites do not yet have any ascribed function (Mayya and Han 2009).

#### **6.2.3.1 PITSLRE serine/threonine protein kinase CDC2L1**

A phosphopeptide from PITSLRE serine/threonine protein kinase CDC2L1 (p58<sup>PITSLRE</sup>) was identified as being 7.4-fold down-regulated at 31°C. p58<sup>PITSLRE</sup> is a p34<sup>cdc2</sup>-related protein kinase that is known to be involved in cell cycle progression, tumorigenesis and pro-apoptotic signalling (Liu et al. 2013).

Although down-regulation of this kinase has been shown to increase growth rate and ectopic expression has been linked with prolonged late telophase, abnormal chromosome segregation, and decreased cell growth rates in CHO cells, this decreased growth rate was due to apoptosis (Lahti et al. 1995; Zhang et al. 2002).

While reduced growth rate is a feature of temperature-shifted culture conditions, it is also associated with extended culture viability (Kumar, Gammell, Clynes 2007). Given the tightly controlled nature of cell cycle progression the observed down-regulation of the phosphorylated peptide of p58<sup>PITSLRE</sup> may not necessarily result in increased growth under these conditions (Fisher et al. 2012).

### **6.2.3.2 Nucleolar phosphoprotein p130**

Gallium oxide phosphopeptide enrichment also enabled the identification of a phosphopeptide from the protein Nucleolar phosphoprotein p130 (Nopp140) which was deemed to be 2.2-fold up-regulated at 31°C. Interestingly this protein is one of the most highly phosphorylated proteins found in mammalian cells. Hyper-phosphorylated, Nopp140 is 140 kDa while in a completely unphosphorylated state its size is reduced to 95 kDa (Kim et al. 2003).

Initially identified as a shuttle protein between the nucleolus and cytoplasm involved in pre-ribosome particle assembly and transport, it is now known that Nopp140 also participates in rDNA transcription through co-localization with RNA polymerase I at the rDNA transcription active foci in the nucleolus (Gao et al. 2011b; Kim et al. 2003). This was further confirmed through the knockdown of Nopp140 which caused the inhibition of rDNA synthesis (Tsai et al. 2008).

While investigating the role of nuclear proteins which oscillate in the cell cycle, it was reported that endogenous levels of Nopp140 are significantly reduced during mitosis compared to interphase (Pai et al. 1995). Interestingly however, phosphorylation of Nopp140 actually increases during mitosis despite a reduction in overall expression (Pai et al. 1995).

In order to definitively determine the role that Nopp140 plays in temperature-shifted CHO SEAP cells, further analysis would be required to know whether the observed up-regulation of the Nopp140 phosphopeptide is as a result of an increase in phosphorylation or an increase in total protein abundance. It has been reported that CHO cells enter G<sub>1</sub> arrest upon temperature-shift (Moore et al. 1997).

Given the complex mechanism by which Nopp140 operates, it is likely that it could be involved in the regulation of cell function on multiple levels and further work would be required to elucidate its precise role in the cellular response to culture temperature reduction (Gao et al. 2011b).

One of the aims of this thesis was to identify changes that occur in the phosphoproteome of CHO cells subject to temperature-shift. Although the use of phosphoprotein enrichment did not initially provide any site-specific information, it nonetheless enabled the identification of proteins and their related processes such as

translation and cell structure that change in direct response to culture temperature reduction. While the number of phosphopeptides identified by Gallium enrichment was lower than expected, it did yield some information on the differential expression of phosphopeptides with site-specific information on proteins involved in translation and the cell cycle thus providing some knowledge of the signalling events that are activated in the phosphoproteome of a temperature-shifted CHO cell.

#### **6.2.4 Qualitative assessment of titanium dioxide and iron oxide phosphopeptide enrichment methods**

Given the nature of the results obtained in section 4.3.3 it was decided that phosphopeptide enrichment would provide more information on the changes taking place within the phosphoproteome of temperature-shifted CHO cells. Studies have shown that the use of different phosphopeptide enrichment methods, such as titanium dioxide and iron oxide, has provided increased coverage of the phosphoproteome (Bodenmiller et al. 2007; Carrascal et al. 2008; Collins, Yu, Choudhary 2007). Titanium and iron phosphopeptide enrichment strategies have also proven to be suitable for quantitative label-free LC/MS applications (Casado and Cutillas 2011; Montoya et al. 2011). In this study, titanium dioxide and iron oxide enrichment of temperature-shifted CHO whole cell lysates permitted the identification of 867 and 960 phosphopeptides respectively. The use of two complementary enrichment methods has provided a greater overview of the CHO phosphoproteome under temperature-shifted conditions enabling the cataloguing of 1,271 phosphopeptides and providing information on the differential expression of 83 phosphopeptides.

Quantitative label-free LC/MS analysis of phosphopeptide enriched whole cell lysates has been successfully employed to identify changes in phosphopeptide abundance in a number of studies (Casado and Cutillas 2011; Gunaratne et al. 2010; Manes et al. 2011). A gallium oxide phosphopeptide enrichment methodology was used in the identification of hormone dependent signalling pathways in renal medullary thick ascending limb (mTAL) cells (Gunaratne et al. 2010). Using this approach, 654 phosphopeptides were identified, of which 76 were differentially phosphorylated. Label-free analysis of titanium dioxide phosphopeptide enriched acute myeloid leukaemia cells treated with the sodium pervanadate resulted in the quantification of 937 phosphopeptides from a total of 1,537 phosphopeptides identified (Montoya et al. 2011). It should be remembered however that sodium pervanadate causes irreversible tyrosine phosphatase inhibition within the cell. This would result in an accumulation of hyper-phosphorylated tyrosine residues and is likely to increase the number of phosphopeptides identified (Boeri Erba et al. 2007).

During the course of this investigation a number of multiple phosphorylated peptide species were identified. Generally speaking, monophosphorylated peptides are more

commonly identified using Gallium and Titanium enrichment methods while Iron based enrichment favours multiply-phosphorylated species (Aryal, Olson, Ross 2008; Zhao et al. 2013). In this study however no such bias was observed. Both titanium and iron based columns enriched similar numbers of singly and multiply phosphorylated phosphopeptides.

Of the 960 phosphopeptides enriched using iron oxide, 471 of these triggered MS3 events while 537 MS3 events were triggered for the 867 phosphopeptides enriched using titanium dioxide. The neutral loss of a phosphate group is widely accepted as a form of ‘validation’ of the phosphorylated state of the phosphopeptide (as detailed section 1.5.7) (Yu et al. 2007). Although it was somewhat surprising that more MS3 events were not observed, there may be a number of reasons for this including the combined neutral loss of phosphate and water, and no dominant neutral loss of phosphate (Sweet et al. 2009). It has been reported that when using high mass accuracy instruments, such as the one used in this study, that MS3 offers no additional benefit in increasing the overall number of confidently identified phosphopeptides and has only limited value in determining the site of phosphorylation localization (Villén, Beausoleil, Gygi 2008). Nonetheless, the use of MS3 information in this study provides additional confidence in the identification of phosphopeptides particularly given the lack of information available on site-specific phosphorylation in CHO compared to other species such as Human and Mouse (Hornbeck et al. 2012).

This section demonstrates that the evaluation of a protocol for the enrichment of phosphopeptides in CHO cells, as laid out in the aims of this thesis, has been successfully completed and that the complementary use of titanium dioxide and iron phosphopeptide enrichment methods achieved the largest numbers of phosphopeptides with site-specific information.

### **6.2.5 Differentially-expressed phosphopeptides common to titanium dioxide and iron oxide phosphopeptide enrichment methods**

Using the complementary approach detailed in this study to obtain relative quantitative information on changes in the abundance of phosphopeptides offers a unique insight into the active response to reduced culture temperature conditions compared to other proteomic profiling studies that simply detail changes in the cellular machinery (Baik et al. 2006; Kaufmann et al. 1999; Kumar et al. 2008a; Underhill and Smales 2007). Of the 83 differentially-expressed phosphopeptides identified using both titanium dioxide and iron oxide enrichment, 29 were up-regulated and 54 down-regulated upon temperature-shift. While many differentially-expressed phosphopeptides were unique to each enrichment method, some of those identified were common to both techniques.

#### **6.2.5.1 eIF4g3**

One such protein was eIF4g3 (also called eIF4gII, but not to be confused with eIF4g2), was identified as 3.7-fold down-regulated at 31°C compared to 37°C. eIF4g3, along with eIF4g1, form the two functional homologs of eIF4g, part of the cap binding protein complex of eIF4f (Ivanov et al. 2011).

It has been shown that eIF4g3 becomes phosphorylated during mitosis inhibiting cap-dependent translation through a reduction in its interaction with eIF4e (Pyronnet, Dostie, Sonenberg 2001). In addition to phosphorylation, cleavage of eIF4g3, resulting in loss of function, is also known to cause inhibition of cap-dependent translation (Marcet-Palacios et al. 2011). These reports show the pivotal role that eIF4g3 plays in initiation of protein translation.

Transcriptomic analysis of B cell lymphocytes displaying increased proliferation and reduced Ig production from transgenic mice expressing a dominant active downstream regulatory element antagonist modulator (DREAM) mutation found that although increased proliferation was associated with a decrease in the expression of the cell cycle regulator Klf9, the reduction in protein synthesis was linked to the down-regulation of the eIF4g3 gene (Savignac et al. 2010).



Identified by both titanium dioxide and iron oxide phosphopeptide enrichment this investigation found phosphorylation of eIF4g3 to decrease upon temperature-shift. This would imply that a reduction in the phosphorylation of eIF4g3 would facilitate binding with eIF4e, thus enabling cap-dependent translation. Based on the reports discussed, this indicates that the reduction in the phosphorylation of eIF4g3 under temperature-shifted conditions may form a critical component of the mechanism involved in the increased productivity of temperature-shifted CHO cells.

#### **6.2.5.2 Protein NDRG1**

Another protein identified as having increased phosphorylation at 31°C by both titanium dioxide and iron oxide enrichment methodologies was protein NDRG1 (N-Myc down-regulated gene 1). NDRG1 is highly conserved across a number of multicellular organisms, typically expressed in response to cell stress and is known as a tumour metastasis suppressor in several different types of cancer including breast, colon, pancreatic and prostate (Ghalayini et al. 2013; McCaig et al. 2011).

Increased expression of NDRG1 has been shown to result in cell cycle arrest (Li and Chen 2011). The phosphorylation of NDRG1 also appears to be involved in the cell cycle arrest mechanism (Ghalayini et al. 2013; McCaig et al. 2011). In this study, we found NDRG1 to have increased phosphorylation upon temperature-shift. Although cell cycle arrest in temperature-shifted CHO cells has been attributed to the activation of the ATR-p53-p21 pathway (Roobol et al. 2011), cellular stress has been shown to result in the expression of NDRG1 in a p53 dependent manner (Roobol et al. 2011; Stein et al. 2004). This would suggest that NDRG1 may be involved in the cell cycle arrest of temperature-shifted CHO cells, however further work would be required to determine the precise mechanisms by which this would occur.

### **6.2.5.3 GO analysis of differentially-expressed phosphopeptides common to titanium dioxide and iron oxide phosphopeptide enrichment methods**

Expanded phosphoproteome coverage through the use of two enrichment methods was also of benefit for identifying key cellular processes and potential pathways that altered as a result of temperature-shift. For example, DAVID analysis of all the differentially-expressed phosphopeptides (Section 4.3.8.2) identified was the enrichment of the membrane organisation pathways at 31°C compared to 37°C. Interestingly, changes in the relative abundance of membrane proteins have been observed in another study which profiled the proteome of temperature-shift, non-producing CHO-K1 cells (Kumar et al. 2008a). Alterations in membrane rigidity have been cited as a means by which cells may detect the reduction of cell culture temperature (Roobol et al. 2011). Such alterations occur as a result of changes in membrane lipid composition causing membrane fluidity to decrease as temperature is reduced (Los and Murata 2004).

Changes in the abundance of phosphorylation on membrane organisation proteins are an encouraging indicator that active cell signalling events are taking place. DAVID analysis also identified the enrichment of proteins involved in cell division pathways, all of which were down-regulated at 31°C. This of course makes sense given the substantial reduction in cell growth under reduced culture temperature conditions. Among the proteins identified in this GO enrichment were Borealin (discussed in more detail in section 6.3.7.3), Anaphase-promoting complex subunit 1 and Serine/threonine-protein kinase Nek1, all of which showed decreased phosphorylation at 31°C compared to the 37°C culture.

#### **6.2.5.3.1 Serine/threonine-protein kinase Nek1**

Nek1 was identified as being 2.3-fold down-regulated in the temperature-shifted CHO SEAP cells. Although it is known that Nek1 is involved in DNA damage repair and cell cycle progression (Fry et al. 2012), reports of the precise mechanisms by which Nek1 operates and even its exact function are often conflicting. While the silencing of Nek1 has been reported to slow down the repair of DNA as a result of exposure to genotoxic agents and blocks DNA damage-induced cell cycle arrest (Pelegrini et al. 2010), it has also been shown that the loss of Nek1 caused a severe

decrease in cell proliferation, most likely due to increased time spent in the S-phase of the cell cycle (Patil et al. 2013). The role of Nek1 as a regulatory kinase for the proper activation of ATM/ATR target kinases Chk1 and Chk2 has been reported (Chen et al. 2008). It is not clear whether Nek1 regulation of Chk1 and Chk2 is as a result of direct or indirect interaction between the kinases.

Given that Nek1 can shuttle from the cytoplasm to the nucleus (Hilton, White, Quarmby 2009), it is possible that Nek1 also aids in the translocation of Chk1/Chk2 to the nuclei where they function (Chen et al. 2008). Although both Crk1 and Crk2 are both involved in DNA damage repair, it is Crk1 that plays the greater role in DNA damage repair while Chk2 is essential for G<sub>2</sub>-M arrest (Huang et al. 2008). Despite its involvement in DNA damage repair via the ATR pathway (Chen et al. 2008), it has also been proposed that Nek1 can regulate checkpoint control and DNA damage response in pathways independent of ATR (Chen et al. 2011). This serves as a prime example of the complexity of trying to decipher signalling pathways within the cell.

#### **6.2.5.3.2 Anaphase-promoting complex subunit 1**

Quantitative label-free LC/MS analysis of titanium dioxide and iron oxide phosphopeptide enrichments both found ANAPC1 to be down-regulated at 31°C by 1.5 and 1.9-fold respectively. ANAPC1 is the largest subunit of anaphase-promoting complex/cyclosome (APC/C), a ubiquitin ligase that targets specific proteins for proteasomal degradation during mitosis (Acquaviva et al. 2004; Jörgensen et al. 2001). Two known mitotic cyclins that the APC complex targets for degradation are cyclins A and B (Chang, Xu, Luo 2003).

It has been shown that cyclin B degradation is required for exit from mitosis (Skoufias et al. 2007). ANAPC1 acts as a scaffold protein that binds Cdc23, Apc5, and Apc4 (Thornton et al. 2006). Once this subunit is formed, it then goes on to bind Cdc16, Cdc27 and Cdc26. It has been suggested that activation of the complex is initiated by phosphorylation (Golan, Yudkovsky, Hershko 2002). In the context of this study the down-regulation of phosphorylated ANAPC1 would imply loss of proteasomal function and/or disassembly of the complex.

### **6.2.5.3.3 Differential expression of phosphopeptides in proteins involved in transcription**

Protein function analysis of all differentially-expressed phosphopeptides also identified a significant number of proteins involved in the regulation of transcription. This observation also makes sense given the changes that take place in the transcriptome of temperature-shifted CHO cells (Baik et al. 2006; Kantardjieff et al. 2010; Nissom et al. 2006; Yee, Gerdtzen, Hu 2009).

Among the proteins identified as having increased levels of phosphorylation was Cyclic AMP-dependent transcription factor ATF-7, zinc finger protein with KRAB and SCAN domains 3 and AT-rich interactive domain-containing protein 5B. Cyclic AMP-dependent transcription factor ATF-2, DNA excision repair protein ERCC-6 and transcription intermediary factor 1-beta displayed decreased levels of phosphorylation upon temperature-shift (some of these proteins are discussed further in section 6.4.6). Due to the limitations of our current knowledge of many of these proteins, whether the changes in phosphorylation will positively or negatively impact on transcription remains unclear.

The identification of signalling events involved in the regulation of transcription raises some interesting questions regarding the control of transcription in temperature-shifted cells; are such phosphorylation events involved in the increased transcription of the recombinant molecule, attenuating the translation of endogenous cellular proteins or in some way facilitating the transcription and/or binding of cold shock proteins to cis elements in the promoter regions of target genes leading to enhanced transcription? Furthering our understanding of these signalling mechanisms could provide cell engineering targets that would enable the preferential translation of specific gene targets.

### **6.2.5.3.4 TRIM28 and ZKSCAN3**

Identified as being 1.6-fold up-regulated at 31°C using titanium dioxide enrichment, Transcription intermediary factor 1-beta (TRIM28) is a transcriptional core-repressor protein involved in a wide variety of functions such as cell growth, apoptosis and DNA repair (Iyengar and Farnham 2011). The precise mechanisms by which TRIM28 operates are poorly understood and, given its upregulation in a variety of

cancers (such as breast cancer) and its down-regulated in other cancer types (including lung cancer), indicates that the role it plays in influencing cellular function appears to be cell line specific (Chen et al. 2012a). Despite this, it has been shown that phosphorylation of TRIM28 on Ser473 leads to a reduction in its activity (King 2013).

The increase in the phosphorylation TRIM28 at 31°C potentially points to a reduction in its activity at reduced culture temperatures. One of the means by which TRIM28 mediates transcriptional control is through interaction with Krüppel associated box (KRAB) zinc finger proteins (Fitzgerald et al. 2013).

Interestingly, Zinc finger protein with KRAB and SCAN domains 3 (ZKSCAN3) was identified as being 7.1-fold decreased at 31°C using iron oxide enrichment. It is possible that a reduction in the activity of TRIM28 could lead to a decrease in the activity of ZKSCAN3. This could make sense in the context of temperature-shift given that ZKSCAN3 has been reported as having increased expression in multiple myeloma, prostate and colorectal cancers and knockout of ZKSCAN 3 in two colon cancer cell lines has been shown to inhibit cell growth (Yang et al. 2010; Zhang et al. 2012), the exact function of ZKSCAN3 in CHO remains unknown and alterations in levels of phosphorylation may result in the activation of alternative pathways.

It must be remembered that phosphorylation cell signalling events are known to be very complex with many layers of regulation (Nguyen, Kolch, Kholodenko 2013). It is quite likely that the changes in phosphorylation observed could be as a result of other cell signalling events taking place. Of course these too may be of interest, but the fact remains that many of these phosphorylation events are uncharacterised and knowledge of their end function poorly understood (Hynes et al. 2013). It is therefore difficult to provide a definitive blueprint of all phosphorylation signalling events and their outcomes based on this investigation alone. Nonetheless proteins involved in cellular processes including transcription, translation, negative regulation of cell cycle progression and cytoskeletal organisation have been identified through the use of GO analysis. Further insights could be obtained through the integration of this data set with other published profiling studies to obtain a systems biology overview of the cellular response to temperature-shift.

Overlapping this data with transcriptomic (Clarke et al. 2011; Kantardjieff et al. 2010; Yee, Gerdtzen, Hu 2009) and proteomic (Baik et al. 2006; Kumar et al. 2008a; Meleady et al. 2011) studies of temperature-shifted CHO cultures could reveal regulatory connections and subnetworks that would not be apparent from the analysis of this phosphoproteomic dataset alone (Jünger and Aebersold 2014).

The data obtained in this study could also be used map the peptide sequences identified onto positions in the CHO genome that code for them, thus validating the expression of a given gene (Kumar and Mann 2009). This is particularly relevant given the recent release of the CHO genome (Xu et al. 2011), where many of the predicted proteins do not yet have any direct experimental information associated with them. The utilisation of the phosphoproteomic data in this thesis, along with data from other profiling proteomic studies (Baycin-Hizal et al. 2012; Lim et al. 2013; Meleady et al. 2011) could ultimately help refine the CHO genome model by providing evidence of gene expression at the protein level.

#### **6.2.6 Differentially-expressed phosphopeptides unique to titanium dioxide enrichment**

While the previous section discussed differentially-expressed phosphopeptides identified in both titanium dioxide and iron oxide enrichment strategies, this section will now describe some of the differentially-expressed phosphopeptides that were only obtained by titanium dioxide enrichment. Due to the nature of many of the proteins identified as a result of titanium dioxide enrichment, the precise role that phosphorylation / dephosphorylation of these proteins played in their function and the mechanisms which they became phosphorylated were often poorly understood. Functional and mechanistic information was however gathered on two phosphoproteins uniquely identified by titanium dioxide enrichment as having differentially-expressed phosphorylation.

#### **6.2.6.1 Programmed cell death 4 (PDCD4)**

2.1-fold up-regulated upon temperature-shift, PDCD4 is phosphorylated and regulated by AKT and has been shown to be up-regulated in lung tumours with slower growth rate while conversely being down-regulated in an aggressive, fast growing cancers such as skin cancer (Matsushashi et al. 2007; Palamarchuk et al. 2005). Although upregulation has also been associated with apoptosis, it is possible, given the complexity of apoptotic signalling pathways, that PDCD4 may in some way be involved in the reduced growth response to temperature-shift without impacting on cell viability (Matsushashi et al. 2007). Whether it is directly involved in growth reduction, or its upregulation occurs as a consequence of slowed growth rate remains unclear. The precise role of phosphorylated PDCD4 remains largely undefined, however it is known that following mitogen stimulation phosphorylation of PDCD4 by p70<sup>S6K1</sup> has been shown to result in its proteasomal degradation (Schmid et al. 2011). While degradation of PDCD4 would potentially lead to an increase in the growth of CHO cells, this is clearly not the case in those cells under reduced temperature conditions.

#### **6.2.6.2 Zyxin**

Zyxin, an actin regulatory protein that stabilizes sites of actin-membrane interaction, was identified as being 1.5-fold down-regulated at 31°C (Call et al. 2011). In addition to being found at focal adhesion sites, zyxin can also translocate to the nucleus; in this way zyxin regulates cell-cell adhesion, cytoskeleton organisation and signal transduction (Keicher et al. 2004; Nix and Beckerle 1997). Given its role as a 'mechanotransducer', zyxin is said to be involved in the active cellular respond to environmental conditions (Crone et al. 2011). Mutation studies inhibiting phosphorylation of zyxin by acinus-S resulted in the loss of its head tail release mechanism leaving a cell unable to disassociate itself from another cell or other substrate once bound (Call et al. 2011). Conversely, AKT-induced phosphorylation of zyxin enabling binding to acinus-S, thus preventing cleavage of acinus-S by caspases, results in inhibition of apoptotic chromatin condensation (Chan et al. 2007). The functional description provided of this phosphoprotein provides further evidence of active changes occurring in the cytoskeleton of temperature-shifted CHO cells.

### **6.2.7 Differentially-expressed phosphopeptides unique to iron oxide enrichment**

Label-free LC/MS analysis of iron oxide phosphopeptide enriched CHO SEAP cells enabled the identification of 60 differentially-expressed phosphopeptides of which 22 were up-regulated at 31°C. The following section outlines the role that phosphorylation plays in the function of some of the phosphoproteins uniquely identified by iron oxide enrichment.

#### **6.2.7.1 ATF-2 and ATF-7**

ATF-2 was identified as 7.4-fold down-regulated and ATF-7 was found to be 9.8-fold up-regulated at 31°C compared to 37°C. ATF-2 and ATF-7 are highly conserved members of the AP-1 transcription factor family (Diring et al. 2011). Cellular stress is known to induce both ATF-2 and ATF-7 (Bogoyevitch and Kobe 2006; Walczynski et al. 2013). This occurs via SAPK/JNK and p38 signalling cascades in the case of ATF-2 and only by p38, but not JNK, in the case of ATF-7 (Bogoyevitch and Kobe 2006; Walczynski et al. 2013).

While phosphorylation of ATF-2 can result in its degradation as well as induction of activity the observed decrease in phosphorylation of Thr-69 would appear to suggest that reduced culture temperature conditions results in a reduction of the active form of ATF-2 (Fuchs, Tappin, Ronai 2000). The increased phosphorylation of ATF-7 indicates that this transcription factor is likely to play a different role in the regulation of transcription to ATF-2. This is highly possible given the diverse range of functions that both ATF-2 and ATF-7 are known to be involved in (Diring et al. 2011).



### **6.2.7.2 DNA excision repair protein ERCC-6**

Phosphorylation of DNA excision repair protein ERCC-6 was decreased 2-fold at 31°C. DNA excision repair protein ERCC-6, a ubiquitin-regulated nucleotide excision repair factor, is involved in the repair of DNA damaged as a result of exposure to UV or oxidative stress (Baas et al. 2010). Oxidative stress has been shown to be reduced in cells grown under mildly hypothermic conditions (Slikker et al. 2001). This reduction in oxidative stress may therefore cause a down-regulation of oxidative stress induced DNA repair mechanisms.

### **6.2.7.3 Differential expression of phosphopeptides in proteins involved in cell division**

Girdin, increased 5.1-fold at 31°C, associates with actin and tubulin in the cytoskeleton to help regulate cell shape (Enomoto, Ping, Takahashi 2006; Liu et al. 2012). Silencing of girdin results in disassembly of stress filaments which, interestingly, zyxin also plays a role in the formation of (discussed section 6.3.6.2) (Kitamura et al. 2006). Phosphorylation of Girdin by AKT plays an important role in cytoskeletal reorganisation by regulating its association with the cell membrane. As such, girdin plays an essential role in cell division (Enomoto et al. 2005; Mao et al. 2012).

Another phosphoprotein known to be involved in cell division was Borealin. Decreased 3.3-fold at 31°C, Borealin is a regulator of chromosome alignment, spindle assembly and cytokinesis (Ruchaud, Carmena, Earnshaw 2007). Knockout of Borealin results in delayed mitotic progression, incorrect spindle assembly and restricts chromosome disassembly in anaphase (Gassmann et al. 2004). Expression of Borealin is cell cycle- dependent with levels peaking during the G2/M phases of the cell cycle (Date et al. 2012).

Upon exiting the M phase of the cell cycle, Borealin becomes dephosphorylated, however the precise phosphatase involved in this event is as yet unknown (Date et al. 2012; Kumar, Gammell, Clynes 2007) 3.3-fold down-regulation of Borealin phosphorylation is an interesting result; it provides further evidence of alterations in the cell cycle mechanisms of temperature-shifted CHO cells.

p58<sup>PITSLRE</sup> and Nek1 were among a number of proteins with differentially-expressed phosphopeptides that were identified as being involved in the regulation of cell cycle progression through GO analysis. Transcriptome analysis of temperature-shift and NaBu treated CHO cells has previously shown down-regulation of a number of genes involved in cell cycle regulation (Kantardjieff et al. 2010). Analysis of the phosphoproteomic data presented in this thesis shows the proteins involved in the cell cycle are evenly split between those with increased phosphorylation and those with decreased phosphorylation. While this is in contrast to the proteins involved in cell division, all of which displayed reduced phosphorylation, the fact that not all cell cycle related proteins became phosphorylated or dephosphorylated is likely to be as a result of the sheer complexity of cell cycle signalling events.

One of the main objectives of this thesis was to identify site-specific changes in the phosphorylation of phosphoproteins occurring in direct response to reduced culture temperature conditions. The identification of the differential expression of phosphopeptides in proteins such as eIF4g3, NDRG1, Zyxin and AFT-2 using a combination of titanium dioxide and iron phosphopeptide enrichment strategies shows that this has been achieved, thus providing further insights into the cellular response to reduced culture temperature.

#### **6.2.8 Utilisation of CHO specific proteomic databases in the analysis of LC/MS data in a phosphoproteomic study**

Until recently proteomic profiling studies of CHO cell lines relied on sequence homology between CHO and other well annotated species in a protein database such as human, mouse and rat, when searching raw mass spectral data (Baik et al. 2006; Kaufmann et al. 1999; Kumar et al. 2008a). This was due to the lack of CHO genome sequence information available, the absence of which was noted in a number of papers at the time (Griffin et al. 2007; Seth et al. 2007; Wlaschin and Hu 2007). Reliance on sequence homology means that potentially interesting proteins may go unidentified thus frustrating efforts to decipher signalling pathways (Meleady et al. 2012). Amino acid substitutions are common modifications found in proteins when direct comparisons are made of the same protein between different species (Liska and Shevchenko 2003; Meleady et al. 2012).

In terms of the detection of post translational modifications such as phosphorylation, correct amino acid sequence information is vital for peptide (and subsequent protein) identification. Accurate amino acid identification becomes even more critical when trying to obtain site-specific phosphorylation information, particularly in instances where a peptide contains a number of amino acids that may be phosphorylated as this requires the correct assignment of 'site determining ions' (Boersema, Mohammed, Heck 2009). The availability of two complementary CHO protein databases has undoubtedly been a significant asset to this investigation.

It has for the first time enabled the identification of 1,307 phosphopeptides and 1,480 phosphorylation sites in a CHO cell line. Reliance on sequence homology would almost certainly have offered reduced coverage of the phosphoproteome and limited the information obtained on the differential regulation of phosphorylation in temperature-shifted CHO cells.

### **6.2.9 Summary**

GO analysis of the 114 differentially-expressed proteins identified by quantitative label-free LC/MS analysis of CHO SEAP whole cell lysate samples showed enrichment of pathways involved in translation, RNA processing and ribosomal biogenesis while cytoskeleton organisation pathways were found to be enriched following similar analysis of phosphoprotein-enriched CHO samples.

Although gallium oxide phosphopeptide enrichment of phosphoprotein-enriched whole cell lysates provided site-specific phosphorylation information on a number of phosphopeptides, by far the greatest contribution came from titanium dioxide and iron oxide phosphopeptide enrichment of whole cell lysates. This strategy identified 82 differentially-expressed phosphopeptides (29 up-regulated and 53 down-regulated at 31°C compared to 37°C). GO analysis of the proteins identified revealed down-regulation of cell cycle and cell division pathways and upregulation of pathways involved in membrane organisation.

Further protein fractionation or validation by western blot would be required to definitively determine whether the changes observed are as a result of changes in phosphorylation.

## **Conclusions and Future Plans**

## Conclusions

### Chapter 1: Host Cell Protein

- This work shows that culture temperature and time in culture can result in the increased expression of enzymes with the potential to negatively impact on product quality. This included the identification of Proteases such as MMP19, MMP9 and Cathepsin D and the glycosidases Alpha-glucosidase and  $\alpha$ -N-acetyl-galactosaminidase in both cultures. In general, proteases were found to be up-regulated at the late phase in culture compared to early culture in both the IgG producing and non-producing CHO cell lines. The accumulation of proteases at this time point could negatively impact on product quality. Furthermore, proteases were identified as being down-regulated in the temperature shifted culture compared to the non-temperature-shifted culture indicating that the use of temperature shift could help reduce product loss in instances where products were susceptible to proteolytic activity.
- One important finding from this work was that significant differences in the HCP profile of a non-producing CHO-K1 cell line and an IgG secreting DP12 cell line. In addition, the reduction of culture temperature, was found to result in changes in the HCP profile of both IgG-secreting and non-secreting CHO cell lines and a non-secreting hybridoma cell line. This result has implications for the ability of HCP ELISA assays to detect contaminating HCPs in the final drug substance across diverse CHO cell lines and different culture conditions.
- The utilization of two CHO specific databases in this investigation enabled the identification of 1,136 HCPs produced by two CHO cell lines. The use of the BOKU- Bielefeld CHO specific proteomic database in addition to the NCBI CHO database in this profiling study enabled an additional 171 proteins to be identified in the secretome of the CHO-K1 cell line and an additional 162 were identified in the DP12 cell line.

Therefore, this study supports the use of both databases in the analysis of conditioned media samples from CHO cells as it increases proteome coverage and the level of confidence in the proteins identified. Ultimately, a merged database would be of value.

## **Chapter 2: Phosphoproteomic analysis of CHO cells subject to temperature shift**

- Phosphoproteomic profiling of CHO cells subject to temperature shift using antibody array based technology initially provided some interesting leads, possibly implicating the phosphorylation of Erk, Jnk and AKT in response to temperature shift. Subsequent follow up studies by western blot proved problematic due to non-specific binding and, in some instances, failure to identify the correct protein of interest. Thereby demonstrating the lack of hamster specific antibodies available to validate this work. Some targets, such as Erk, warrant further investigation based on the results obtained in this study, but it may be interesting to generate hamster specific antibodies as the anti-human antibodies available often gave poor results in CHO.
- GO analysis of differentially regulated phosphoproteins identified through quantitative label-free LC/MS analysis of phosphoprotein-enriched CHO SEAP samples indicated enrichment of pathways involved in cell cycle progression, protein biosynthesis and protein transport. In addition, GO analysis of differentially-expressed phosphopeptides identified following LC/MS analysis of titanium dioxide and iron oxide phosphopeptide enriched samples from temperature-shifted CHO whole cell lysate samples showed enrichment of pathways involved in transcription, translation, negative regulation of cell cycle progression and cytoskeletal organisation. Taken together, these results show some of the biological processes that are involved in the cellular response to temperature shift.

- Until the recent advent of CHO specific proteomic databases, reliable identification of site-specific phosphorylation was difficult. Identification was reliant on homology between the location of a specific phosphosite and the surrounding amino acid sequence in a CHO protein and its location in the same protein from another well annotated species. The use of two CHO specific data bases in this thesis has enabled greater coverage of the CHO phosphoproteome leading to the accurate identification of site-specific phosphorylation on 1,307 phosphopeptides (1,480 phosphosites in total).

## **Future work**

### **Proteomic profiling of Host Cell Proteins (HCPs) from CHO cells**

- An important finding of this work was the differential expression of proteases and glycosidases over time in culture and between temperature-shifted and non-temperature-shifted conditions. A follow up investigation to examine the impact that these enzymes have on product quality would be beneficial. Initially, zymography and glycosidase activity assays would be used to identify proteases and glycosidases that are functionally active in the conditioned media of CHO cells. A panel of well characterised recombinant therapeutics could be individually exposed to the enzymatic activity of each protease/glycosidase found to be active in the media of the CHO cell. Each 'model' product could then be analysed by LC/MS, capillary electrophoresis and SDS PAGE to assess the effect of the exposure to respective proteases/glycosidases on product integrity.

### **Phosphoproteomic characterisation of CHO cells subject to temperature shift**

- The phosphoproteomic profiling of CHO cells identified a number of phosphoproteins potentially involved in the response to temperature shift. Further insights into the cellular response to temperature shift could be obtained through the amalgamation of this data set with published datasets from other profiling studies. Overlapping the phosphoproteomic data in this thesis with datasets from transcriptomic (Clarke et al. 2011; Kantardjieff et al. 2010; Yee, Gerdtzen, Hu 2009) and proteomic (Baik et al. 2006; Kumar et al. 2008a; Meleady et al. 2011) analysis of temperature shifted CHO cells could help identify cell signalling mechanisms involved in the reduced growth rate and/or increased specific productivity seen in many temperature shifted CHO cell cultures.



## Bibliography

Acquaviva C, Herzog F, Kraft C, Pines J. 2004. The anaphase promoting complex/cyclosome is recruited to centromeres by the spindle assembly checkpoint. *Nat Cell Biol* 6(9):892-8.

Agrawal V and Bal M. 2012. Strategies for rapid production of therapeutic proteins in mammalian cells. *BioProcess Int* 10(4):32-48.

Ahlemann M, Zeidler R, Lang S, Mack B, Münz M, Gires O. 2006. Carcinoma-associated eIF3i overexpression facilitates mTOR-dependent growth transformation. *Mol Carcinog* 45(12):957-67.

Aldhamen YA, Seregin SS, Rastall DP, Aylsworth CF, Pepelyayeva Y, Busuito CJ, Godbehere-Roosa S, Kim S, Amalfitano A. 2013. Endoplasmic reticulum aminopeptidase-1 functions regulate key aspects of the innate immune response. *PloS One* 8(7):e69539.

Al-Fageeh MB and Smales CM. 2006. Control and regulation of the cellular responses to cold shock: The responses in yeast and mammalian systems. *Biochem J* 397(Pt 2):247.

Al-Fageeh MB and Smales CM. 2013. Alternative promoters regulate cold inducible RNA-binding (CIRP) gene expression and enhance transgene expression in mammalian cells. *Mol Biotechnol* 54(2):238-49.

Al-Fageeh MB, Marchant RJ, Carden MJ, Smales CM. 2006. The cold-shock response in cultured mammalian cells: Harnessing the response for the improvement of recombinant protein production. *Biotechnol Bioeng* 93(5):829-35.

Altamirano C, Berrios J, Vergara M, Becerra S. 2013. Advances in improving mammalian cells metabolism for recombinant protein production. *EJB* 16(3):10-.

Altelaar AM, Munoz J, Heck AJ. 2013. Next-generation proteomics: Towards an integrative view of proteome dynamics. *Nature Reviews Genetics* 14(1):35-48.

America AHP and Cordewener JHG. 2008. Comparative LC-MS: A landscape of peaks and valleys. *Proteomics* 8(4):731-49.

Angel TE, Aryal UK, Hengel SM, Baker ES, Kelly RT, Robinson EW, Smith RD. 2012. Mass spectrometry-based proteomics: Existing capabilities and future directions. *Chem Soc Rev* 41(10):3912-28.

Annes JP, Munger JS, Rifkin DB. 2003. Making sense of latent TGF $\beta$  activation. *J Cell Sci* 116(2):217-24.

Arachea BT, Sun Z, Potente N, Malik R, Isailovic D, Viola RE. 2012. Detergent selection for enhanced extraction of membrane proteins. *Protein Expr Purif* 86(1):12-20.

Arden N and Betenbaugh MJ. 2004. Life and death in mammalian cell culture: Strategies for apoptosis inhibition. *Trends Biotechnol* 22(4):174-80.

- Arima K, Kinoshita A, Mishima H, Kanazawa N, Kaneko T, Mizushima T, Ichinose K, Nakamura H, Tsujino A, Kawakami A. 2011. Proteasome assembly defect due to a proteasome subunit beta type 8 (PSMB8) mutation causes the autoinflammatory disorder, nakajo-nishimura syndrome. *Proceedings of the National Academy of Sciences* 108(36):14914-9.
- Armirotti A and Damonte G. 2010. Achievements and perspectives of top-down proteomics. *Proteomics* 10(20):3566-76.
- Aryal UK and Ross AR. 2010. Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry. *Rapid Communications in Mass Spectrometry* 24(2):219-31.
- Aryal UK, Olson DJ, Ross AR. 2008. Optimization of immobilized gallium (III) ion affinity chromatography for selective binding and recovery of phosphopeptides from protein digests. *Journal of Biomolecular Techniques: JBT* 19(5):296.
- Aumiller JJ, Hollister JR, Jarvis DL. 2006. Molecular cloning and functional characterization of  $\beta$ -N-acetylglucosaminidase genes from Sf9 cells. *Protein Expr Purif* 47(2):571-90.
- Baas DC, Despriet DD, Gorgels TG, Bergeron-Sawitzke J, Uitterlinden AG, Hofman A, van Duijn CM, Merriam JE, Smith RT, Barile GR. 2010. The ERCC6 gene and age-related macular degeneration. *PloS One* 5(11):e13786.
- Bahr SM, Borgschulte T, Kayser KJ, Lin N. 2009. Using microarray technology to select housekeeping genes in chinese hamster ovary cells. *Biotechnol Bioeng* 104(5):1041-6.
- Baik JY, Joo EJ, Kim YH, Lee GM. 2008. Limitations to the comparative proteomic analysis of thrombopoietin producing chinese hamster ovary cells treated with sodium butyrate. *J Biotechnol* 133(4):461-8.
- Baik JY, Lee MS, An SR, Yoon SK, Joo EJ, Kim YH, Park HW, Lee GM. 2006. Initial transcriptome and proteome analyses of low culture temperature-induced expression in CHO cells producing erythropoietin. *Biotechnol Bioeng* 93(2):361-71.
- Baik JY, Ha TK, Kim YH, Lee GM. 2011. Proteomic understanding of intracellular responses of recombinant chinese hamster ovary cells adapted to grow in serum-free suspension culture. *Biotechnol Prog* 27(6):1680-8.
- Barnes LM and Dickson AJ. 2006. Mammalian cell factories for efficient and stable protein expression. *Curr Opin Biotechnol* 17(4):381-6.
- Barron N, Sanchez N, Kelly P, Clynes M. 2011. MicroRNAs: Tiny targets for engineering CHO cell phenotypes? *Biotechnol Lett* 33(1):11-21.
- Baxter RC and Martin JL. 1989. Binding proteins for the insulin-like growth factors: Structure, regulation and function. *Prog Growth Factor Res* 1(1):49-68.
- Baycin-Hizal D, Tabb DL, Chaerkady R, Chen L, Lewis NE, Nagarajan H, Sarkaria V, Kumar A, Wolozny D, Colao J. 2012. Proteomic analysis of chinese hamster ovary cells. *Journal of Proteome Research* 11(11):5265-76.

- Beatson R, Sproviero D, Maher J, Wilkie S, Taylor-Papadimitriou J, Burchell JM. 2011. Transforming growth factor-beta 1 is constitutively secreted by chinese hamster ovary cells and is functional in human cells. *Biotechnol Bioeng* 108(11):2759-64.
- Beausoleil H, Labrie V, Dubreuil JD. 2002. Trypan blue uptake by chinese hamster ovary cultured epithelial cells: A cellular model to study *Escherichia coli* STb enterotoxin. *Toxicon* 40(2):185-91.
- Beausoleil SA, Villén J, Gerber SA, Rush J, Gygi SP. 2006. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat Biotechnol* 24(10):1285-92.
- Beausoleil SA, Jedrychowski M, Schwartz D, Elias JE, Villen J, Li J, Cohn MA, Cantley LC, Gygi SP. 2004. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A* 101(33):12130-5.
- Beltran L and Cutillas PR. 2012. Advances in phosphopeptide enrichment techniques for phosphoproteomics. *Amino Acids* 43(3):1009-24.
- Berger J, Hauber J, Hauber R, Geiger R, Cullen BR. 1988. Secreted placental alkaline phosphatase: A powerful new quantitative indicator of gene expression in eukaryotic cells. *Gene* 66(1):1-10.
- Blagoev B, Ong S, Kratchmarova I, Mann M. 2004. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nat Biotechnol* 22(9):1139-45.
- Blom N, Kreegipuu A, Brunak S. 1998. PhosphoBase: A database of phosphorylation sites. *Nucleic Acids Res* 26(1):382-6.
- Bodenmiller B, Mueller LN, Mueller M, Domon B, Aebersold R. 2007. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nature Methods* 4(3):231-7.
- Boeri Erba E, Matthiesen R, Bunkenborg J, Schulze WX, Di Stefano P, Cabodi S, Tarone G, Defilippi P, Jensen ON. 2007. Quantitation of multisite EGF receptor phosphorylation using mass spectrometry and a novel normalization approach. *Journal of Proteome Research* 6(7):2768-85.
- Boersema PJ, Mohammed S, Heck AJ. 2009. Phosphopeptide fragmentation and analysis by mass spectrometry. *Journal of Mass Spectrometry* 44(6):861-78.
- Bogoyevitch MA and Kobe B. 2006. Uses for JNK: The many and varied substrates of the c-jun N-terminal kinases. *Microbiology and Molecular Biology Reviews* 70(4):1061-95.
- Boja ES and Fales HM. 2001. Overalkylation of a protein digest with iodoacetamide. *Anal Chem* 73(15):3576-82.
- Bollati-Fogolín M, Forno G, Nimtz M, Conradt HS, Etcheverrigaray M, Kratje R. 2005. Temperature reduction in cultures of hGM-CSF-expressing CHO cells: Effect on productivity and product quality. *Biotechnol Prog* 21(1):17-21.

Borissenko L and Groll M. 2007. 20S proteasome and its inhibitors: Crystallographic knowledge for drug development. *Chem Rev* 107(3):687-717.

Borth N, Mattanovich D, Kunert R, Katinger H. 2005. Effect of increased expression of protein disulfide isomerase and heavy chain binding protein on antibody secretion in a recombinant CHO cell line. *Biotechnol Prog* 21(1):106-11.

Brochu C, Cabrita MA, Melanson BD, Hamill JD, Lau R, Pratt MC, McKay BC. 2013. NF- $\kappa$ B-dependent role for cold-inducible RNA binding protein in regulating interleukin 1 $\beta$ . *PloS One* 8(2):e57426.

Brownridge P and Beynon RJ. 2011. The importance of the digest: Proteolysis and absolute quantification in proteomics. *Methods* 54(4):351-60.

Budiman ME, Bubenik JL, Miniard AC, Middleton LM, Gerber CA, Cash A, Driscoll DM. 2009. Eukaryotic initiation factor 4a3 is a selenium-regulated RNA-binding protein that selectively inhibits selenocysteine incorporation. *Mol Cell* 35(4):479-89.

Bulavin DV, Saito S, Hollander MC, Sakaguchi K, Anderson CW, Appella E, Fornace AJ. 1999. Phosphorylation of human p53 by p38 kinase coordinates N-terminal phosphorylation and apoptosis in response to UV radiation. *EMBO J* 18(23):6845-54.

Busby WH, Nam TJ, Moralez A, Smith C, Jennings M, Clemmons DR. 2000. The complement component C1s is the protease that accounts for cleavage of insulin-like growth factor-binding protein-5 in fibroblast medium. *J Biol Chem* 275(48):37638.

Call GS, Chung JY, Davis JA, Price BD, Primavera TS, Thomson NC, Wagner MV, Hansen MD. 2011. Zyxin phosphorylation at serine 142 modulates the zyxin head-tail interaction to alter cell-cell adhesion. *Biochem Biophys Res Commun* 404(3):780-4.

Carlage T, Hincapie M, Zang L, Lyubarskaya Y, Madden H, Mhatre R, Hancock WS. 2009. Proteomic profiling of a high-producing chinese hamster ovary cell culture. *Anal Chem* 81(17):7357-62.

Carlage T, Kshirsagar R, Zang L, Janakiraman V, Hincapie M, Lyubarskaya Y, Weiskopf A, Hancock WS. 2012. Analysis of dynamic changes in the proteome of a bcl-XL overexpressing chinese hamster ovary cell culture during exponential and stationary phases. *Biotechnol Prog* 28(3):814-23.

Carrascal M, Ovelheiro D, Casas V, Gay M, Abian J. 2008. Phosphorylation analysis of primary human T lymphocytes using sequential IMAC and titanium oxide enrichment. *Journal of Proteome Research* 7(12):5167-76.

Casado P and Cutillas PR. 2011. A self-validating quantitative mass spectrometry method for assessing the accuracy of high-content phosphoproteomic experiments. *Molecular & Cellular Proteomics* 10(1).

Chalkley RJ and Clauser KR. 2012. Modification site localization scoring: Strategies and performance. *Mol Cell Proteomics* 11(5):3-14.

- Champion K, Madden H, Dougherty J, Shacter E. 2005. Defining your product profile and maintaining control over it, part 2. *BioProcess Int* 3:52-7.
- Champion KM, Arnott D, Henzel WJ, Hermes S, Weikert S, Stults J, Vanderlaan M, Krummen L. 1999. A two-dimensional protein map of chinese hamster ovary cells. .
- Chan C, Liu X, Tang X, Fu H, Ye K. 2007. Akt phosphorylation of zyxin mediates its interaction with acinus-S and prevents acinus-triggered chromatin condensation. *Cell Death & Differentiation* 14(9):1688-99.
- Chang DC, Xu N, Luo KQ. 2003. Degradation of cyclin B is required for the onset of anaphase in mammalian cells. *J Biol Chem* 278(39):37865-73.
- Chao MP, Jaiswal S, Weissman-Tsukamoto R, Alizadeh AA, Gentles AJ, Volkmer J, Weiskopf K, Willingham SB, Raveh T, Park CY, et al. 2010. Calreticulin is the dominant pro-phagocytic signal on multiple human cancers and is counterbalanced by CD47. *Sci Transl Med* 2(63):63ra94.
- Chappell SA, Owens GC, Mauro VP. 2001. A 5' leader of Rbm3, a cold stress-induced mRNA, mediates internal initiation of translation with increased efficiency under conditions of mild hypothermia. *J Biol Chem* 276(40):36917-22.
- Chelius D and Bondarenko PV. 2002. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *Journal of Proteome Research* 1(4):317-23.
- Chen EI, Cociorva D, Norris JL, Yates JR. 2007. Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *Journal of Proteome Research* 6(7):2529-38.
- Chen L, Chen D, Kurtyka C, Rawal B, Fulp WJ, Haura EB, Cress WD. 2012a. Tripartite motif containing 28 (Trim28) can regulate cell proliferation by bridging HDAC1/E2F interactions. *J Biol Chem* 287(48):40106-18.
- Chen S, Kung M, Hu T, Chen H, Wu J, Kuo H, Tsai H, Lin Y, Wen Z, Liu J. 2012b. Hepatoma-derived growth factor regulates breast cancer cell invasion by modulating epithelial–mesenchymal transition. *J Pathol* 228(2):158-69.
- Chen Y, Chen P, Chen C, Jiang X, Riley DJ. 2008. Never-in-mitosis related kinase 1 functions in DNA damage response and checkpoint control. *Cell Cycle* 7(20):3194-201.
- Chen Y, Chen CF, Riley DJ, Chen PL. 2011. Nek1 kinase functions in DNA damage response and checkpoint control through a pathway independent of ATM and ATR. *Cell Cycle* 10(4):655-63.
- Chen YX, Li Y, Wang WM, Zhang W, Chen XN, Xie YY, Lu J, Huang QH, Chen N. 2010. Phosphoproteomic study of human tubular epithelial cell in response to transforming growth factor-beta-1-induced epithelial-to-mesenchymal transition. *Am J Nephrol* 31(1):24-35.
- Chen Z, Southwick K, Thulin CD. 2004. Initial analysis of the phosphoproteome of chinese hamster ovary cells using electrophoresis. *J Biomol Tech* 15(4):249-56.

Chen Z, Wu B, Liu H, Liu X, Huang P. 2004. Temperature shift as a process optimization step for the production of pro-urokinase by a recombinant chinese hamster ovary cell line in high-density perfusion culture. *Journal of Bioscience and Bioengineering* 97(4):239-43.

Chenau J, Michelland S, de Fraipont F, Josserand V, Coll JL, Favrot MC, Seve M. 2009. The cell line secretome, a suitable tool for investigating proteins released in vivo by tumors: Application to the study of p53-modulated proteins secreted in lung cancer cells. *Journal of Proteome Research* 8(10):4579-91.

Chiang GG and Sisk WP. 2005. Bcl-xL mediates increased production of humanized monoclonal antibodies in chinese hamster ovary cells. *Biotechnol Bioeng* 91(7):779-92.

Chiba S. 1997. Molecular mechanism in alpha-glucosidase and glucoamylase. *Biosci Biotechnol Biochem* 61(8):1233-9.

Chip S, Zelmer A, Ogunshola OO, Felderhoff-Mueser U, Nitsch C, Bühner C, Wellmann S. 2011. The RNA-binding protein RBM3 is involved in hypothermia induced neuroprotection. *Neurobiol Dis* 43(2):388-96.

Chitteti BR and Peng Z. 2007. Proteome and phosphoproteome dynamic change during cell dedifferentiation in arabidopsis. *Proteomics* 7(9):1473-500.

Choe L, D'Ascenzo M, Relkin NR, Pappin D, Ross P, Williamson B, Guertin S, Pribil P, Lee KH. 2007. 8-Plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for alzheimer's disease. *Proteomics* 7(20):3651-60.

Chu EH. 2004. Early days of mammalian somatic cell genetics: The beginning of experimental mutagenesis. *Mutation Research/Reviews in Mutation Research* 566(1):1-8.

Chu L and Robinson DK. 2001. Industrial choices for protein production by large-scale cell culture. *Curr Opin Biotechnol* 12(2):180-7.

Chun C, Heineken K, Szeto D, Ryll T, Chamow S, Chung JD. 2003. Application of factorial design to accelerate identification of CHO growth factor requirements. *Biotechnol Prog* 19(1):52-7.

Chung JY, Lim SW, Hong YJ, Hwang SO, Lee GM. 2004. Effect of doxycycline-regulated calnexin and calreticulin expression on specific thrombopoietin productivity of recombinant chinese hamster ovary cells. *Biotechnol Bioeng* 85(5):539-46.

Chuppa S, Tsai YS, Yoon S, Shackelford S, Rozales C, Bhat R, Tsay G, Matanguihan C, Konstantinov K, Naveh D. 1997. Fermentor temperature as a tool for control of high-density perfusion cultures of mammalian cells. *Biotechnol Bioeng* 55(2):328-38.

Clark KJR, Chaplin FWR, Harcum SW. 2004. Temperature effects on Product-Quality-Related enzymes in batch CHO cell cultures producing recombinant tPA. *Biotechnol Prog* 20(6):1888-92.

- Clarke C, Doolan P, Barron N, Meleady P, O'Sullivan F, Gammell P, Melville M, Leonard M, Clynes M. 2011. Large scale microarray profiling and coexpression network analysis of CHO cells identifies transcriptional modules associated with growth and productivity. *J Biotechnol* 155(3):350-9.
- Cohen P. 2001. The role of protein phosphorylation in human health and disease. *European Journal of Biochemistry* 268(19):5001-10.
- Collins MO, Yu L, Choudhary JS. 2007. Analysis of protein phosphorylation on a proteome-scale. *Proteomics* 7(16):2751-68.
- Collins MO, Wright JC, Jones M, Rayner JC, Choudhary JS. 2014. Confident and sensitive phosphoproteomics using combinations of collision induced dissociation and electron transfer dissociation. *Journal of Proteomics* 103:1-14.
- Cong W, Hwang S, Jin L, Choi J. 2008. Sensitive fluorescent staining for proteomic analysis of proteins in 1-D and 2-D SDS-PAGE and its comparison with SYPRO ruby by PMF. *Electrophoresis* 29(21):4304-15.
- Conus S and Simon H. 2010. Cathepsins and their involvement in immune responses. *Swiss Med Wkly* 140:w13042.
- Cooper B. 2011. The problem with peptide presumption and low mascot scoring. *Journal of Proteome Research* 10(3):1432-5.
- Cost GJ, Freyvert Y, Vafiadis A, Santiago Y, Miller JC, Rebar E, Collingwood TN, Snowden A, Gregory PD. 2010. BAK and BAX deletion using zinc-finger nucleases yields apoptosis-resistant CHO cells. *Biotechnol Bioeng* 105(2):330-40.
- Cottrell JS and London U. 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20(18):3551-67.
- Cox J and Mann M. 2011. Quantitative, high-resolution proteomics for data-driven systems biology. *Annu Rev Biochem* 80:273-99.
- Cox J and Mann M. 2007. Is proteomics the new genomics? *Cell* 130(3):395-8.
- Crone J, Glas C, Schultheiss K, Moehlenbrink J, Krieghoff-Henning E, Hofmann TG. 2011. Zyxin is a critical regulator of the apoptotic HIPK2-p53 signaling axis. *Cancer Res* 71(6):2350-9.
- Cui X, Hawari F, Alsaaty S, Lawrence M, Combs CA, Geng W, Rouhani FN, Miskinis D, Levine SJ. 2002. Identification of ARTS-1 as a novel TNFR1-binding protein that promotes TNFR1 ectodomain shedding. *J Clin Invest* 110(4):515-26.
- D'Ambrosio C, Salzano AM, Arena S, Renzone G, Scaloni A. 2007. Analytical methodologies for the detection and structural characterization of phosphorylated proteins. *Journal of Chromatography B* 849(1):163-80.
- Date D, Dreier MR, Borton MT, Bekier ME, Taylor WR. 2012. Effects of phosphatase and proteasome inhibitors on borealin phosphorylation and degradation. *J Biochem* 151(4):361-9.

- Datta P, Linhardt RJ, Sharfstein ST. 2013. An'omics approach towards CHO cell engineering. *Biotechnol Bioeng* 110(5):1255-71.
- De Jesus M and Wurm FM. 2011. Manufacturing recombinant proteins in kg-ton quantities using animal cells in bioreactors. *European Journal of Pharmaceutics and Biopharmaceutics* 78(2):184-8.
- de Muga SV, Timpson P, Cubells L, Evans R, Hayes T, Rentero C, Hegemann A, Reverter M, Leschner J, Pol A. 2008. Annexin A6 inhibits ras signalling in breast cancer cells. *Oncogene* 28(3):363-77.
- de Nobel H, Lawrie L, Brul S, Klis F, Davis M, Alloush H, Coote P. 2001. Parallel and comparative analysis of the proteome and transcriptome of sorbic acid-stressed *saccharomyces cerevisiae*. *Yeast* 18(15):1413-28.
- Deaven L and Petersen D. 1973. The chromosomes of CHO, an aneuploid chinese hamster cell line: G-band, C-band, and autoradiographic analyses. *Chromosoma* 41(2):129-44.
- Dephoure N, Gould KL, Gygi SP, Kellogg DR. 2013. Mapping and analysis of phosphorylation sites: A quick guide for cell biologists. *Mol Biol Cell* 24(5):535-42.
- Deutsch EW. 2010. Mass spectrometer output file format mzML. In: *Proteome bioinformatics*. Springer. 319 p.
- Deutsch EW. 2012. File formats commonly used in mass spectrometry proteomics. *Mol Cell Proteomics* 11(12):1612-21.
- Dick Jr LW, Qiu D, Mahon D, Adamo M, Cheng KC. 2008. C-terminal lysine variants in fully human monoclonal antibodies: Investigation of test methods and possible causes. *Biotechnol Bioeng* 100(6):1132-43.
- Dietmair S, Nielsen LK, Timmins NE. 2011. Engineering a mammalian super producer. *Journal of Chemical Technology and Biotechnology* 86(7):905-14.
- Dinkel H, Chica C, Via A, Gould CM, Jensen LJ, Gibson TJ, Diella F. 2011. Phospho.ELM: A database of phosphorylation sites--update 2011. *Nucleic Acids Res* 39(Database issue):D261-7.
- Dinnis DM, Stansfield SH, Schlatter S, Smales CM, Alete D, Birch JR, Racher AJ, Marshall CT, Nielsen LK, James DC. 2006. Functional proteomic analysis of GS-NS0 murine myeloma cell lines with varying recombinant monoclonal antibody production rate. *Biotechnol Bioeng* 94(5):830-41.
- Diring J, Camuzeaux B, Donzeau M, Vigneron M, Rosa-Calatrava M, Kedinger C, Chatton B. 2011. A cytoplasmic negative regulator isoform of ATF7 impairs ATF7 and ATF2 phosphorylation and transcriptional activity. *PloS One* 6(8):e23351.
- Dix MM, Simon GM, Wang C, Okerberg E, Patricelli MP, Cravatt BF. 2012. Functional interplay between caspase cleavage and phosphorylation sculpts the apoptotic proteome. *Cell* 150(2):426-40.



Dokduang H, Juntana S, Techasen A, Namwat N, Yongvanit P, Khuntikeo N, Riggins GJ, Loilome W. 2013. Survey of activated kinase proteins reveals potential targets for cholangiocarcinoma treatment. *Tumor Biol* 34(6):3519-28.

Doneanu Catalin, Xenopoulos Alex, Fadgen Keith, Murphy Jim, Prentice Holly, Stapels Martha and Chen Weibin. 2012. Analysis of host-cell proteins in biotherapeutic proteins by comprehensive online two-dimensional liquid chromatography/mass spectrometry. *MAbsLandes Bioscience*. 24 p.

Dong W, Wang T, Wang F, Zhang J. 2011. Simple, time-saving dye staining of proteins for sodium dodecyl sulfate–polyacrylamide gel electrophoresis using coomassie blue. *PloS One* 6(8):e22394.

Dong Z, Qi J, Peng H, Liu J, Zhang JT. 2013. Spectrin domain of eukaryotic initiation factor 3a is the docking site for formation of the a:B:I:G subcomplex. *J Biol Chem* 288(39):27951-9.

Donoghue PM, Hughes C, Vissers JPC, Langridge JI, Dunn MJ. 2008. Nonionic detergent phase extraction for the proteomic analysis of heart membrane proteins using label-free LC-MS. *Proteomics* 8(18):3895-905.

Doolan P, Meleady P, Barron N, Henry M, Gallagher R, Gammell P, Melville M, Sinacore M, McCarthy K, Leonard M. 2010. Microarray and proteomics expression profiling identifies several candidates, including the valosin-containing protein (VCP), involved in regulating high cellular growth rate in production CHO cell lines. *Biotechnol Bioeng* 106(1):42-56.

Dorai H, Santiago A, Campbell M, Tang QM, Lewis MJ, Wang Y, Lu Q, Wu S, Hancock W. 2011. Characterization of the proteases involved in the N-terminal clipping of glucagon-like-peptide-1-antibody fusion proteins. *Biotechnol Prog* 27(1):220-31.

Douillard-Guilloux G, Mouly V, Caillaud C, Richard E. 2009. Immortalization of murine muscle cells from lysosomal  $\alpha$ -glucosidase deficient mice: A new tool to study pathophysiology and assess therapeutic strategies for pompe disease. *Biochem Biophys Res Commun* 388(2):333-8.

Dowell JA, Johnson JA, Li L. 2009. Identification of astrocyte secreted proteins with a combination of shotgun proteomics and bioinformatics. *Journal of Proteome Research* 8(8):4135-43.

Dresios J, Aschrafi A, Owens GC, Vanderklish PW, Edelman GM, Mauro VP. 2005. Cold stress-induced protein Rbm3 binds 60S ribosomal subunits, alters microRNA levels, and enhances global protein synthesis. *Proc Natl Acad Sci U S A* 102(6):1865-70.

Dulla K, Daub H, Hornberger R, Nigg EA, Korner R. 2010. Quantitative site-specific phosphorylation dynamics of human protein kinases during mitotic progression. *Mol Cell Proteomics* 9(6):1167-81.

Dumaz N and Meek DW. 1999. Serine 15 phosphorylation stimulates p53 transactivation but does not directly influence interaction with HDM2. *EMBO J* 18(24):7002-10.

- Dutton RL, Scharer JM, Moo-Young M. 1999. Hybridoma growth and productivity: Effects of conditioned medium and of inoculum size. *Cytotechnology* 29(1):1-10.
- Eaton LC. 1995. Host cell contaminant protein assay development for recombinant biopharmaceuticals. *Journal of Chromatography A* 705(1):105-14.
- Ebisuya M, Kondoh K, Nishida E. 2005. The duration, magnitude and compartmentalization of ERK MAP kinase activity: Mechanisms for providing signaling specificity. *J Cell Sci* 118(14):2997-3002.
- Elliott P, Hohmann A, Spanos J. 2003. Protease expression in the supernatant of chinese hamster ovary cells grown in serum-free culture. *Biotechnol Lett* 25(22):1949-52.
- Eng JK, McCormack AL, Yates JR. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 5(11):976-89.
- Engholm-Keller K and Larsen MR. 2013. Technologies and challenges in large-scale phosphoproteomics. *Proteomics* 13(6):910-31.
- Enomoto A, Ping J, Takahashi M. 2006. Girdin, a novel Actin-Binding protein, and its family of proteins possess versatile functions in the akt and wnt signaling pathways. *Ann N Y Acad Sci* 1086(1):169-84.
- Enomoto A, Murakami H, Asai N, Morone N, Watanabe T, Kawai K, Murakumo Y, Usukura J, Kaibuchi K, Takahashi M. 2005. Akt/PKB regulates actin organization and cell motility via Girdin/APE. *Developmental Cell* 9(3):389-402.
- Eon-Duval A, Broly H, Gleixner R. 2012. Quality attributes of recombinant therapeutic proteins: An assessment of impact on safety and efficacy as part of a quality by design development approach. *Biotechnol Prog* 28(3):608-22.
- Eriksson U, Hassel J, Lullau E, Haggstrom L. 2005. Metalloproteinase activity is the sole factor responsible for the growth-promoting effect of conditioned medium in trichoplusia ni insect cell cultures. *J Biotechnol* 119(1):76-86.
- Ermolenko D and Makhatadze G. 2002. Bacterial cold-shock proteins. *Cellular and Molecular Life Sciences CMLS* 59(11):1902-13.
- Even MS, Sandusky CB, Barnard ND. 2006. Serum-free hybridoma culture: Ethical, scientific and safety considerations. *Trends Biotechnol* 24(3):105-8.
- Eymann C, Becher D, Bernhardt J, Gronau K, Klutzny A, Hecker M. 2007. Dynamics of protein phosphorylation on Ser/Thr/Tyr in bacillus subtilis. *Proteomics* 7(19):3509-26.
- Eyrich B, Sickmann A, Zahedi RP. 2011. Catch me if you can: Mass spectrometry-based phosphoproteomics and quantification strategies. *Proteomics* 11(4):554-70.
- Fan F, Wey JS, McCarty MF, Belcheva A, Liu W, Bauer TW, Somcio RJ, Wu Y, Hooper A, Hicklin DJ. 2005. Expression and function of vascular endothelial growth factor receptor-1 on human colorectal cancer cells. *Oncogene* 24(16):2647-53.

- Farady CJ, Sellers BD, Jacobson MP, Craik CS. 2009. Improving the species cross-reactivity of an antibody using computational design. *Bioorg Med Chem Lett* 19(14):3744-7.
- Farajollahi MM, Cook DB, Hamzehlou S, Self CH. 2012. Reduction of non-specific binding in immunoassays requiring long incubations. *Scandinavian Journal of Clinical & Laboratory Investigation* 72(7):531-9.
- Fernald A, Pisania A, Abdelgadir E, Milling J, Marvell T, Steininger B. 2009. Scale-up and comparison studies evaluating disposable bioreactors and probes. .
- Ferry AL, Vanderklish PW, Dupont-Versteegden EE. 2011. Enhanced survival of skeletal muscle myoblasts in response to overexpression of cold shock protein RBM3. *American Journal of Physiology-Cell Physiology* 301(2):C392-402.
- Fíla J and Honys D. 2012. Enrichment techniques employed in phosphoproteomics. *Amino Acids* 43(3):1025-47.
- Fisher D, Krasinska L, Coudreuse D, Novák B. 2012. Phosphorylation network dynamics in the control of cell cycle transitions. *J Cell Sci* 125(20):4703-11.
- Fitzgerald S, Sheehan KM, O'Grady A, Kenny D, O'Kennedy R, Kay EW, Kijanka GS. 2013. Relationship between epithelial and stromal TRIM28 expression predicts survival in colorectal cancer patients. *J Gastroenterol Hepatol* 28(6):967-74.
- Flanagan L, Whyte L, Chatterjee N, Tenniswood M. 2010. Effects of clusterin overexpression on metastatic progression and therapy in breast cancer. *BMC Cancer* 10(1):107.
- Follman DK and Fahrner RL. 2004. Factorial screening of antibody purification processes using three chromatography steps without protein A. *Journal of Chromatography A* 1024(1-2):79-85.
- Fowlkes JL and Winkler MK. 2002. Exploring the interface between metalloproteinase activity and growth factor and cytokine bioavailability. *Cytokine Growth Factor Rev* 13(3):277-87.
- Fowlkes JL, Enghild JJ, Suzuki K, Nagase H. 1994. Matrix metalloproteinases degrade insulin-like growth factor-binding protein-3 in dermal fibroblast cultures. *J Biol Chem* 269(41):25742.
- Fox SR, Patel UA, Yap MGS, Wang DIC. 2004. Maximizing interferon- $\gamma$  production by chinese hamster ovary cells through temperature shift optimization: Experimental and modeling. *Biotechnol Bioeng* 85(2):177-84.
- Fränzel B and Wolters DA. 2011. Advanced MudPIT as a next step toward high proteome coverage. *Proteomics* 11(18):3651-6.
- Frese CK, Altelaar AM, Hennrich ML, Nolting D, Zeller M, Griep-Raming J, Heck AJ, Mohammed S. 2011. Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-orbitrap velos. *Journal of Proteome Research* 10(5):2377-88.

- Froehlich JW, Chu CS, Tang N, Waddell K, Grimm R, Lebrilla CB. 2011. Label-free liquid chromatography-tandem mass spectrometry analysis with automated phosphopeptide enrichment reveals dynamic human milk protein phosphorylation during lactation. *Anal Biochem* 408(1):136-46.
- Fry AM, O'Regan L, Sabir SR, Bayliss R. 2012. Cell cycle regulation by the NEK family of protein kinases. *J Cell Sci* 125(19):4423-33.
- Fuchs SY, Tappin I, Ronai Z. 2000. Stability of the ATF2 transcription factor is regulated by phosphorylation and dephosphorylation. *J Biol Chem* 275(17):12560-4.
- Fujimoto N, Terlizzi J, Aho S, Brittingham R, Fertala A, Oyama N, McGrath JA, Uitto J. 2006. Extracellular matrix protein 1 inhibits the activity of matrix metalloproteinase 9 through high-affinity protein/protein interactions. *Exp Dermatol* 15(4):300-7.
- Fujioka S, Sundal C, Strongosky AJ, Castanedes MC, Rademakers R, Ross OA, Vilariño-Güell C, Farrer MJ, Wszolek ZK, Dickson DW. 2013. Sequence variants in eukaryotic translation initiation factor 4-gamma (eIF4G1) are associated with lewy body dementia. *Acta Neuropathol* 125(3):425-38.
- Fujita J. 1999. Cold shock response in mammalian cells. *J Mol Microbiol Biotechnol* 1(2):243-55.
- Gammell P, Barron N, Kumar N, Clynes M. 2007. Initial identification of low temperature and culture stage induction of miRNA expression in suspension CHO-K1 cells. *Journal of Biotechnology*, 130(3):213-8.
- Gandin V, Senft D, Topisirovic I, Ze'ev AR. 2013. RACK1 function in cell motility and protein synthesis. *Genes & Cancer* 4(9-10):369-77.
- Gao SX, Zhang Y, Stansberry-Perkins K, Buko A, Bai S, Nguyen V, Brader ML. 2011a. Fragmentation of a highly purified monoclonal antibody attributed to residual CHO cell protease activity. *Biotechnol Bioeng* 108(4):977-82.
- Gao XS, Li W, Song H, Liu SN. 2011b. Identification of nucleolar and coiled-body phosphoprotein 1 (NOLC1) minimal promoter regulated by NF- $\kappa$ B and CREB. *Biochemistry and Molecular Biology Reports* 44(1):70-5.
- Garrigue-Antar L, Hartigan N, Kadler KE. 2002. Post-translational modification of bone morphogenetic protein-1 is required for secretion and stability of the protein. *J Biol Chem* 277(45):43327-34.
- Gassmann R, Carvalho A, Henzing AJ, Ruchaud S, Hudson DF, Honda R, Nigg EA, Gerloff DL, Earnshaw WC. 2004. Borealin a novel chromosomal passenger required for stability of the bipolar mitotic spindle. *J Cell Biol* 166(2):179-91.
- Gates MB, Tomer KB, Deterding LJ. 2010. Comparison of metal and metal oxide media for phosphopeptide enrichment prior to mass spectrometric analyses. *J Am Soc Mass Spectrom* 21(10):1649-59.
- Geiger T, Wehner A, Schaab C, Cox J, Mann M. 2012. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol Cell Proteomics* 11(3):M111.014050.

Getie-Kebtie M, Lazarev A, Eichelberger M, Alterman M. 2011. Label-free mass spectrometry-based relative quantification of proteins separated by one-dimensional gel electrophoresis. *Anal Biochem* 409(2):202-12.

Ghalayini MK, Dong Q, Richardson DR, Assinder SJ. 2013. Proteolytic cleavage and truncation of NDRG1 in human prostate cancer cells, but not normal prostate epithelial cells. *Biosci Rep* .

Gnad F, Gunawardena J, Mann M. 2011. PHOSIDA 2011: The posttranslational modification database. *Nucleic Acids Res* 39(Database issue):D253-60.

Gnad F, Forner F, Zielinska DF, Birney E, Gunawardena J, Mann M. 2010. Evolutionary constraints of phosphorylation in eukaryotes, prokaryotes, and mitochondria. *Mol Cell Proteomics* 9(12):2642-53.

Golan A, Yudkovsky Y, Hershko A. 2002. The cyclin-ubiquitin ligase activity of cyclosome/APC is jointly activated by protein kinases Cdk1-cyclin B and plk. *J Biol Chem* 277(18):15552-7.

Gold LI, Eggleton P, Sweetwyne MT, Van Duyn LB, Greives MR, Naylor SM, Michalak M, Murphy-Ullrich JE. 2010. Calreticulin: Non-endoplasmic reticulum functions in physiology and disease. *FASEB J* 24(3):665-83.

Gondi CS and Rao JS. 2013. Cathepsin B as a cancer target. *Expert Opinion on Therapeutic Targets* 17(3):281-91.

Gong X, Li D, Li X, Fang Q, Han X, Wu Y, Yang S, Shen BQ. 2006. Fed-batch culture optimization of a growth-associated hybridoma cell line in chemically defined protein-free media. *Cytotechnology* 52(1):25-38.

Gordus A and MacBeath G. 2006. Circumventing the problems caused by protein diversity in microarrays: Implications for protein interaction networks. *J Am Chem Soc* 128(42):13668-9.

Gramer MJ and Goochee CF. 1993. Glycosidase activities in chinese hamster ovary cell lysate and cell culture supernatant. *Biotechnol Prog* 9(4):366-73.

Griffin TJ, Seth G, Xie H, Bandhakavi S, Hu W. 2007. Advancing mammalian cell culture engineering using genome-scale technologies. *Trends Biotechnol* 25(9):401-8.

Grillberger L, Kreil TR, Nasr S, Reiter M. 2009. Emerging trends in plasma-free manufacturing of recombinant protein therapeutics expressed in mammalian cells. *Biotechnology Journal* 4(2):186-201.

Gronborg M, Kristiansen TZ, Stensballe A, Andersen JS, Ohara O, Mann M, Jensen ON, Pandey A. 2002. A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: Identification of a novel protein, frigg, as a protein kinase A substrate. *Mol Cell Proteomics* 1(7):517-27.

Grossmann J, Roschitzki B, Panse C, Fortes C, Barkow-Oesterreicher S, Rutishauser D, Schlapbach R. 2010. Implementation and evaluation of relative and absolute

quantification in shotgun proteomics with label-free methods. *Journal of Proteomics* 73(9):1740-6.

Grzeskowiak JK, Tscheliessnig A, Toh PC, Chusainow J, Lee YY, Wong N, Jungbauer A. 2009. 2-D DIGE to expedite downstream process development for human monoclonal antibody purification. *Protein Expr Purif* 66(1):58-65.

Gunaratne R, Braucht DWW, Rinschen MM, Chou C, Hoffert JD, Pisitkun T, Knepper MA. 2010. Quantitative phosphoproteomic analysis reveals cAMP/vasopressin-dependent signaling pathways in native renal thick ascending limb cells. *Proc Natl Acad Sci U S A* 107(35):15653-8.

Gygi SP, Rochon Y, Franza BR, Aebersold R. 1999. Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol* 19(3):1720-30.

Gygi S, Corthals G, Zhang Y, Rochon Y, Aebersold R. 2000. Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A* 97(17):9390-5.

Gygi S, Rist B, Gerber S, Turecek F, Gelb M, Aebersold R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* 17(10):994-9.

Hacker DL, De Jesus M, Wurm FM. 2009. 25 years of recombinant proteins from reactor-grown cells—Where do we go from here? *Biotechnol Adv* 27(6):1023-7.

Hanash S and Taguchi A. 2010. The grand challenge to decipher the cancer proteome. *Nature Reviews Cancer* 10(9):652-60.

Hansen K, Kjalke M, Rasmussen PB, Kongerslev L, Ezban M. 1997. Proteolytic cleavage of recombinant two-chain factor VIII during cell culture production is mediated by protease (s) from lysed cells. the use of pulse labelling directly in production medium. *Cytotechnology* 24(3):227-34.

Hardt M, Thomas LR, Dixon SE, Newport G, Agabian N, Prakobphol A, Hall SC, Witkowska HE, Fisher SJ. 2005. Toward defining the human parotid gland salivary proteome and peptidome: Identification and characterization using 2D SDS-PAGE, ultrafiltration, HPLC, and mass spectrometry. *Biochemistry (N Y)* 44(8):2885-99.

Harsha H and Pandey A. 2010. Phosphoproteomics in cancer. *Molecular Oncology* 4(6):482-95.

Hattan SJ and Parker KC. 2006. Methodology utilizing MS signal intensity and LC retention time for quantitative analysis and precursor ion selection in proteomic LC-MALDI analyses. *Anal Chem* 78(23):7986-96.

Hayduk EJ, Choe LH, Lee KH. 2004. A two-dimensional electrophoresis map of chinese hamster ovary cell proteins based on fluorescence staining. *Electrophoresis* 25(15):2545-56.

Hendrick V, Winnepenninckx P, Abdelkafi C, Vandeputte O, Cherlet M, Marique T, Renemann G, Loa A, Kretzmer G, Werenne J. 2001. Increased productivity of recombinant tissular plasminogen activator (t-PA) by butyrate and shift of temperature: A cell cycle phases analysis. *Cytotechnology* 36(1-3):71-83.

- Herper M. 2010. The world's most expensive drugs. *Forbes.Com*. Retrieved July 4:2012.
- Hilton LK, White MC, Quarmby LM. 2009. The NIMA-related kinase NEK1 cycles through the nucleus. *Biochem Biophys Res Commun* 389(1):52-6.
- Hintze BB and Aufenvenne K. 2007. The protease domain of procollagen C-proteinase (BMP1) lacks substrate selectivity, which is conferred by non-proteolytic domains. *Biol Chem* 388:513-21.
- Hoefsloot LH, Hoogeveen-Westerveld M, Kroos MA, van Beeumen J, Reuser AJ, Oostra BA. 1988. Primary structure and processing of lysosomal alpha-glucosidase; homology with the intestinal sucrase-isomaltase complex. *EMBO J* 7(6):1697-704.
- Hoffman K. 2000. Strategies for host cell protein analysis. *BIOPHARM-EUGENE*-13(5):38-47.
- Hogwood CE, Tait AS, Koloteva-Levine N, Bracewell DG, Smales CM. 2013. The dynamics of the CHO host cell protein profile during clarification and protein A capture in a platform antibody purification process. *Biotechnol Bioeng* 110(1):240-51.
- Holt C and Yandell M. 2011. MAKER2: An annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491,2105-12-491.
- Hong JK, Kim Y, Yoon SK, Lee GM. 2007. Down-regulation of cold-inducible RNA-binding protein does not improve hypothermic growth of chinese hamster ovary cells producing erythropoietin. *Metab Eng* 9(2):208-16.
- Hopkins DR, Keles S, Greenspan DS. 2007. The bone morphogenetic protein 1/Tolloid-like metalloproteinases. *Matrix Biology* 26(7):508-23.
- Hornbeck PV, Kornhauser JM, Tkachev S, Zhang B, Skrzypek E, Murray B, Latham V, Sullivan M. 2012. PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* 40(Database issue):D261-70.
- Hossler P, Khattak SF, Li ZJ. 2009. Optimal and consistent protein glycosylation in mammalian cell culture. *Glycobiology* 19(9):936-49.
- HS Lu C, Liu K, Tan LP, Yao SQ. 2012. Current chemical biology tools for studying protein phosphorylation and dephosphorylation. *Chemistry-A European Journal* 18(1):28-39.
- Huang Y, Hu W, Rustandi E, Chang K, Yusuf-Makagiansar H, Ryll T. 2010. Maximizing productivity of CHO cell-based fed-batch culture using chemically defined media conditions and typical manufacturing equipment. *Biotechnol Prog* 26(5):1400-10.
- Huang M, Miao ZH, Zhu H, Cai YJ, Lu W, Ding J. 2008. Chk1 and Chk2 are differentially involved in homologous recombination repair and cell cycle arrest in response to DNA double-strand breaks induced by camptothecins. *Mol Cancer Ther* 7(6):1440-9.

- Huggett B and Lähteenmaki R. 2012. Public biotech 2011—the numbers. *Nat Biotechnol* 30(8):751.
- Hwang SO, Chung JY, Lee GM. 2003. Effect of Doxycycline-Regulated ERp57 expression on specific thrombopoietin productivity of recombinant CHO cells. *Biotechnol Prog* 19(1):179-84.
- Hynes NE, Ingham PW, Lim WA, Marshall CJ, Massagué J, Pawson T. 2013. Signalling change: Signal transduction through the decades. *Nature Reviews Molecular Cell Biology* 14(6):393-8.
- Imai Y, Busby Jr WH, Smith CE, Clarke JB, Garmong AJ, Horwitz GD, Rees C, Clemmons DR. 1997. Protease-resistant form of insulin-like growth factor-binding protein 5 is an inhibitor of insulin-like growth factor-I actions on porcine smooth muscle cells in culture. *J Clin Invest* 100(10):2596.
- Issaq HJ and Veenstra TD. 2008. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): Advances and perspectives. *BioTechniques* 44(5):697.
- Ivanov IP, Firth AE, Michel AM, Atkins JF, Baranov PV. 2011. Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res* 39(10):4220-34.
- Iyengar S and Farnham PJ. 2011. KAP1 protein: An enigmatic master regulator of the genome. *J Biol Chem* 286(30):26267-76.
- Jayapal KP, Wlaschin KF, Hu W, Yap MG. 2007. Recombinant protein therapeutics from CHO cells-20 years and counting. *Chem Eng Prog* 103(10):40.
- Jefferis R. 2005. Glycosylation of recombinant antibody therapeutics. *Biotechnol Prog* 21(1):11-6.
- Jiang CC, Lai F, Thorne RF, Yang F, Liu H, Hersey P, Zhang XD. 2011. MEK-independent survival of B-RAFV600E melanoma cells selected for resistance to apoptosis induced by the RAF inhibitor PLX4720. *Clinical Cancer Research* 17(4):721-30.
- Jin M, Szapiel N, Zhang J, Hickey J, Ghose S. 2010. Profiling of host cell proteins by two-dimensional difference gel electrophoresis (2D-DIGE): Implications for downstream process development. *Biotechnol Bioeng* 105(2):306-16.
- Jones RB, Gordus A, Krall JA, MacBeath G. 2005. A quantitative protein interaction network for the ErbB receptors using protein microarrays. *Nature* 439(7073):168-74.
- Jørgensen P, Gräslund S, Betz R, Ståhl S, Larsson C, Höög C. 2001. Characterisation of the human APC1, the largest subunit of the anaphase-promoting complex. *Gene* 262(1):51-9.
- Jünger MA and Aebersold R. 2014. Mass spectrometry-driven phosphoproteomics: Patterning the systems biology mosaic. *Wiley Interdisciplinary Reviews: Developmental Biology* 3(1):83-112.
- Käll L and Vitek O. 2011. Computational mass Spectrometry–Based proteomics. *PLoS Computational Biology* 7(12):e1002277.



Kang TH, Bae K, Yu M, Kim W, Hwang H, Jung H, Lee PY, Kang S, Yoon T, Park SG. 2007. Phosphoproteomic analysis of neuronal cell death by glutamate-induced oxidative stress. *Proteomics* 7(15):2624-35.

Kantardjieff A, Jacob NM, Yee JC, Epstein E, Kok Y, Philp R, Betenbaugh M, Hu W. 2010. Transcriptome and proteome analysis of chinese hamster ovary cells under low temperature and butyrate treatment. *J Biotechnol* 145(2):143-59.

Kao YH, Hewitt DP, Trexler-Schmidt M, Laird MW. 2010. Mechanism of antibody reduction in cell culture production processes. *Biotechnol Bioeng* .

Karl DW, Donovan M, Flickinger MC. 1990. A novel acid proteinase released by hybridoma cells. *Cytotechnology* 3(2):157-69.

Kaufmann H, Mazur X, Fussenegger M, Bailey JE. 1999. Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells. *Biotechnol Bioeng* 63(5):573-82.

Keenan J, Pearson D, Clynes M. 2006. The role of recombinant proteins in the development of serum-free media. *Cytotechnology* 50(1):49-56.

Keicher C, Gambaryan S, Schulze E, Marcus K, Meyer HE, Butt E. 2004. Phosphorylation of mouse LASP-1 on threonine 156 by cAMP- and cGMP-dependent protein kinase. *Biochem Biophys Res Commun* 324(1):308-16.

Khatib A and Geraldine S. 2006. Growth factors: To cleave or not to cleave. In: *Regulation of carcinogenesis, angiogenesis and metastasis by the proprotein convertases (PCs)*. Springer. 121 p.

Kim HJ and Bar-Sagi D. 2004. Modulation of signalling by sprouty: A developing story. *Nature Reviews Molecular Cell Biology* 5(6):441-50.

Kim NS and Lee GM. 2000. Overexpression of bcl-2 inhibits sodium butyrate-induced apoptosis in chinese hamster ovary cells resulting in enhanced humanized antibody production. *Biotechnol Bioeng* 71(3):184-93.

Kim S and Coulombe PA. 2010. Emerging role for the cytoskeleton as an organizer and regulator of translation. *Nature Reviews Molecular Cell Biology* 11(1):75-81.

Kim W, Tokunaga M, Ozaki H, Ishibashi T, Honda K, Kajiura H, Fujiyama K, Asano R, Kumagai I, Omasa T. 2010. Glycosylation pattern of humanized IgG-like bispecific antibody produced by recombinant CHO cells. *Appl Microbiol Biotechnol* 85(3):535-42.

Kim Y, Park B, Lee S, Ahn JO, Jung J, Lee HW, Lee GM, Lee EG. 2012. Effect of mitochondrial and ER-targeted bcl-2 overexpression on apoptosis in recombinant chinese hamster ovary cells treated with sodium butyrate. *Process Biochemistry* 47(12):2518-22.

Kim Y, Han YK, Kim JY, Lee EG, Lee HW, Lee GM. 2011. Effect of constitutively active ras overexpression on cell growth in recombinant chinese hamster ovary cells. *Biotechnol Prog* 27(2):577-80.

- Kim Y, Jin Y, Vukoti KM, Park JK, Kim EE, Lee K, Yu YG. 2003. Purification and characterization of human nucleolar phosphoprotein 140 expressed in *Escherichia coli*. *Protein Expr Purif* 31(2):260-4.
- Kinder M, Greenplate AR, Grugan KD, Soring KL, Heeringa KA, McCarthy SG, Bannish G, Perpetua M, Lynch F, Jordan RE, et al. 2013. Engineered protease-resistant antibodies with selectable cell-killing functions. *J Biol Chem* 288(43):30843-54.
- King CA. 2013. Kaposi's sarcoma-associated herpesvirus kaposin B induces unique monophosphorylation of STAT3 at serine 727 and MK2-mediated inactivation of the STAT3 transcriptional repressor TRIM28. *J Virol* 87(15):8779-91.
- Kitamura Tomoya, Maeda Kengo, Enomoto Atushi, Kondo Takahisa, Takahasi Masahide and Murohara Toyooki. 2006. A novel akt substrate girdin plays essential roles in signaling pathways after VEGF-stimulation in endothelial cells. *Circulation* Lippincott Williams & Wilkins 530 Walnut St, Philadelphia, PA 19106-3621 USA. 33 p.
- Klass B, Grobbelaar A, Rolfe K. 2009. Transforming growth factor  $\beta$ 1 signalling, wound healing and repair: A multifunctional cytokine with clinical implications for wound repair, a delicate balance. *Postgrad Med J* 85(999):9-14.
- Kloetzel P and Ossendorp F. 2004. Proteasome and peptidase function in MHC-class-I-mediated antigen presentation. *Curr Opin Immunol* 16(1):76-81.
- Kokkonen N, Rivinoja A, Kauppila A, Suokas M, Kellokumpu I, Kellokumpu S. 2004. Defective acidification of intracellular organelles results in aberrant secretion of cathepsin D in cancer cells. *J Biol Chem* 279(38):39982-8.
- Korke R, Gatti MdL, Lau ALY, Lim JWE, Seow TK, Chung MCM, Hu W. 2004. Large scale gene expression profiling of metabolic shift of mammalian cells in culture. *J Biotechnol* 107(1):1-17.
- Koterba KL, Borgschulte T, Laird MW. 2012. Thioredoxin 1 is responsible for antibody disulfide reduction in CHO cell culture. *J Biotechnol* 157(1):261-7.
- Kou T, Fan L, Zhou Y, Ye Z, Liu X, Zhao L, Tan W. 2011. Detailed understanding of enhanced specific productivity in chinese hamster ovary cells at low culture temperature. *Journal of Bioscience and Bioengineering* 111(3):365-9.
- Kowanetz M and Ferrara N. 2006. Vascular endothelial growth factor signaling pathways: Therapeutic perspective. *Clinical Cancer Research* 12(17):5018-22.
- Krampe B and Al-Rubeai M. 2010. Cell death in mammalian cell culture: Molecular mechanisms and cell line engineering strategies. *Cytotechnology* 62(3):175-88.
- Krawitz DC, Forrest W, Moreno GT, Kittleson J, Champion KM. 2006. Proteomic studies support the use of multi-product immunoassays to monitor host cell protein impurities. *Proteomics* 6(1):94-110.
- Kridel SJ, Chen E, Kotra LP, Howard EW, Mobashery S, Smith JW. 2001. Substrate hydrolysis by matrix metalloproteinase-9. *J Biol Chem* 276(23):20572-8.

- Krishnamoorthy T, Pavitt GD, Zhang F, Dever TE, Hinnebusch AG. 2001. Tight binding of the phosphorylated alpha subunit of initiation factor 2 (eIF2alpha) to the regulatory subunits of guanine nucleotide exchange factor eIF2B is required for inhibition of translation initiation. *Mol Cell Biol* 21(15):5018-30.
- Kubiczkova L, Sedlarikova L, Hajek R, Sevcikova S. 2012. TGF- $\beta$ —an excellent servant but a bad master. *Journal of Translational Medicine* 10(1):1-24.
- Kumar C and Mann M. 2009. Bioinformatics analysis of mass spectrometry-based proteomics data sets. *FEBS Lett* 583(11):1703-12.
- Kumar N, Gammell P, Clynes M. 2007. Proliferation control strategies to improve productivity and survival during CHO based production culture - A summary of recent methods employed and the effects of proliferation control in product secreting CHO cell lines. *Cytotechnology* 53(1-3):33-46.
- Kumar N, Gammell P, Meleady P, Henry M, Clynes M. 2008a. Differential protein expression following low temperature culture of suspension CHO-K1 cells. *Bmc Biotechnology* 8:42.
- Kumar N, Maurya P, Ganunell P, Dowling P, Clynes M, Meleady P. 2008b. Proteomic profiling of secreted proteins from CHO cells using surface-enhanced laser desorption ionization time-of-flight mass spectrometry. *Biotechnol Prog* 24(1):273-8.
- Lahti JM, Xiang J, Heath LS, Campana D, Kidd VJ. 1995. PITSLRE protein kinase activity is associated with apoptosis. *Mol Cell Biol* 15(1):1-11.
- Larsen MR, Thingholm TE, Jensen ON, Roepstorff P, Jorgensen TJ. 2005. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* 4(7):873-86.
- Lawlor K, Nazarian A, Lacomis L, Tempst P, Villanueva J. 2009. Pathway-based biomarker search by high-throughput proteomics profiling of secretomes. *Journal of Proteome Research* 8(3):1489-503.
- Lee Jr J and McCubrey J. 2002. The Raf/MEK/ERK signal transduction cascade as a target for chemotherapeutic intervention in leukemia. *Leukemia: Official Journal of the Leukemia Society of America, Leukemia Research Fund, UK* 16(4):486-507.
- Lee H and Tomarev SI. 2007. Optimedin induces expression of N-cadherin and stimulates aggregation of NGF-stimulated PC12 cells. *Exp Cell Res* 313(1):98-108.
- Lee JS, Park HJ, Kim YH, Lee GM. 2010. Protein reference mapping of dihydrofolate reductase-deficient CHO DG44 cell lines using 2-dimensional electrophoresis. *Proteomics* 10(12):2292-302.
- Lee YH, Tan HT, Chung M. 2010. Subcellular fractionation methods and strategies for proteomics. *Proteomics* 10(22):3935-56.
- Leskov KS, Klovov DY, Li J, Kinsella TJ, Boothman DA. 2003. Synthesis and functional analyses of nuclear clusterin, a cell death protein. *J Biol Chem* 278(13):11590-600.

- Lesslie D, Summy J, Parikh N, Fan F, Trevino J, Sawyer T, Metcalf C, Shakespeare W, Hicklin D, Ellis L. 2006. Vascular endothelial growth factor receptor-1 mediates migration of human colorectal carcinoma cells by activation of src family kinases. *Br J Cancer* 94(11):1710-7.
- Levin Y, Schwarz E, Wang L, Leweke FM, Bahn S. 2007. Label-free LC-MS/MS quantitative proteomics for large-scale biomarker discovery in complex samples. *Journal of Separation Science* 30(14):2198-203.
- Lewis NE, Liu X, Li Y, Nagarajan H, Yerganian G, O'Brien E, Bordbar A, Roth AM, Rosenbloom J, Bian C. 2013. Genomic landscapes of chinese hamster ovary cell lines as revealed by the cricetus griseus draft genome. *Nat Biotechnol* 31(8):759-65.
- Li D, Han Z, Liu J, Zhang X, Ren J, Yan L, Liu H, Xu Z. 2013. Upregulation of nucleus HDGF predicts poor prognostic outcome in patients with penile squamous cell carcinoma bypass VEGF-A and ki-67. *Medical Oncology* 30(4):1-8.
- Li F, Seillier-Moisewitsch F, Korostyshevskiy VR. 2011. Region-based statistical analysis of 2D PAGE images. *Comput Stat Data Anal* 55(11):3059-72.
- Li Q and Chen H. 2011. Transcriptional silencing of N-myc downstream-regulated gene 1 (NDRG1) in metastatic colon cancer cell line SW620. *Clin Exp Metastasis* 28(2):127-35.
- Liang H, Ding X, Zhou C, Zhang Y, Xu M, Zhang C, Xu L. 2012. Knockdown of eukaryotic translation initiation factors 3B (EIF3B) inhibits proliferation and promotes apoptosis in glioblastoma cells. *Neurological Sciences* 33(5):1057-62.
- Licari PJ, Jarvis DL, Bailey JE. 1993. Insect cell hosts for baculovirus expression vectors contain endogenous exoglycosidase activity. *Biotechnol Prog* 9(2):146-52.
- Liew JCJ, Tan WS, Alitheen NBM, Chan ES, Tey BT. 2010. Over-expression of the X-linked inhibitor of apoptosis protein (XIAP) delays serum deprivation-induced apoptosis in CHO-K1 cells. *Journal of Bioscience and Bioengineering* 110(3):338-44.
- Lim UM, Yap MG, Lim Y, Goh LT, Ng SK. 2013. Identification of autocrine growth factors secreted by CHO cells for applications in single cell cloning media. *Journal of Proteome Research* .
- Lim Y, Wong NS, Lee YY, Ku SC, Wong DC, Yap MG. 2010. Engineering mammalian cells in bioprocessing—current achievements and future perspectives. *Biotechnol Appl Biochem* 55(4):175-89.
- Lin C, Whang EE, Moalem J, Ruan DT. 2012. Strategic combination therapy overcomes tyrosine kinase coactivation in adrenocortical carcinoma. *Surgery* 152(6):1045-50.
- Lin J, Xie Z, Zhu H, Qian J. 2010. Understanding protein phosphorylation on a systems level. *Brief Funct Genomics* 9(1):32-42.

Lindskog E, Svensson I, Häggström L. 2006. A homologue of cathepsin L identified in conditioned medium from Sf9 insect cells. *Appl Microbiol Biotechnol* 71(4):444-9.

Linger R, Cohen R, Cummings C, Sather S, Migdall-Wilson J, Middleton D, Lu X, Baron A, Franklin W, Merrick D. 2013. Mer or axl receptor tyrosine kinase inhibition promotes apoptosis, blocks growth and enhances chemosensitivity of human non-small cell lung cancer. *Oncogene* 32(29):3420-31.

Lipscomb ML, Palomares LA, Hernández V, Ramírez OT, Kompala DS. 2005. Effect of production method and gene amplification on the glycosylation pattern of a secreted reporter protein in CHO cells. *Biotechnol Prog* 21(1):40-9.

Liska AJ and Shevchenko A. 2003. Expanding the organismal scope of proteomics: Cross-species protein identification by mass spectrometry and its implications. *Proteomics* 3(1):19-28.

Liu A, Zhang Z, Li A, Xue J. 2010. Effects of hypothermia and cerebral ischemia on cold-inducible RNA-binding protein mRNA expression in rat brain. *Brain Res* 1347:104-10.

Liu C, Zhang Y, Xu H, Zhang R, Li H, Lu P, Jin F. 2012. Girdin protein: A new potential distant metastasis predictor of breast cancer. *Medical Oncology* 29(3):1554-60.

Liu H, Sadygov RG, Yates III JR. 2004. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal Chem* 76(14):4193-201.

Liu Hongcheng and May Kimberly. 2012. Disulfide bond structures of IgG molecules: Structural variations, chemical modifications and possible impacts to stability and biological function. *mAbs* Landes Bioscience. 17 p.

Liu X, Cheng C, Shao B, Wu X, Ji Y, Lu X, Shen A. 2013. LPS-stimulating astrocyte-conditioned medium causes neuronal apoptosis via increasing CDK11p58 expression in PC12 cells through downregulating AKT pathway. *Cell Mol Neurobiol* :1-9.

Lo A, Tang Y, Chen L, Li L. 2013. Automation of dimethylation after guanidination labeling chemistry and its compatibility with common buffers and surfactants for mass spectrometry-based shotgun quantitative proteome analysis. *Anal Chim Acta* 788:81-8.

López-Otín C and Overall CM. 2002. Protease degradomics: A new challenge for proteomics. *Nature Reviews Molecular Cell Biology* 3(7):509-19.

Los DA and Murata N. 2004. Membrane fluidity and its roles in the perception of environmental signals. *Biochimica Et Biophysica Acta (BBA)-Biomembranes* 1666(1):142-57.

Lu Y, Boer JM, Barsova RM, Favorova O, Goel A, Müller M, Feskens EJ. 2012. TGFB1 genetic polymorphisms and coronary heart disease risk: A meta-analysis. *BMC Medical Genetics* 13(1):39.

- Luo J, Zhang J, Ren D, Tsai W, Li F, Amanullah A, Hudson T. 2012. Probing of C-terminal lysine variation in a recombinant monoclonal antibody production using chinese hamster ovary cells with chemically defined media. *Biotechnol Bioeng* 109(9):2306-15.
- Mahuran DJ. 1999. Biochemical consequences of mutations causing the GM2 gangliosidosis. *Biochimica Et Biophysica Acta (BBA)-Molecular Basis of Disease* 1455(2):105-38.
- Majors BS, Betenbaugh MJ, Pederson NE, Chiang GG. 2009. Mcl-1 overexpression leads to higher viabilities and increased production of humanized monoclonal antibody in chinese hamster ovary cells. *Biotechnol Prog* 25(4):1161-8.
- Makridakis M, Roubelakis MG, Bitsika V, Dimuccio V, Samiotaki M, Kossida S, Panayotou G, Coleman J, Candiano G, Anagnou NP. 2010. Analysis of secreted proteins for the study of bladder cancer cell aggressiveness. *Journal of Proteome Research* 9(6):3243-59.
- Mamidipudi V and Cartwright C. 2009. A novel pro-apoptotic function of RACK1: Suppression of src activity in the intrinsic and akt pathways. *Oncogene* 28(50):4421-33.
- Manes NP, Dong L, Zhou W, Du X, Reghu N, Kool AC, Choi D, Bailey CL, Petricoin EF, Liotta LA. 2011. Discovery of mouse spleen signaling responses to anthrax using label-free quantitative phosphoproteomics via mass spectrometry. *Molecular & Cellular Proteomics* 10(3).
- Mann M and Kelleher NL. 2008. Precision proteomics: The case for high resolution and high mass accuracy. *Proceedings of the National Academy of Sciences* 105(47):18132-8.
- Manning BD and Cantley LC. 2007. AKT/PKB signaling: Navigating downstream. *Cell* 129(7):1261.
- Mao J, Xu Z, Fang Y, Wang H, Xu J, Ye J, Zheng S, Zhu Y. 2008. Hepatoma-derived growth factor involved in the carcinogenesis of gastric epithelial cells through promotion of cell proliferation by Erk1/2 activation. *Cancer Science* 99(11):2120-7.
- Mao J, Jiang P, Cui S, Ren Y, Zhao J, Yin X, Enomoto A, Liu H, Hou L, Takahashi M. 2012. Girdin locates in centrosome and midbody and plays an important role in cell division. *Cancer Science* 103(10):1780-7.
- Marcet-Palacios M, Duggan BL, Shostak I, Barry M, Geskes T, Wilkins JA, Yanagiya A, Sonenberg N, Bleackley RC. 2011. Granzyme B inhibits vaccinia virus production through proteolytic cleavage of eukaryotic initiation factor 4 gamma 3. *PLoS Pathogens* 7(12):e1002447.
- Marchant RJ, Al-Fageeh MB, Underhill MF, Racher AJ, Smales CM. 2008. Metabolic rates, growth phase, and mRNA levels influence cell-specific antibody production levels from in vitro-cultured mammalian cells at sub-physiological temperatures. *Mol Biotechnol* 39(1):69-77.

- Masson O, Bach A, Derocq D, Prébois C, Laurent-Matha V, Pattingre S, Liaudet-Coopman E. 2010. Pathophysiological functions of cathepsin D: Targeting its catalytic activity versus its protein binding activity? *Biochimie* 92(11):1635-43.
- Masterton RJ, Roobol A, Al-Fageeh MB, Carden MJ, Smales CM. 2010. Post-translational events of a model reporter protein proceed with higher fidelity and accuracy upon mild hypothermic culturing of chinese hamster ovary cells. *Biotechnol Bioeng* 105(1):215-20.
- Masterton RJ and Smales CM. 2014. The impact of process temperature on mammalian cell lines and the implications for the production of recombinant proteins in CHO cells. *Pharmaceutical Bioprocessing* 2(1):49-61.
- Masutani M, Sonenberg N, Yokoyama S, Imataka H. 2007. Reconstitution reveals the functional core of mammalian eIF3. *EMBO J* 26(14):3373-83.
- Matijasevic Z, Snyder J, Ludlum D. 1998. Hypothermia causes a reversible, p53-mediated cell cycle arrest in cultured fibroblasts. *Oncol Res* 10(11-12):605-10.
- Matsushashi S, Narisawa Y, Ozaki I, Mizuta T. 2007. Expression patterns of programmed cell death 4 protein in normal human skin and some representative skin lesions. *Exp Dermatol* 16(3):179-84.
- Mayya V and Han DK. 2009. Phosphoproteomics by mass spectrometry: Insights, implications, applications and limitations. .
- McCaig C, Potter L, Abramczyk O, Murray JT. 2011. Phosphorylation of NDRG1 is temporally and spatially controlled during the cell cycle. *Biochem Biophys Res Commun* 411(2):227-34.
- Melanson JE, Avery SL, Pinto DM. 2006. High-coverage quantitative proteomics using amine-specific isotopic labeling. *Proteomics* 6(16):4466-74.
- Meleady P, Doolan P, Henry M, Barron N, Keenan J, O'Sullivan F, Clarke C, Gammell P, Melville M, Leonard M. 2011. Sustained productivity in recombinant chinese hamster ovary (CHO) cell lines: Proteome analysis of the molecular basis for a process-related phenotype. *BMC Biotechnology* 11(1):78.
- Meleady P, Hoffrogge R, Henry M, Rupp O, Bort JH, Clarke C, Brinkrolf K, Kelly S, Mueller B, Doolan P, et al. 2012. Utilization and evaluation of CHO-specific sequence databases for mass spectrometry based proteomics. *Biotechnol Bioeng* 109(6):1386-94.
- Mendoza MC, Er EE, Blenis J. 2011. The ras-ERK and PI3K-mTOR pathways: Cross-talk and compensation. *Trends Biochem Sci* 36(6):320-8.
- Merregaert J, Van Langen J, Hansen U, Ponsaerts P, El Ghalbzouri A, Steenackers E, Van Ostade X, Sercu S. 2010. Phospholipid scramblase 1 is secreted by a lipid raft-dependent pathway and interacts with the extracellular matrix protein 1 in the dermal epidermal junction zone of human skin. *J Biol Chem* 285(48):37823-37.
- Miljkovic-Licina M, Hammel P, Garrido-Urbani S, Lee BP, Meguenani M, Chaabane C, Bochaton-Piallat ML, Imhof BA. 2012. Targeting olfactomedin-like 3

inhibits tumor growth by impairing angiogenesis and pericyte coverage. *Mol Cancer Ther* 11(12):2588-99.

Mirgorodskaya E, Braeuer C, Fucini P, Lehrach H, Gobom J. 2005. Nanoflow liquid chromatography coupled to matrix-assisted laser desorption/ionization mass spectrometry: Sample preparation, data analysis, and application to the analysis of complex peptide mixtures. *Proteomics* 5(2):399-408.

Mohamad SB, Nagasawa H, Uto Y, Hori H. 2002. Tumor cell alpha-N-acetylgalactosaminidase activity and its involvement in GcMAF-related macrophage activation. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology* 132(1):1-8.

Mohan C, Park SH, Chung JY, Lee GM. 2007. Effect of doxycycline-regulated protein disulfide isomerase expression on the specific productivity of recombinant CHO cells: Thrombopoietin and antibody. *Biotechnol Bioeng* 98(3):611-5.

Mohan S, Nakao Y, Honda Y, Landale E, Leser U, Dony C, Lang K, Baylink DJ. 1995. Studies on the mechanisms by which insulin-like growth factor (IGF) binding protein-4 (IGFBP-4) and IGFBP-5 modulate IGF actions in bone cells. *J Biol Chem* 270(35):20424.

Molinari M, Eriksson KK, Calanca V, Galli C, Cresswell P, Michalak M, Helenius A. 2004. Contrasting functions of calreticulin and calnexin in glycoprotein folding and ER quality control. *Mol Cell* 13(1):125-35.

Mols J, Peeters-Joris C, Agathos SN, Schneider YJ. 2004. Origin of rice protein hydrolysates added to protein-free media alters secretion and extracellular proteolysis of recombinant interferon- $\gamma$  as well as CHO-320 cell growth. *Biotechnol Lett* 26(13):1043-6.

Montoya A, Beltran L, Casado P, Rodríguez-Prados J, Cutillas PR. 2011. Characterization of a TiO<sub>2</sub> enrichment method for label-free quantitative phosphoproteomics. *Methods* 54(4):370-8.

Moore A, Mercer J, Dutina G, Donahue CJ, Bauer KD, Mather JP, Etcheverry T, Ryll T. 1997. Effects of temperature shift on cell cycle, apoptosis and nucleotide pools in CHO cell batch cultures. *Cytotechnology* 23(1-3):47-54.

Moruz L, Pichler P, Stranzl T, Mechtler K, Käll L. 2013. Optimized nonlinear gradients for reversed-phase liquid chromatography in shotgun proteomics. *Anal Chem* 85(16):7777-85.

Mu D, Cambier S, Fjellbirkeland L, Baron JL, Munger JS, Kawakatsu H, Sheppard D, Broaddus VC, Nishimura SL. 2002. The integrin  $\alpha\beta 8$  mediates epithelial homeostasis through MT1-MMP-dependent activation of TGF- $\beta 1$ . *J Cell Biol* 157(3):493-507.

Muchamuel T, Basler M, Aujay MA, Suzuki E, Kalim KW, Lauer C, Sylvain C, Ring ER, Shields J, Jiang J. 2009. A selective inhibitor of the immunoproteasome subunit LMP7 blocks cytokine production and attenuates progression of experimental arthritis. *Nat Med* 15(7):781-7.



Munoz J, Low TY, Kok YJ, Chin A, Frese CK, Ding V, Choo A, Heck AJ. 2011. The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. *Molecular Systems Biology* 7(1).

Mysliwy J, Dingley AJ, Sedlacek R, Grötzinger J. 2006. Structural characterization and binding properties of the hemopexin-like domain of the matrixmetalloproteinase-19. *Protein Expr Purif* 46(2):406-13.

Neville DC, Townsend RR, Rozanas CR, Verkman A, Price EM, Gruis DB. 1997. Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. *Protein Science* 6(11):2436-45.

Newman RH, Hu J, Rho H, Xie Z, Woodard C, Neiswinger J, Cooper C, Shirley M, Clark HM, Hu S. 2013. Construction of human activity-based phosphorylation networks. *Molecular Systems Biology* 9(1).

Nguyen LK, Kolch W, Kholodenko BN. 2013. When ubiquitination meets phosphorylation: A systems biology perspective of EGFR/MAPK signalling. *Cell Communication and Signaling* 11(1):52.

Niklew M, Hochkirch U, Melikyan A, Moritz T, Kurzawski S, Schlüter H, Ebner I, Linscheid MW. 2010. Phosphopeptide screening using nanocrystalline titanium dioxide films as affinity matrix-assisted laser desorption ionization targets in mass spectrometry. *Anal Chem* 82(3):1047-53.

Nilsson CL, Dillon R, Devakumar A, Shi SD, Greig M, Rogers JC, Krastins B, Rosenblatt M, Kilmer G, Major M. 2009. Quantitative phosphoproteomic analysis of the STAT3/IL-6/HIF1 $\alpha$  signaling network: An initial study in GSC11 glioblastoma stem cells. *Journal of Proteome Research* 9(1):430-43.

Nishiyama H, Itoh K, Kaneko Y, Kishishita M, Yoshida O, Fujita J. 1997. A glycine-rich RNA-binding protein mediating cold-inducible suppression of mammalian cell growth. *J Cell Biol* 137(4):899.

Nissom PM, Sanny A, Kok YJ, Hiang YT, Chuah SH, Shing TK, Lee YY, Wong TK, Hu W, Sim MYG. 2006. Transcriptome and proteome profiling to understanding the biology of high productivity CHO cells. *Mol Biotechnol* 34(2):125-40.

Nita-Lazar A, Saito-Benz H, White FM. 2008. Quantitative phosphoproteomics by mass spectrometry: Past, present, and future. *Proteomics* 8(21):4433-43.

Nix DA and Beckerle MC. 1997. Nuclear–cytoplasmic shuttling of the focal contact protein, zyxin: A potential mechanism for communication between sites of cell adhesion and the nucleus. *J Cell Biol* 138(5):1139-47.

Noh SM, Sathyamurthy M, Lee GM. 2013. Development of recombinant chinese hamster ovary cell lines for therapeutic protein production. *Current Opinion in Chemical Engineering* .

Nuutinen T, Suuronen T, Kauppinen A, Salminen A. 2009. Clusterin: A forgotten player in alzheimer's disease. *Brain Res Rev* 61(2):89-104.

O'Callaghan PM and James DC. 2008. Systems biotechnology of mammalian cell factories. *Briefings in Functional Genomics & Proteomics* 7(2):95-110.

Oda Y, Aishima S, Morimatsu K, Hayashi A, Shindo K, Fujino M, Mizuuchi Y, Hattori M, Tanaka M, Oda Y. 2013. Differential ezrin and phosphorylated ezrin expression profiles between pancreatic intraepithelial neoplasia, intraductal papillary mucinous neoplasm, and invasive ductal carcinoma of the pancreas. *Hum Pathol* .

Ohnishi T, Wang X, Ohnishi K, Takahashi A. 1998. p53-dependent induction of WAF1 by cold shock in human glioblastoma cells. *Oncogene* 16(11):1507-11.

Old WM, Meyer-Arendt K, Aveline-Wolf L, Pierce KG, Mendoza A, Sevinsky JR, Resing KA, Ahn NG. 2005. Comparison of label-free methods for quantifying human proteins by shotgun proteomics. *Molecular & Cellular Proteomics* 4(10):1487.

Olsen JV and Macek B. 2009. High accuracy mass spectrometry in large-scale analysis of protein phosphorylation. In: *Mass spectrometry of proteins and peptides*. Springer. 131 p.

Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, Mann M. 2006. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127(3):635-48.

Olsen JV and Mann M. 2013. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Mol Cell Proteomics* 12(12):3444-52.

Olsen JV, Vermeulen M, Santamaria A, Kumar C, Miller ML, Jensen LJ, Gnäd F, Cox J, Jensen TS, Nigg EA, et al. 2010. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Sci Signal* 3(104):ra3.

Öncül AA, Kalmbach A, Genzel Y, Reichl U, Thévenin D. 2010. Characterization of flow conditions in 2 L and 20 L wave bioreactors® using computational fluid dynamics. *Biotechnol Prog* 26(1):101-10.

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, 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(5):376.

Ooi B, Mukhopadhyay A, Masilamani J, Do D, Lim C, Cao X, Lim I, Mao L, Ren H, Nakamura H. 2010. Hepatoma-derived growth factor and its role in keloid pathogenesis. *J Cell Mol Med* 14(6a):1328-37.

Opstal A, Bijvelt J, Donselaar E, Humbel BM, Boonstra J. 2012. Inhibition of protein kinase B activity induces cell cycle arrest and apoptosis during early G1 phase in CHO cells. *Cell Biol Int* 36(4):357-65.

Orsburn BC, Stockwin LH, Newton DL. 2011. Challenges in plasma membrane phosphoproteomics. *Expert Review of Proteomics* 8(4):483-94.

- Ow SY, Salim M, Noirel J, Evans C, Rehman I, Wright PC. 2009. iTRAQ underestimation in simple and complex mixtures: “the good, the bad and the ugly”. *Journal of Proteome Research* 8(11):5347-55.
- Pai C, Chen H, Sheu H, Yeh N. 1995. Cell-cycle-dependent alterations of a highly phosphorylated nucleolar protein p130 are associated with nucleologenesis. *J Cell Sci* 108(5):1911-20.
- Pak S, Hunt S, Bridges M, Sleight M, Gray P. 1996. Super-CHO—A cell line capable of autocrine growth under fully defined protein-free conditions. *Cytotechnology* 22(1):139-46.
- Palamarchuk A, Efanov A, Maximov V, Aqeilan RI, Croce CM, Pekarsky Y. 2005. Akt phosphorylates and regulates Pcd4 tumor suppressor protein. *Cancer Res* 65(24):11282-6.
- Park H, Kim I, Kim I, Kim K, Kim H. 2000. Expression of carbamoyl phosphate synthetase I and ornithine transcarbamoylase genes in chinese hamster ovary< i> dhfr</i>-cells decreases accumulation of ammonium ion in culture media. *J Biotechnol* 81(2):129-40.
- Patel VJ, Thalassinos K, Slade SE, Connolly JB, Crombie A, Murrell JC, Scrivens JH. 2009. A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *Journal of Proteome Research* 8(7):3752-9.
- Patil M, Pabla N, Ding H, Dong Z. 2013. Nek1 interacts with Ku80 to assist chromatin loading of replication factors and S-phase progression. *Cell Cycle* 12(16):2608-16.
- Pelegri AL, Moura DJ, Brenner BL, Ledur PF, Maques GP, Henriques JA, Saffi J, Lenz G. 2010. Nek1 silencing slows down DNA repair and blocks DNA damage-induced cell cycle arrest. *Mutagenesis* 25(5):447-54.
- Peter CP, Suzuki Y, Büchs J. 2006. Hydromechanical stress in shake flasks: Correlation for the maximum local energy dissipation rate. *Biotechnol Bioeng* 93(6):1164-76.
- Picotti P, Bodenmiller B, Mueller LN, Domon B, Aebersold R. 2009. Full dynamic range proteome analysis of< i> S. cerevisiae</i> by targeted proteomics. *Cell* 138(4):795-806.
- Piersma SR, Fiedler U, Span S, Lingnau A, Pham TV, Hoffmann S, Kubbutat MH, Jiménez CR. 2010. Workflow comparison for label-free, quantitative secretome proteomics for cancer biomarker discovery: Method evaluation, differential analysis, and verification in serum. *Journal of Proteome Research* 9(4):1913-22.
- Pilotte J, Dupont-Versteegden EE, Vanderklish PW. 2011. Widespread regulation of miRNA biogenesis at the dicer step by the cold-inducible RNA-binding protein, RBM3. *PloS One* 6(12):e28446.
- Pitt JJ. 2009. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *The Clinical Biochemist Reviews* 30(1):19.

- Posewitz MC and Tempst P. 1999. Immobilized gallium (III) affinity chromatography of phosphopeptides. *Anal Chem* 71(14):2883-92.
- Ptacek J, Devgan G, Michaud G, Zhu H, Zhu X, Fasolo J, Guo H, Jona G, Breitskreutz A, Sopko R. 2005. Global analysis of protein phosphorylation in yeast. *Nature* 438(7068):679-84.
- Puck TT, Cieciura SJ, Robinson A. 1958. Genetics of somatic mammalian cells III. long-term cultivation of euploid cells from human and animal subjects. *J Exp Med* 108(6):945-56.
- Pyronnet S, Dostie J, Sonenberg N. 2001. Suppression of cap-dependent translation in mitosis. *Genes Dev* 15(16):2083-93.
- Qian WJ, Jacobs JM, Liu T, Camp DG, Smith RD. 2006. Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Molecular & Cellular Proteomics* 5(10):1727.
- Qiao L, Bi H, Busnel J, Hojeij M, Mendez M, Liu B, Girault HH. 2010. Controlling the specific enrichment of multi-phosphorylated peptides on oxide materials: Aluminium foil as a target plate for laser desorption ionization mass spectrometry. *Chemical Science* 1(3):374-82.
- Rader R. 2012. FDA biopharmaceutical product approvals and trends. Biotechnology Information Institute, [www.Biopharma.com/approvals\\_2011.html](http://www.Biopharma.com/approvals_2011.html), Accessed Jan 16.
- Raghavan M, Wijeyesakere SJ, Peters LR, Del Cid N. 2013. Calreticulin in the immune system: Ins and outs. *Trends Immunol* 34(1):13-21.
- Rathore AS, Sobacke S, Kocot T, Morgan D, Dufield R, Mozier N. 2003. Analysis for residual host cell proteins and DNA in process streams of a recombinant protein product expressed in escherichia coli cells. *J Pharm Biomed Anal* 32(6):1199-211.
- Ray K, Hines CS, Coll-Rodriguez J, Rodgers DW. 2004. Crystal structure of human thimet oligopeptidase provides insight into substrate recognition, regulation, and localization. *J Biol Chem* 279(19):20480-9.
- Rebecchi KR, Wenke JL, Go EP, Desaire H. 2009. Label-free quantitation: A new glycoproteomics approach. *J Am Soc Mass Spectrom* 20(6):1048-59.
- Reid C, Tait A, Baldascini H, Mohindra A, Racher A, Bilsborough S, Smales C, Hoare M. 2010. Rapid whole monoclonal antibody analysis by mass spectrometry: An ultra scale-down study of the effect of harvesting by centrifugation on the post-translational modification profile. *Biotechnol Bioeng* 107(1):85-95.
- Reynolds JE, Li J, Craig RW, Eastman A. 1996. BCL-2 and MCL-1 expression in chinese hamster ovary cells inhibits intracellular acidification and apoptosis induced by staurosporine. *Exp Cell Res* 225(2):430-6.
- Robert F, Bierau H, Rossi M, Agugiaro D, Soranzo T, Broly H, Mitchell-Logean C. 2009. Degradation of an Fc-fusion recombinant protein by host cell proteases: Identification of a CHO cathepsin D protease. *Biotechnol Bioeng* 104(6):1132-41.

- Rock KL and Goldberg AL. 1999. Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annu Rev Immunol* 17(1):739-79.
- Rodrigues ME, Costa AR, Henriques M, Azeredo J, Oliveira R. 2010. Technological progresses in monoclonal antibody production systems. *Biotechnol Prog* 26(2):332-51.
- Rodriguez J, Spearman M, Huzel N, Butler M. 2005. Enhanced production of monomeric interferon- $\beta$  by CHO cells through the control of culture conditions. *Biotechnol Prog* 21(1):22-30.
- Roepstorff P and Fohlman J. 1984. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Mass Spectrom* 11:601.
- Rogers LD and Foster LJ. 2009. Phosphoproteomics—finally fulfilling the promise? *Molecular Biosystems* 5(10):1122-9.
- Roobol A, Carden MJ, Newsam RJ, Smales CM. 2009. Biochemical insights into the mechanisms central to the response of mammalian cells to cold stress and subsequent rewarming. *Febs Journal* 276(1):286-302.
- Roobol A, Roobol J, Carden M, Bastide A, Willis A, Dunn W, Goodacre R, Smales C. 2011. ATR (ataxia telangiectasia mutated-and Rad3-related kinase) is activated by mild hypothermia in mammalian cells and subsequently activates p53. *Biochem J* 435:499-508.
- Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S. 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics* 3(12):1154.
- Ruan Y, Sun L, Hao Y, Wang L, Xu J, Zhang W, Xie J, Guo L, Zhou L, Yun X, et al. 2012. Ribosomal RACK1 promotes chemoresistance and growth in human hepatocellular carcinoma. *J Clin Invest* 122(7):2554-66.
- Ruchaud S, Carmena M, Earnshaw WC. 2007. Chromosomal passengers: Conducting cell division. *Nature Reviews Molecular Cell Biology* 8(10):798-812.
- Russo LC, Castro LM, Gozzo FC, Ferro ES. 2012. Inhibition of thimet oligopeptidase by siRNA alters specific intracellular peptides and potentiates isoproterenol signal transduction. *FEBS Lett* 586(19):3287-92.
- Sadygov RG, Cociorva D, Yates JR. 2004. Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book. *Nature Methods* 1(3):195-202.
- Saito K, Fukuda N, Matsumoto T, Iribe Y, Tsunemi A, Kazama T, Yoshida-Noro C, Hayashi N. 2010. Moderate low temperature preserves the stemness of neural stem cells and suppresses apoptosis of the cells via activation of the cold-inducible RNA binding protein. *Brain Res* 1358:20-9.
- Sakai J, Kojima S, Yanagi K, Kanaoka M. 2005.  $^{18}\text{O}$ -labeling quantitative proteomics using an ion trap mass spectrometer. *Proteomics* 5(1):16-23.

- Sakurai T, Itoh K, Liu Y, Higashitsuji H, Sumitomo Y, Sakamaki K, Fujita J. 2005. Low temperature protects mammalian cells from apoptosis initiated by various stimuli in vitro. *Exp Cell Res* 309(2):264-72.
- Sandberg H, Lütkemeyer D, Kuprin S, Wrangel M, Almstedt A, Persson P, Ek V, Mikaelsson M. 2006. Mapping and partial characterization of proteases expressed by a CHO production cell line. *Biotechnol Bioeng* 95(5):961-71.
- Sandin M, Krogh M, Hansson K, Levander F. 2011. Generic workflow for quality assessment of quantitative label-free LC-MS analysis. *Proteomics* 11(6):1114-24.
- Saric T, Chang S, Hattori A, York IA, Markant S, Rock KL, Tsujimoto M, Goldberg AL. 2002. An IFN- $\gamma$ -induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. *Nat Immunol* 3(12):1169-76.
- Sarma AD, Oehrle NW, Emerich DW. 2008. Plant protein isolation and stabilization for enhanced resolution of two-dimensional polyacrylamide gel electrophoresis. *Anal Biochem* 379(2):192-5.
- Satoh M, Hosoi S, Sato S. 1990. Chinese hamster ovary cells continuously secrete a cysteine endopeptidase. In *In Vitro Cellular & Developmental Biology-Plant* 26(11):1101-4.
- Savignac M, Mellström B, Bébin A, Oliveros JC, Delpy L, Pinaud E, Naranjo JR. 2010. Increased B cell proliferation and reduced ig production in DREAM transgenic mice. *The Journal of Immunology* 185(12):7527-36.
- Savino E, Hu B, Sellers J, Sobjak A, Majewski N, Fenton S, Yang T. 2011a. Development of an in-house, process-specific ELISA for detecting HCP in a therapeutic antibody, part 1. *BioProcess Intl* 9:38-47.
- Savino E, Hu B, Sellers J, Sobjak A, Majewski N, Fenton S, Yang T. 2011b. BioProcess technical. *BioProcess International* 9(4):68-75.
- Savitski MM, Nielsen ML, Zubarev RA. 2006. ModifiComb, a new proteomic tool for mapping stoichiometric post-translational modifications, finding novel types of modifications, and fingerprinting complex protein mixtures. *Mol Cell Proteomics* 5(5):935-48.
- Schatz SM, Kerschbaumer RJ, Gerstenbauer G, Kral M, Dorner F, Scheiflinger F. 2003. Higher expression of Fab antibody fragments in a CHO cell line at reduced temperature. *Biotechnol Bioeng* 84(4):433-8.
- Schmid T, Bajer MM, Blees JS, Eifler LK, Milke L, Rübsamen D, Schulz K, Weigert A, Baker AR, Colburn NH. 2011. Inflammation-induced loss of Pdcd4 is mediated by phosphorylation-dependent degradation. *Carcinogenesis* 32(10):1427-33.
- Schroder M, Matischak K, Friedl P. 2004. Serum- and protein-free media formulations for the Chinese hamster ovary cell line DUKXB11. *J Biotechnol* 108(3):279-92.

- Schulenberg B, Aggeler R, Beechem JM, Capaldi RA, Patton WF. 2003. Analysis of steady-state protein phosphorylation in mitochondria using a novel fluorescent phosphosensor dye. *J Biol Chem* 278(29):27251-5.
- Schwertner D and Kirchner M. 2010. Are generic HCP assays outdated? *BioProcess International* .
- Segu ZM and Mechref Y. 2010. Characterizing protein glycosylation sites through higher-energy C-trap dissociation. *Rapid Communications in Mass Spectrometry* 24(9):1217-25.
- Seth G, Charaniya S, Wiaschin KF, Hu W. 2007. In pursuit of a super producer - alternative paths to high producing recombinant mammalian cells. *Curr Opin Biotechnol* 18(6):557-64.
- Shankar G, Pendley C, Stein KE. 2007. A risk-based bioanalytical strategy for the assessment of antibody immune responses against biological drugs. *Nat Biotechnol* 25(5):555-61.
- Shukla AA and Hinckley P. 2008. Host cell protein clearance during protein A chromatography: Development of an improved column wash step. *Biotechnol Prog* 24(5):1115-21.
- Shukla AA, Hubbard B, Tressel T, Guhan S, Low D. 2007. Downstream processing of monoclonal antibodies--application of platform approaches. *Journal of Chromatography B* 848(1):28-39.
- Shukla AA and Gottschalk U. 2012. Single-use disposable technologies for biopharmaceutical manufacturing. *Trends Biotechnol* .
- Simpson RJ. 2007. Assays with protein arrays. *Cold Spring Harbor Protocols* 2007(3):pdb. top2.
- Sinacore MS, Drapeau D, Adamson S. 2000. Adaptation of mammalian cells to growth in serum-free media. *Mol Biotechnol* 15(3):249-57.
- Sirover MA. 2005. New nuclear functions of the glycolytic protein, glyceraldehyde-3-phosphate dehydrogenase, in mammalian cells. *J Cell Biochem* 95(1):45-52.
- Skoufias DA, Indorato RL, Lacroix F, Panopoulos A, Margolis RL. 2007. Mitosis persists in the absence of Cdk1 activity when proteolysis or protein phosphatase activity is suppressed. *J Cell Biol* 179(4):671-85.
- Slade PG, Hajivandi M, Bartel CM, Gorfien SF. 2012. Identifying the CHO secretome using mucin-type O-linked glycosylation and click-chemistry. *Journal of Proteome Research* 11(12):6175-86.
- Slikker W, Desai VG, Duhart H, Feuers R, Imam SZ. 2001. Hypothermia enhances *bcl-2* expression and protects against oxidative stress-induced cell death in chinese hamster ovary cells. *Free Radical Biology and Medicine* 31(3):405-11.
- Smejkal GB, Robinson MH, Lazarev A. 2004. Comparison of fluorescent stains: Relative photostability and differential staining of proteins in two-dimensional gels. *Electrophoresis* 25(15):2511-9.

- Solá RJ and Griebenow K. 2010. Glycosylation of therapeutic proteins. *BioDrugs* 24(1):9-21.
- Spens E and Häggström L. 2007. Defined protein and animal component-free NS0 fed-batch culture. *Biotechnol Bioeng* 98(6):1183-94.
- Spens E and Häggström L. 2005. Protease activity in protein-free NS0 myeloma cell cultures. *In Vitro Cellular & Developmental Biology-Animal* 41(10):330-6.
- Spens E and Häggström L. 2009. Proliferation of NS0 cells in protein-free medium: The role of cell-derived proteins, known growth factors and cellular receptors. *J Biotechnol* 141(3):123-9.
- Spens E and Häggström L. 2005. Defined Protein-Free NS0 myeloma cell cultures: Stimulation of proliferation by conditioned medium factors. *Biotechnol Prog* 21(1):87-95.
- Stapulionis R, Kolli S, Deutscher MP. 1997. Efficient mammalian protein synthesis requires an intact F-actin system. *J Biol Chem* 272(40):24980-6.
- Stastna M and Van Eyk JE. 2012. Analysis of protein isoforms: Can we do it better? *Proteomics* 12(19-20):2937-48.
- Steen H and Mann M. 2004. The ABC's (and XYZ's) of peptide sequencing. *Nature Reviews Molecular Cell Biology* 5(9):699-711.
- Stein S, Thomas EK, Herzog B, Westfall MD, Rocheleau JV, Jackson RS, Wang M, Liang P. 2004. NDRG1 is necessary for p53-dependent apoptosis. *J Biol Chem* 279(47):48930-40.
- Stevenson SE, Chu Y, Ozias-Akins P, Thelen JJ. 2009. Validation of gel-free, label-free quantitative proteomics approaches: Applications for seed allergen profiling. *Journal of Proteomics* 72(3):555-66.
- Stoevesandt O, Taussig MJ, He M. 2009. Protein microarrays: High-throughput tools for proteomics. .
- Stracke JO, Hutton M, Stewart M, Pendas AM, Smith B, Lopez-Otin C, Murphy G, Knauper V. 2000. Biochemical characterization of the catalytic domain of human matrix metalloproteinase 19. evidence for a role as a potent basement membrane degrading enzyme. *J Biol Chem* 275(20):14809-16.
- Strojnik T, Kavalar R, Trinkaus M, Lah TT. 2005. Cathepsin L in glioma progression: Comparison with cathepsin B. *Cancer Detect Prev* 29(5):448-55.
- Strome SE, Sausville EA, Mann D. 2007. A mechanistic perspective of monoclonal antibodies in cancer therapy beyond target-related effects. *Oncologist* 12(9):1084-95.
- Sunley K and Butler M. 2010. Strategies for the enhancement of recombinant protein production from mammalian cells by growth arrest. *Biotechnol Adv* 28(3):385-94.
- Sunstrom NAS, Gay RD, Wong DC, Kitchen NA, DeBoer L, Gray PP. 2000. Insulin-Like growth Factor-I and transferrin mediate growth and survival of chinese hamster ovary cells. *Biotechnol Prog* 16(5):698-702.



- Suriano G, Oliveira C, Ferreira P, Machado JC, Bordin MC, De Wever O, Bruyneel EA, Moguilevsky N, Grehan N, Porter TR. 2003. Identification of CDH1 germline missense mutations associated with functional inactivation of the E-cadherin protein in young gastric cancer probands. *Hum Mol Genet* 12(5):575-82.
- Swaney DL, Wenger CD, Coon JJ. 2010. Value of using multiple proteases for large-scale mass spectrometry-based proteomics. *Journal of Proteome Research* 9(3):1323-9.
- Sweet SM, Bailey CM, Cunningham DL, Heath JK, Cooper HJ. 2009. Large scale localization of protein phosphorylation by use of electron capture dissociation mass spectrometry. *Mol Cell Proteomics* 8(5):904-12.
- Tait A, Hogwood C, Smales C, Bracewell D. 2012. Host cell protein dynamics in the supernatant of a mAb producing CHO cell line. *Biotechnol Bioeng* 109(4):971-82.
- Tammela T, Enholm B, Alitalo K, Paavonen K. 2005. The biology of vascular endothelial growth factors. *Cardiovasc Res* 65(3):550-63.
- Tan HK, Lee MM, Yap MGS, Wang DIC. 2008. Overexpression of cold-inducible RNA-binding protein increases interferon-gamma production in chinese-hamster ovary cells. *Biotechnol Appl Biochem* 49:247-57.
- Tarrant RD, Velez-Suberbie ML, Tait AS, Smales CM, Bracewell DG. 2012. Host cell protein adsorption characteristics during protein a chromatography. *Biotechnol Prog* 28(4):1037-44.
- Taus T, Köcher T, Pichler P, Paschke C, Schmidt A, Henrich C, Mechtler K. 2011. Universal and confident phosphorylation site localization using phosphoRS. *Journal of Proteome Research* 10(12):5354-62.
- Teige M, Weidemann R, Kretzmer G. 1994. Problems with serum-free production of antithrombin III regarding proteolytic activity and product quality. *J Biotechnol* 34(1):101-5.
- Thiel KA. 2004. Biomanufacturing, from bust to boom... to bubble? *Nat Biotechnol* 22(11):1365-72.
- Thingholm TE, Jensen ON, Robinson PJ, Larsen MR. 2008. SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Molecular & Cellular Proteomics* 7(4):661-71.
- Thömmes J and Etzel M. 2007. Alternatives to chromatographic separations. *Biotechnol Prog* 23(1):42-5.
- Thornton BR, Ng TM, Matyskiela ME, Carroll CW, Morgan DO, Toczyski DP. 2006. An architectural map of the anaphase-promoting complex. *Genes Dev* 20(4):449-60.
- Tisdale EJ. 2002. Glyceraldehyde-3-phosphate dehydrogenase is phosphorylated by protein kinase C $\alpha$  and plays a role in microtubule dynamics in the early secretory pathway. *J Biol Chem* 277(5):3334-41.

Tisdale EJ and Artalejo CR. 2007. A GAPDH mutant defective in Src-Dependent tyrosine phosphorylation impedes Rab2-Mediated events. *Traffic* 8(6):733-41.

Tong G, Endersfelder S, Rosenthal L, Wollersheim S, Sauer IM, Bühner C, Berger F, Schmitt KRL. 2013. Effects of moderate and deep hypothermia on RNA-binding proteins RBM3 and CIRP expressions in murine hippocampal brain slices. *Brain Res* .

Tran BQ, Hernandez C, Waridel P, Potts A, Barblan J, Lisacek F, Quadroni M. 2010. Addressing trypsin bias in large scale (phospho) proteome analysis by size exclusion chromatography and secondary digestion of large post-trypsin peptides. *Journal of Proteome Research* 10(2):800-11.

Trexler M, Bányai L, Patthy L. 2001. A human protein containing multiple types of protease-inhibitory modules. *Proceedings of the National Academy of Sciences* 98(7):3705.

Trexler-Schmidt M, Sargis S, Chiu J, Sze-Khoo S, Mun M, Kao YH, Laird MW. 2010. Identification and prevention of antibody disulfide bond reduction during cell culture manufacturing. *Biotechnol Bioeng* 106(3):452-61.

Tropak MB and Mahuran D. 2007. Lending a helping hand, screening chemical libraries for compounds that enhance  $\beta$ -hexosaminidase A activity in GM2 gangliosidosis cells. *FEBS Journal* 274(19):4951-61.

Tsai Y, Lin C, Chen H, Lee K, Hsu C, Yang S, Yeh N. 2008. Chromatin tethering effects of hNopp140 are involved in the spatial organization of nucleolus and the rRNA gene transcription. *J Biomed Sci* 15(4):471-86.

Tscheliessnig AL, Konrath J, Bates R, Jungbauer A. 2013. Host cell protein analysis in therapeutic protein bioprocessing—methods and applications. *Biotechnology Journal* .

Tse-Wen Chang. 1983. Binding of cells to matrixes of distinct antibodies coated on solid surface. *J Immunol Methods* 65(1–2):217-23.

Turtoi A, Mazzucchelli G, De Pauw E. 2010. Isotope coded protein label quantification of serum proteins--comparison with the label-free LC-MS and validation using the MRM approach. *Talanta* 80(4):1487-95.

Ullah M, Burns T, Bhalla A, Beltz H, Greller G, Adams T. 2008. Disposable bioreactors for cells and microbes. .

Underhill MF and Smales CM. 2007. The cold-shock response in mammalian cells: Investigating the HeLa cell cold-shock proteome. *Cytotechnology* 53(1):47-53.

Underhill MF, Birch JR, Smales CM, Naylor LH. 2005. eIF2 $\alpha$  phosphorylation, stress perception, and the shutdown of global protein synthesis in cultured CHO cells. *Biotechnol Bioeng* 89(7):805-14.

Usukura K, Kasamatsu A, Okamoto A, Kouzu Y, Higo M, Koike H, Sakamoto Y, Ogawara K, Shiiba M, Tanzawa H. 2013. Tripeptidyl peptidase II in human oral squamous cell carcinoma. *J Cancer Res Clin Oncol* 139(1):123-30.

- Van Erp R, Adorf M, Van Sommeren A, Gribnau T. 1991. Monitoring of the production of monoclonal antibodies by hybridomas. part II: Characterization and purification of acid proteases present in cell culture supernatant. *J Biotechnol* 20(3):249-61.
- Velez-Suberbie ML, Tarrant RD, Tait AS, Spencer DI, Bracewell DG. 2012. Impact of aeration strategy on CHO cell performance during antibody production. *Biotechnol Prog* .
- Vergara M, Becerra S, Diaz-Barrera A, Berrios J, Altamirano C. 2012. Simultaneous environmental manipulations in semi-perfusion cultures of CHO cells producing rhTPA. *EJB* 15(6):2-.
- Verrecchia F and Mauviel A. 2007. Transforming growth factor-beta and fibrosis. *World Journal of Gastroenterology* 13(22):3056-62.
- Villén J, Beausoleil SA, Gygi SP. 2008. Evaluation of the utility of neutral-loss-dependent MS3 strategies in large-scale phosphorylation analysis. *Proteomics* 8(21):4444-52.
- Villen J, Beausoleil SA, Gerber SA, Gygi SP. 2007. Large-scale phosphorylation analysis of mouse liver. *Proc Natl Acad Sci U S A* 104(5):1488-93.
- Volarević S and Thomas G. 2000. Role of S6 phosphorylation and S6 kinase in cell growth. *Prog Nucleic Acid Res Mol Biol* 65:101-27.
- von Marschall Z and Fisher LW. 2010. Dentin sialophosphoprotein (DSPP) is cleaved into its two natural dentin matrix products by three isoforms of bone morphogenetic protein-1 (BMP1). *Matrix Biology* 29(4):295-303.
- Wagstaff JL, Masterton RJ, Povey JF, Smales CM, Howard MJ. 2013. <sup>1</sup>H NMR spectroscopy profiling of metabolic reprogramming of chinese hamster ovary cells upon a temperature shift during culture. *PLoS One* 8(10):e77195.
- Walczynski J, Lyons S, Jones N, Breitwieser W. 2013. Sensitisation of c-MYC-induced B-lymphoma cells to apoptosis by ATF2. *Oncogene* .
- Walsh G and Jefferis R. 2006. Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol* 24(10):1241-52.
- Walther TC and Mann M. 2010. Mass spectrometry-based proteomics in cell biology. *J Cell Biol* 190(4):491-500.
- Waltz E. 2005. GlaxoSmithKline cancer drug threatens herceptin market. *Nat Biotechnol* 23(12):1453-4.
- Wang X, Zhang S, Zhang J, Lam E, Liu X, Sun J, Feng L, Lu H, Yu J, Jin H. 2013. Annexin A6 is down-regulated through promoter methylation in gastric cancer. *American Journal of Translational Research* 5(5):555.
- Wang Y, Chen W, Wu J, Guo Y, Xia X. 2007. Highly efficient and selective enrichment of phosphopeptides using porous anodic alumina membrane for MALDI-TOF MS analysis. *J Am Soc Mass Spectrom* 18(8):1387-95.

Wang Z, Chen J, Sun J, Cui Z, Wu H. 2012. RNA interference-mediated silencing of eukaryotic translation initiation factor 3, subunit B (EIF3B) gene expression inhibits proliferation of colon cancer cells. *World Journal of Surgical Oncology* 10(1):1-9.

Waszkiewicz N, Szajda SD, Waszkiewicz M, Wojtulewska-Supron A, Szulc A, Kępką A, Chojnowska S, Dadan J, Ładny JR, Zwierz K. 2013. The activity of serum beta-galactosidase in colon cancer patients with a history of alcohol and nicotine dependence: Preliminary data. *Advances in Hygiene & Experimental Medicine/Postepy Higieny i Medycyny Doswiadczalnej* 67.

Wei YC, Naderi S, Meshram M, Budman H, Scharer JM, Ingalls BP, McConkey BJ. 2011. Proteomics analysis of chinese hamster ovary cells undergoing apoptosis during prolonged cultivation. *Cytotechnology* 63(6):663-77.

Wickström D, Wagner S, Simonsson P, Pop O, Baars L, Ytterberg AJ, van Wijk KJ, Luirink J, de Gier JL. 2011. Characterization of the consequences of YidC depletion on the inner membrane proteome of *E. coli* using 2D blue Native/SDS-PAGE. *J Mol Biol* 409(2):124-35.

Wiera G, Wójtowicz T, Lebida K, Piotrowska A, Drulis-Fajdasz D, Gomulkiwicz A, Gendosz D, Podhorska-Okolów M, Capogna M, Wilczyński G. 2012. Long term potentiation affects intracellular metalloproteinases activity in the mossy fiber—CA3 pathway. *Molecular and Cellular Neuroscience* 50(2):147-59.

Wiese H, Kuhlmann K, Wiese S, Stoepel NS, Pawlas M, Meyer HE, Stephan C, Eisenacher M, Drepper F, Warscheid B. 2013. Comparison of alternative MS/MS and bioinformatics approaches for confident phosphorylation site localization. *Journal of Proteome Research* .

Wimmer K, Harant H, Reiter M, Blüml G, Gaida T, Katinger H. 1994. Two-dimensional gel electrophoresis for controlling and comparing culture supernatants of mammalian cell culture productions systems. *Cytotechnology* 16(3):137-46.

Winefield RD, Williams TD, Himes RH. 2009. A label-free mass spectrometry method for the quantification of protein isotopes. *Anal Biochem* 395(2):217-23.

Wlaschin KF and Hu W. 2007. A scaffold for the chinese hamster genome. *Biotechnol Bioeng* 98(2):429-39.

Wolter T and Richter A. 2005. Assays for controlling host-cell impurities in biopharmaceuticals. *Bioprocess International* 2:2-6.

Woolley JF and Al-Rubeai M. 2009. The isolation and identification of a secreted biomarker associated with cell stress in serum-free CHO cell culture. *Biotechnol Bioeng* 104(3):590-600.

Wu CC, MacCoss MJ, Howell KE, Yates JR. 2003. A method for the comprehensive proteomic analysis of membrane proteins. *Nat Biotechnol* 21(5):532-8.

Wu M, Massaeli H, Durston M, Mesaali N. 2007. Differential expression and activity of matrix metalloproteinase-2 and-9 in the calreticulin deficient cells. *Matrix Biology* 26(6):463-72.

- Wu Y, Wang Y, Sun Y, Zhang L, Wang D, Ren F, Chang D, Chang Z, Jia B. 2010a. RACK1 promotes bax oligomerization and dissociates the interaction of bax and bcl-XL. *Cell Signal* 22(10):1495-501.
- Wu SJ, Luo J, O'Neil KT, Kang J, Lacy ER, Canziani G, Baker A, Huang M, Tang QM, Raju TS, et al. 2010b. Structure-based engineering of a monoclonal antibody for improved solubility. *Protein Eng Des Sel* 23(8):643-51.
- Wurm FM. 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat Biotechnol* 22(11):1393-8.
- Xie J, Xu L, Xie Y, Zhang H, Cai W, Zhou F, Shen Z, Li E. 2009. Roles of ezrin in the growth and invasiveness of esophageal squamous carcinoma cells. *International Journal of Cancer* 124(11):2549-58.
- Xu X, Nagarajan H, Lewis NE, Pan S, Cai Z, Liu X, Chen W, Xie M, Wang W, Hammond S. 2011. The genomic sequence of the chinese hamster ovary (CHO)-K1 cell line. *Nat Biotechnol* 29(8):735-41.
- Xue H, Lu B, Lai M. 2008. The cancer secretome: A reservoir of biomarkers. *Journal of Translational Medicine* 6(1):52.
- Xue S and Barna M. 2012. Specialized ribosomes: A new frontier in gene regulation and organismal biology. *Nature Reviews Molecular Cell Biology* 13(6):355-69.
- Xue Y, Gao X, Cao J, Liu Z, Jin C, Wen L, Yao X, Ren J. 2010. A summary of computational resources for protein phosphorylation. *Current Protein and Peptide Science* 11(6):485-96.
- Yang C and Carrier F. 2001. The UV-inducible RNA-binding protein A18 (A18 hnRNP) plays a protective role in the genotoxic stress response. *J Biol Chem* 276(50):47277.
- Yang L, Wang H, Kornblau S, Graber D, Zhang N, Matthews J, Wang M, Weber D, Thomas S, Shah J. 2010. Evidence of a role for the novel zinc-finger transcription factor ZKSCAN3 in modulating cyclin D2 expression in multiple myeloma. *Oncogene* 30(11):1329-40.
- Yates JR, Ruse CI, Nakorchevsky A. 2009. Proteomics by mass spectrometry: Approaches, advances, and applications. *Annu Rev Biomed Eng* 11:49-79.
- Yates JR, 3rd and Kelleher NL. 2013. Top down proteomics. *Anal Chem* 85(13):6151.
- Yee JC, Gerdtzen ZP, Hu WS. 2009. Comparative transcriptome analysis to unveil genes affecting recombinant protein productivity in mammalian cells. *Biotechnol Bioeng* 102(1):246-63.
- Yee JC, de Leon Gatti M, Philp RJ, Yap M, Hu WS. 2008. Genomic and proteomic exploration of CHO and hybridoma cells under sodium butyrate treatment. *Biotechnol Bioeng* 99(5):1186-204.
- Yogalingam G, Hwang S, Ferreira JC, Mochly-Rosen D. 2013. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) phosphorylation by protein kinase C delta

( $\delta$ PKC) inhibits mitochondrial elimination by lysosomal-like structures following ischemia and reoxygenation-induced injury. *J Biol Chem* .

Yoon SK, Hwang SO, Lee GM. 2004. Enhancing effect of low culture temperature on specific antibody productivity of recombinant chinese hamster ovary cells: Clonal variation. *Biotechnol Prog* 20(6):1683-8.

Yoon SK, Song JY, Lee GM. 2003. Effect of low culture temperature on specific productivity, transcription level, and heterogeneity of erythropoietin in chinese hamster ovary cells. *Biotechnol Bioeng* 82(3):289-98.

Yu L, Zhu Z, Chan KC, Issaq HJ, Dimitrov DS, Veenstra TD. 2007. Improved titanium dioxide enrichment of phosphopeptides from HeLa cells and high confident phosphopeptide identification by cross-validation of MS/MS and MS/MS/MS spectra. *Journal of Proteome Research* 6(11):4150-62.

Zanghi JA, Fussenegger M, Bailey JE. 1999. Serum protects protein-free competent chinese hamster ovary cells against apoptosis induced by nutrient deprivation in batch culture. *Biotechnol Bioeng* 64(1):108-19.

Zarghooni M, Bukovac S, Tropak M, Callahan J, Mahuran D. 2004. An  $\alpha$ -subunit loop structure is required for GM2 activator protein binding by  $\beta$ -hexosaminidase A. *Biochem Biophys Res Commun* 324(3):1048-52.

Zavašnik-Bergant T and Turk B. 2006. Cysteine cathepsins in the immune response. *Tissue Antigens* 67(5):349-55.

Zhang B, VerBerkmoes NC, Langston MA, Uberbacher E, Hettich RL, Samatova NF. 2006. Detecting differential and correlated protein expression in label-free shotgun proteomics. *Journal of Proteome Research* 5(11):2909-18.

Zhang D, Wong LL, Koay E. 2007. Phosphorylation of Ser78 of Hsp27 correlated with HER-2/neu status and lymph node positivity in breast cancer. *Mol Cancer* 6:52.

Zhang Hong and Pelech Steven. 2012. Using protein microarrays to study phosphorylation-mediated signal transduction. *Seminars in cell & developmental biology* Elsevier. 872 p.

Zhang H, Kim JK, Edwards CA, Xu Z, Taichman R, Wang C. 2005. Clusterin inhibits apoptosis by interacting with activated bax. *Nat Cell Biol* 7(9):909-15.

Zhang S, Cai M, Zhang S, Xu S, Chen S, Chen X, Chen C, Gu J. 2002. Interaction of p58PITSLRE, a G2/M-specific protein kinase, with cyclin D3. *J Biol Chem* 277(38):35314-22.

Zhang X, Jing Y, Qin Y, Hunsucker S, Meng H, Sui J, Jiang Y, Gao L, An G, Yang N. 2012. The zinc finger transcription factor ZKSCAN3 promotes prostate cancer cell migration. *Int J Biochem Cell Biol* 44(7):1166-73.

Zhang Y, Fonslow BR, Shan B, Baek M, Yates III JR. 2013. Protein analysis by shotgun/bottom-up proteomics. *Chem Rev* 113(4):2343-94.

Zhang XH, Zhao C, Ma ZA. 2007. The increase of cell-membranous phosphatidylcholines containing polyunsaturated fatty acid residues induces phosphorylation of p53 through activation of ATR. *J Cell Sci* 120(Pt 23):4134-43.

Zhao X, Wang Q, Wang S, Zou X, An M, Zhang X, Ji J. 2013. Citric acid-assisted two-step enrichment with TiO<sub>2</sub> enhances the separation of multi- and mono-phosphorylated peptides and increases phosphoprotein profiling. *Journal of Proteome Research* .

Zhong X and Somers W. 2012. Recent advances in glycosylation modifications in the context of therapeutic glycoproteins. *Integrative Proteomics* .

Zhou H, Di Palma S, Preisinger C, Peng M, Polat AN, Heck AJ, Mohammed S. 2012. Toward a comprehensive characterization of a human cancer cell phosphoproteome. *Journal of Proteome Research* 12(1):260-71.

Zhou JX, Tressel T, Yang X, Seewoester T. 2008. Implementation of advanced technologies in commercial monoclonal antibody production. *Biotechnology Journal* 3(9-10):1185-200.

Zhu L, Zhou R, Mettler S, Wu T, Abbas A, Delaney J, Forte JG. 2007. High turnover of ezrin T567 phosphorylation: Conformation, activity, and cellular function. *American Journal of Physiology-Cell Physiology* 293(3):C874-84.

Zhu H, Hu S, Jona G, Zhu X, Kreiswirth N, Willey BM, Mazzulli T, Liu G, Song Q, Chen P, et al. 2006. Severe acute respiratory syndrome diagnostics using a coronavirus protein microarray. *Proc Natl Acad Sci U S A* 103(11):4011-6.

Zhu-Shimoni J, Yu C, Nishihara J, Wong RM, Gunawan F, Lin M, Krawitz D, Liu P, Sandoval W, Vanderlaan M. 2014. Host cell protein testing by ELISAs and the use of orthogonal methods. *Biotechnol Bioeng* 111(12):2367-79.

Zybailov B, Mosley AL, Sardi ME, Coleman MK, Florens L, Washburn MP. 2006. Statistical analysis of membrane proteome expression changes in *saccharomyces cerevisiae*. *Journal of Proteome Research* 5(9):2339-47.