Proteomic Analysis of Endoplasmic Reticulum Stress in Chinese Hamster Ovary Cells in Culture



A thesis submitted for the award of Ph.D.

By

David Ryan, M.Sc., B.Sc. (Hons)

School of Biotechnology

Dublin City University

The work described herein was conducted under the supervision of

Dr Paula Meleady

School of Biotechnology

December 2025

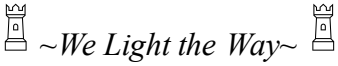
DECLARATION

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, and that I have exercised reasonable care to ensure that the work is original and have conformed to the regulations on the use and declaration of Generative AI, 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. I hereby certify that no Generative Artificial Intelligence (Gen AI) tools have been used in the creation of the thesis.

Signed: David Ryan ID No: 21261088 Date: 28th November 2025

This thesis is dedicated to my grandmother, Jenny Ryan, the person who taught me to follow my own path, despite what others may think, and that the spice of life is just a little bit of divilment.

Acknowledgements



My sincere thanks to my supervisor Dr Paula Meleady for this opportunity and her ongoing support, guidance and patience throughout. Your understanding of when to push me and when to let me go my own way made this process much easier, and your insight and knowledge of this topic allowed me to make the most of the incredible opportunities you gave me.

The advice and Mass Spectrometry knowledge of Dr Michael Henry has been incredibly helpful. Without you Mick, the proteomic work in this thesis would be of a much lower quality, and your patience, expertise and humour made this process entirely more pleasant.

Dr Esen Efeoglu, your training and teaching helped me develop the skills I needed to complete this work and having you to always rely on for support and recommendations when faced with an obstacle was instrumental to my growth as a scientist.

Dr Selvaprakash Karuppuchamy, your guidance on protein preparation and mass spec analysis was much appreciated during the hard work of my ubiquitin study.

My thanks to Gillian Smith and all the technical staff at the Life Sciences Institute for helping facilitate my research.



Christiana Kondylo-Sideri, where to begin. This journey was all the better for having you beside me. You have been a stalwart colleague and friend throughout. I hope the early morning messages asking each other to passage cells, or every post-meeting debrief session meant as much to you as they did to me. It has been a pleasure to grow as a researcher and a person alongside you. However, I hope we never have to share a hotel room again.

My fellow NICB/LSI students and staff, past and present, I thank you for the laughs, distractions and support you gave me. My thanks to Martha Faulkner, for the indelible character that is Muffin; to Niamh Rock, for always reminding me of the bright side; to Dr Dalal AlSultan, for encouraging the worst in me; to Dr Taylor Jade Allen Coyle, for the antics that I could drone on about, some of which were on the fringe of professional; and to Dr Niall Ashfield, the Eddie to my Patsy, the soda to my vodka, love you sis.

 *~Hear Me Roar~* 

To every drag queen, club kid, cabaret artist and burlesque performer I've worked with, thank you for the beauty, pride and glamour you've brought to me.

To Dublin Front Runners, thank you for giving me a community I never imagined I would have, and helping me realise what I am capable of.

 *~ The Old, the True, the Brave~* 

To all the friends that sat through me giving out about things they had no clue about, thank you for your patience and your fake nods of understanding. Your encouragement and acceptance have given me strength during the most challenging times, and your companionship, humour, and perspective reminded me of the world outside the lab. To the recent but rock-solid additions to my life: Conor, Conor, Garret and Tom, I look forward to future adventures.

To those who've been in it for the long run, I could write an essay, but truly, there are not enough words to convey the love and gratitude I have for you all. Claudine, Alexander, Kate, Aisling and Eabha, you are the lights in my life, I treasure you all and I hope I can return your loyalty someday. My world is made infinitely better because I know you.

 *~ Fire and Blood~* 

To all my extended family, especially my grandmother Jenny and my godmother Yvonne, for their steadfast support throughout this process.

To my siblings, Aisling, Seán and Emmett, thank you for putting up with me, and enduring a lifetime of useless trivia and endless rants. I hope the Christmas desserts and cocktails make up for some of it.

I am deeply grateful to my parents, Paul and Teresa, for their unwavering love and support, which has been the foundation of every step I've taken in my continued studies. Your belief in me has always been my greatest encouragement. Having you in my corner has been a vital part of my headstrong confidence and my insistence on living life my own way. You have my eternal love and gratitude, knowing I could not have reached this point without you, and I hope I have made you proud.

Table of Contents

Acknowledgements ..... **iv**

Abbreviations **4**

Research outputs **7**

Abstract **8**

Chapter 1: Introduction **9**

1.1 Chinese hamster ovary cells and their use in the biopharmaceutical industry 9

1.1.1 Biotherapeutic production in CHO cells 9

1.1.2 Optimisation of CHO cells for biopharmaceutical production 13

1.1.3 CHO culture types 15

1.1.4 Key CHO culture conditions 17

1.2 The relationship between culture conditions and CHO productivity 18

1.2.1 The impact of waste product accumulation on productivity 18

1.2.2 Manipulation of CHO culture temperature to increase productivity 19

1.3 Endoplasmic Reticulum (ER) Stress 23

1.3.1 Cellular function of the ER 23

1.3.2 Unfolded Protein Response (UPR) 26

1.3.3 ER-associated degradation (ERAD) 30

1.3.4 The relationship between ER stress and CHO cell culture productivity 32

1.4 Ubiquitin-a crucial regulator of cellular homeostasis 38

1.4.1 Biological structure of ubiquitin 38

1.4.2 The role of ubiquitin in cell signalling pathway regulation 42

1.4.3 The role of ubiquitin in proteasomal degradation targeting 43

1.5 Proteomic methods for uncovering differential protein expression 46

1.5.1 Bottom-up proteomics 46

1.5.2 The use of mass spectrometry in proteomics 46

1.5.3 Quantitative Label-free proteomics 49

1.5.4 Inclusion of PTMs in proteomic studies 50

1.6 Conclusion 56

1.7 References 58

Chapter 2: Paper 1- Label-Free Quantitative Proteomics Analysis of Producer and Non-producer Chinese Hamsters Ovary (CHO) Cells under ER Stress Conditions **82**

2.1 Abstract 86

2.2 Introduction 87

2.3 Materials and Methods 90

2.3.1 Cell Culture 90
2.3.2 Exposure to ER Stressor Thapsigargin 91
2.3.3 Protein Extraction and Western Blotting 91
2.3.4 ELISA 92
2.3.5 Protein extraction and preparation for Mass Spectrometry 93
2.3.6 LC-MS/MS and label-free quantitative differential analysis 94
2.3.7 KEGG pathway and STRING association analysis 95
2.3.8 Statistical Analysis 96
2.4 Results and Discussion 96
2.4.1 CHO cell survival, productivity and expression of stress related proteins under ER stress conditions 96
2.4.2 Proteomic investigation of producer and non-producer CHO cells under ER stress conditions 98
2.4.3 Protein synthesis, folding and secretion under ER stress conditions in producer and non-producer CHO cells. 107
2.4.4 Activation of the Unfolded Protein Response (UPR) under ER stress conditions in producer and non-producer cells 110
2.4.5 Activation of ER associated degradation (ERAD) under ER stress conditions in producer and non-producer cells 112
2.6 References 118
Chapter 3: Paper 2-Proteomic and Ubiquitinated Proteome Insights into ER Stress Responses in Chinese Hamster Ovary Cells Under Mild Hypothermic Conditions .. 125
3.1 Abstract 129
3.2 Introduction 130
3.3 Methodology 137
3.3.1 Cell culture conditions 137
3.3.2 Temperature Shift Culture Conditions 137
3.3.3 ELISA 138
3.3.4 Protein extraction from whole cell lysates 138
3.3.5 Ubiquitinated peptide enrichment 139
3.3.6 LC-MS/MS for Label free quantification (LFQ) 140
3.3.7 Protein identification and differential expression analysis using quantitative label-free LC-MS/MS 141
3.3.8 LC-MS/MS of ubiquitinated peptides and identification of differentially expressed ubiquitinated proteins 142
3.3.9 KEGG pathway and STRING association analysis 143
3.4 Results and Discussion 143
3.4.1 Culture viability and recombinant protein production 143
3.4.2 Gene ontology analysis of differentially expressed proteins in CHO cells under mild hypothermic conditions 146

3.4.3 Gene ontology analysis of differentially expressed ubiquitinated proteins in CHO cells under mild hypothermic conditions 149
3.4.4 Differential expression of ubiquitinated proteins following a downshift in culture temperature 155
3.4.5 Differential expression of ER stress proteins following a downshift in culture temperature 158
3.5 Conclusions 170
3.6 References 172
Chapter 4: Paper 3-Proteomic Analysis of Endoplasmic Reticulum (ER) Protein Expression under Ammonia-Induced Stress in Chinese Hamster Ovary (CHO) Cell Cultures **190**
Chapter 5: Regression and machine learning methods for the analysis of CHO cell mass spectrometry ID (mzID) files and identification of bioprocess relevant stress markers. **237**
5.1. Introduction 237
5.1.1 The growth of proteomics datasets 237
5.1.2 The MZIdentML file format 238
5.1.3 R as a tool for protein bioinformatic analysis 240
5.1.4 Regression 241
5.1.5 Machine Learning 242
5.1.6 ML methods available in R 245
5.2 Methodology 249
5.2.1 MzIdentML file creation 249
5.2.2 Parsing of files into R 249
5.2.3 Logistic regression 250
5.2.4 Random forest 250
5.2.5 Extreme gradient boost 250
5.2.6 Support vector machine 250
5.3 Results 251
5.4 Discussion 258
5.5 Conclusions 264
5.6 References 266
Chapter 6: Overview of major findings and conclusions **275**
Chapter 7: Future Work **284**
Appendix I – Declarations of authorship **1**
Appendix II – Chapter 4 supplementary materials **1**

Abbreviations

AA	Amino acid
ACN	Acetonitrile
AGC	Automatic gain control
ANOVA	One-Way analysis of variance
ATCC	American Type Culture Collection
BSA	Bovine serum albumin
CHO	Chinese Hamster Ovary
CID	Collision induced dissociation
CRIGR	*Cricetulus griseus*
DAVID	Database for Annotation, Visualization, and Integrated Discovery
DDA	Data dependent acquisition
DE	Differential expression
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
DTT	DL-Dithiothreitol
EBI	European Bioinformatics Institute
ELISA	Enzyme-linked immunosorbent assay
EPO	Erythropoietin
ER	Endoplasmic Reticulum
ERAD	Endoplasmic Reticulum Associated Degradation
ESI	Electrospray ionisation
ETD	Electron transfer dissociation
FA	Formic acid
FDR	False discovery rate
GO	Gene ontology

HCD	High collision dissociation
HCP	Host cell protein
HPLC	High pressure liquid chromatography
Ig	Immunoglobulin
IT	Ion trap
KEGG	Kyoto Encyclopaedia of Genes and Genomes
KNN	K-nearest neighbour
LC-MS	Liquid Chromatography-Mass Spectrometry
LF	Label free
m/z	Mass-to-charge
MALDI	Matrix-Assisted Laser Desorption Ionization
MAb	Monoclonal antibody
ML	Machine learning
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
OT	Orbitrap
PBS	Phosphate buffered saline
PPI	Protein-protein interaction
PPM	Parts per million
PSM	Peptide spectral match
PTM	Post-translational modification
RF	Random Forest
RNA	Ribonucleic Acid
RO H_2O	Reverse Osmosis purified Water
RPC	Reverse phase chromatography
SERCA	Sarco endoplasmic reticulum calcium ATPase

STRING	Search Tool for Retrieval of Interacting Genes/Proteins
SUMO	Small ubiquitin-like modifier
SVM	Support vector machine
TFA	Trifluoroacetic acid
TMT	Tandem mass tag
UHP	Ultrapure water
UPR	Unfolded protein response
XCorr	Cross-correlation
XGBoost	Extreme gradient boosting

Research outputs

Papers-First Author

- Ryan, D., Sideri, C.-K., Henry, M., Efeoglu, E. and Meleady, P. (2023)., Label-free quantitative proteomics analysis of producer and non-producer Chinese Hamsters Ovary (CHO) cells under ER stress conditions, Current Research in Biotechnology, Volume 6, 2023, 100141, ISSN 2590-2628, https://doi.org/10.1016/j.crbiot.2023.100141.

- Ryan, D., Sideri, C.-K., Henry, M., Karuppuchamy, S., Efeoglu, E. and Meleady, P. (2026)., Proteomic and Ubiquitinated Proteome Insights into ER Stress Responses in Chinese Hamster Ovary Cells Under Mild Hypothermic Conditions. Biotechnology and Bioengineering, Volume123,1 (2026): 5-25. doi:10.1002/bit.70081

- Ryan, D., Sideri, C.-K., Henry, M., Efeoglu, E. and Meleady, P. (2026)., Proteomic Analysis of Endoplasmic Reticulum (ER) Protein Expression under Ammonia-Induced Stress in Chinese Hamster Ovary (CHO) Cell Cultures. Journal of Proteomics, Volume 328, 2026, 105655, ISSN 1874-3919, https://doi.org/10.1016/j.jprot.2026.105655.

Other publications

- Selvaprakash, K., Sideri, C.-K., Henry, M., Efeoglu, E., Ryan, D. and Meleady, P. (2024), Characterization of the Ubiquitin-Modified Proteome of Recombinant Chinese Hamster Ovary Cells in Response to Endoplasmic Reticulum Stress. Biotechnology. J., 19: e202400413. https://doi-org.dcu.idm.oclc.org/10.1002/biot.202400413

- Selvaprakash, K., Henry, M., Ryan, D. and Meleady, P., 2024. LC-MS/MS Analysis to Study the Ubiquitin-Modified Proteome of Recombinant Chinese Hamster Ovary Cells. In Heterologous Protein Production in CHO Cells: Methods and Protocols (pp. 191-203). New York, NY: Springer US.

Abstract

Title: Proteomic Analysis of Endoplasmic Reticulum Stress in Chinese Hamster Ovary Cells in Culture

Author: David Ryan

Since the 1980s, Chinese hamster ovary (CHO) cells have been central to biopharmaceutical production. These cells can experience endoplasmic reticulum (ER) stress during culture, activating the unfolded protein response (UPR) and potentially limiting recombinant protein yields.

This study used liquid chromatography and tandem mass spectrometry (LC-MS/MS) to analyse proteomic changes in CHO-K1 (non-producer), CHO-SK15-EPO (hEPO producer), CHO-DP12 and NISTCHO (IgG producers) under ER stress. Stress was induced using thapsigargin, reduced culture temperature (37°C to 31°C), and elevated ammonia. Label-free quantitative LC-MS/MS identified differential protein expression in pathways linked to protein folding, UPR, ER-associated degradation (ERAD), RNA transport, aminoacyl-tRNA biosynthesis, and amino acid metabolism. Expression changes were noted in glycosyltransferases, protein disulfide isomerases, and molecular chaperones across both artificially and bioprocess-induced stress conditions.

Immunoaffinity enrichment was applied to study the ubiquitinated proteome in hypothermia-treated cells. Increased ubiquitination of misfolded proteins indicated greater ERAD pathway activity and broader regulation of cellular processes by ubiquitination during cold stress. These findings emphasise the role of post-translational modifications (PTMs) in proteomic studies, as some stress responses were only evident in the ubiquitinated proteome.

Differential expression datasets were analysed using machine learning (Random Forest, Support Vector Machines) to identify a panel of stress-responsive proteins. These proteins represent potential genetic engineering targets to improve stress resistance (HMOX1) or enhance productivity (SRP14). Additionally, the panel could form the basis of an assay for monitoring ER stress in cultures, providing near-real-time feedback on protein processing and aiding bioprocess optimisation.

Chapter 1: Introduction

1.1 Chinese hamster ovary cells and their use in the biopharmaceutical industry

1.1.1 Biotherapeutic production in CHO cells

Chinese hamster ovary (CHO) cells have become indispensable in the biotechnology and pharmaceutical industries as hosts for producing therapeutic proteins and biologics. Since first derived in 1957 (Puck, 1972), the ease of transfection and compatibility with recombinant DNA technologies have made CHO cells an attractive system for producing recombinant proteins (Kim et al., 2012a).

The approval of Genetech's *Activase*, in 1987 by the Food and Drug Administration (FDA), would usher in a new era of biotherapeutic manufacturing, with the CHO cell as the industry's new workhorse. In recent years, the global market has seen intensified demand due to increased biopharmaceutical research and production. In 2025, more than 1,200 biopharmaceutical manufacturing facilities use CHO cell culture, producing over 500 therapeutic biologics, including more than 85% of FDA approved monoclonal antibodies (mAbs) ("CHO Cell Culture Media Market Size, Share | Growth Research - 2033," n.d.).

CHO cells are used to produce a wide range of biopharmaceuticals, including mAbs, clotting factors, hormones and cytokines, examples of which are included in **Figure 1**. The production of mAbs has been particularly successful in CHO cells due to their ability to perform human-like post-translational modifications (PTMs) such as glycosylation (Walsh, 2018). Analysis of the genome sequence of the CHO cell showed that approximately 99% of the genes related to glycan synthesis were homologous between the CHO and human genome (Xu et al., 2011). Proper glycosylation is critical for therapeutic efficacy and reducing immunogenicity (Jefferis, 2005). CHO cells have been used in the production of vaccines for several viral infections, including hepatitis B, respiratory syncytial virus and recurrence of shingles (Liu et al., 2021; Sanchez-Martinez et al., 2024). CHO cells are

transfected with the viral gene of a viral pathogen and the produced protein is then formulated into a subunit vaccine (Strugnell et al., 2011). As it is unable to revert to a virulent strain, this vaccine type is regarded as safer than either inactivated or live-attenuated vaccines (Kushnir et al., 2012).

Recombinant clotting factors, such as Factor VIII for haemophilia treatment, have also been successfully expressed in CHO cells. These complex proteins require extensive PTMs, making CHO cells an ideal expression system (Lialios et al., 2005). CHO cells have been employed in producing cytokines like erythropoietin (EPO), which stimulates red blood cell production in patients with anaemia (Kling et al., 1996). Several types of immunoglobulin G (IgG) have been produced in recombinant CHO cells for use as biopharmaceuticals (Kuwae et al., 2018). Extensive studies exploring the production of these proteins has led to relativity high titres, some up to 9g/L in long term cultures (Bandaranayake and Almo, 2014; Takagi et al., 2017). The CHO cell lines investigated in this thesis produce two different modalities, recombinant human erythropoietin (EPO) and the antibody IgG, alongside a non-producer cell line. A crucial part of optimising CHO production is understanding the stress effect of recombinant production on the cell. The XBP1-related ER pathways have been identified as a bottleneck in IgG production. Increased expression of XBP1 was found to increase overall antibody titre by 40% (Becker et al., 2008). Transcriptomic analysis of producer CHO cells found ER-stress and autophagy genes to be essential targets for genetic engineering. The resulting overexpression of HSP90B1 and DNAJC3 increased mAb titre by 50%, and reduced negative protein aggregation (Barzadd et al., 2022).

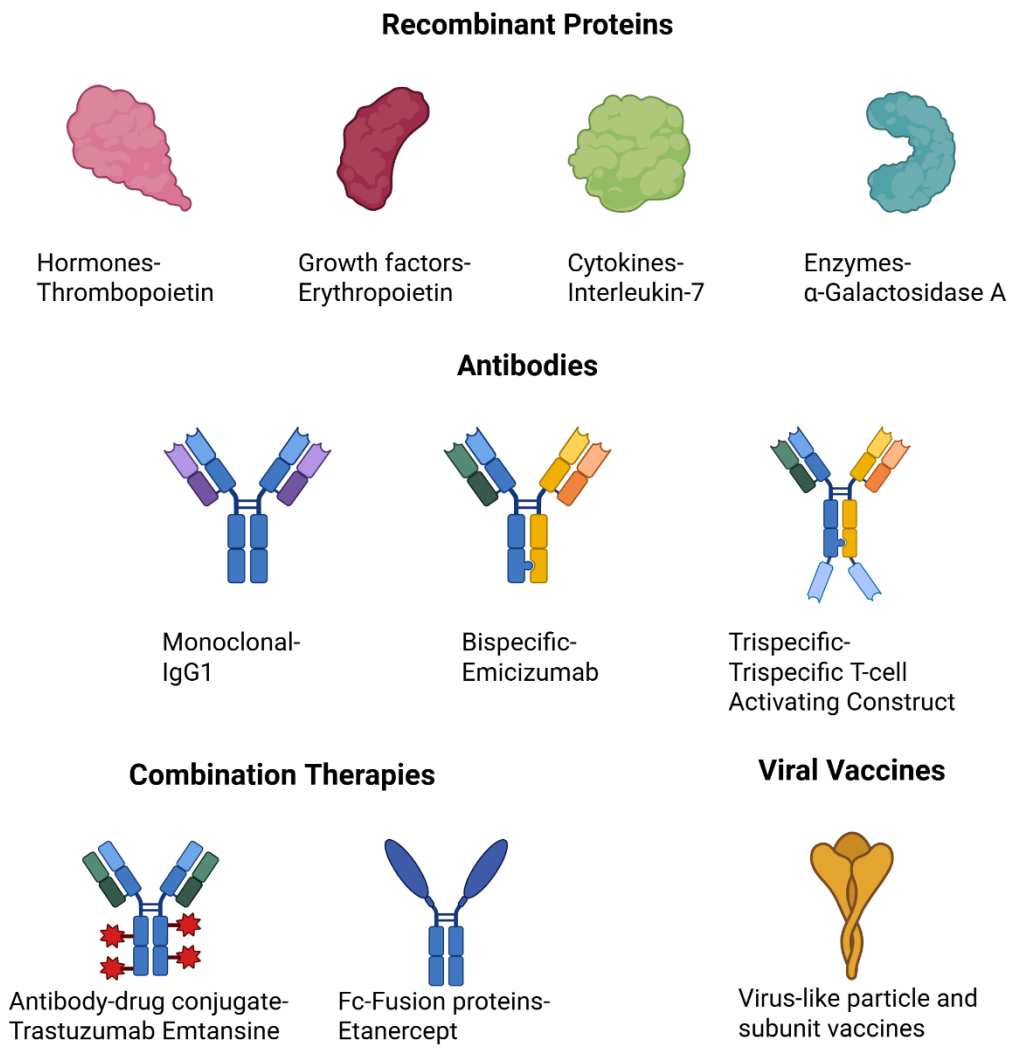


Figure 1: An overview of recombinant biotherapeutic modalities produced in CHO cells. Various biologics are produced in CHO cells, including recombinant enzymes, cytokines and growth factors, monoclonal antibodies (mAbs), bi- and tri-specific antibodies, antibody-drug conjugates (ADCs), Fc-fusion proteins, and viral glycoproteins. A general structure for each modality is included, with an example of this product type made by CHO cells. The variety of CHO cell products requires complex post-translational modifications, proper folding, and secretion, which puts pressure on the ER. *Figure created in BioRender.com*

Certain mAbs have been termed as difficult-to-express (DTE), meaning they produce low volumetric titre and lack certain molecular characteristics important to protein function, such as disulphide bonds (Johari et al., 2015). The exact characteristics that make some mAbs DTE is still being uncovered. Mathias et al (2020) investigated how ER function may be linked to DTE proteins. They identified that within the antibody (IgG1) there was a lack of proper disulfide bridge formation, attributing this to less recognition of the antibody by protein disulfide isomerases (PDI). Therefore, the DTE molecule was unable to fold

correctly, which impaired intracellular transport and led to the protein being degraded by the ubiquitin proteasome system via ER-associated degradation (ERAD) (Mathias et al., 2020). This study directly links ER function and ER stress response mechanisms to product quality, suggesting an avenue for further research into the area of ER function in pursuit of optimising CHO cells for biotherapeutic production. It is possible that the results explored in this thesis could aid in the production of DTE products. With better knowledge of ER stress in CHO cells, we can understand how DTE molecules may put excessive pressure on the ER, leading to errors in folding and protein structure. The ER proteins related to stress may also be potential targets for genetic engineering. This could increase ER folding capacity, opening the options for CHO-derived products.

Bispecific and trispecific antibodies, fusion proteins, and antibody-drug conjugates (ADCs) are considered to be next-generation modalities in CHO-derived products. These products can require the advanced engineering of host cell lines and bioprocesses. These modalities require optimised glycosylation control, enhanced folding capacity, and tailored metabolic pathways to ensure correct assembly and stability (Kim et al., 2012b). Recent strategies involve genome editing, targeted knock-ins, and use of synthetic promoters to improve expression, while intensified perfusion and continuous manufacturing systems address scalability and consistency (Hong et al., 2018). It has been found that bispecific production can lead to ER stress, which can cause aggregation of the recombinant products, which can be mitigated through perfusion culture (Sinharoy et al., 2020). Given the expense and complexity of perfusion culture equipment, cellular adaptations would be more broadly applicable. In order to efficiently adapt CHO cells to the production of these therapeutics, we must fully understand the cell systems utilised in their creation.

1.1.2 Optimisation of CHO cells for biopharmaceutical production

The transition of CHO cells from a research tool to a production workhorse was facilitated by advancements in genetic engineering and cell culture technologies. The development of the dihydrofolate reductase (DHFR) amplification system significantly improved recombinant protein yield (Kaufman and Sharp, 1982). From cell engineering methods to growth medium enhancements, attempts have been made to optimise aspects of CHO production. Pan et al (2017) showed that two IgG producing clones derived from the same parental cell line responded differently to combinations of different mediums and feeds in a fed-batch culture. The two clones had differing metabolic and physiological responses for the different medium-feed combinations. The greatest product titre (1.2 g/L) was obtained with ActiCHO basal medium and feeds, which the study defined as a "rich medium with a rich feed" (Pan et al., 2017). This is just one of the studies that shows the impact of media selection on CHO culture productivity.

The focus of these optimisation studies also aims to improve product quality and expand the range of biotherapeutics produced in CHO cells. The effects of culture temperature, metal ion supplementation, growth factor supplementation and hormone addition on N-glycosylation site-occupancy in CHO cells was studied in 2009, with each condition change allowing for an increasing in site occupancy between 1-4% (Gawlitzek et al., 2009). Proper glycosylation is key to protein function and is a major role of the ER. It has also been found that reducing culture temperature to 35°C and the addition of 80 µM Zn^{2+} can increase mAb titre and increase charge variants of the antibody (Weng et al., 2020). This is a crucial factor of product quality.

There are several state-of-the-art cell biology techniques currently being used to optimise productivity and product quality. The gene sequence of the recombinant antibody is important, with the proportion of heavy- and light- polypeptide chains a difficult factor to

control in transfection (Vanhove et al., 2001). In transient expression of proteins, it is difficult to control the ratio of heavy-chain to light-chain peptides (Schlatter et al., 2005), meaning productivity can be affected, and more stable expression is needed for industry. CHO cell lines have been generated using random integration-based antibody expression, in which expression plasmids are randomly integrated into the host genome after transfection (Ho et al., 2013). Through newer targeted integration methods, the plasmid DNA can be recombined into a pre-determined locus. Recent studies have shown that various targeted integration (TI) methods demonstrate better consistency and similar or better protein productivity than random integration. Carver et al. (2020) developed a TI host utilising their patented 'landing pad at a transcriptional hotspot' which lead to increased mAb titre (Carver et al., 2020). TI using BXB1 integrase was also shown to increase product titre compared to random integration (Inniss et al., 2017). Synthetic promoters have been shown to create single multigene expression vectors co-expressing recombinant genes at a predictable relative stoichiometry in mammalian cells (Patel et al., 2021). This allows for increased transcriptional activity, with a wider application range to traditional CMV and hEF-1α promoters. More recently, CRISPR-Cas9 genome editing has revolutionised CHO cell engineering by enabling precise genetic modifications to enhance productivity and reduce undesirable by-products (Fischer and Otte, 2019).

Recent advancements have focussed on improving the scalability and efficiency of CHO cell-based production systems. High-throughput screening techniques and omics technologies, such as transcriptomics and proteomics, are being used to better understand CHO cell biology and identify novel targets for genetic engineering (Kuo et al., 2018; Fischer et al., 2015). Single-cell RNA sequencing has provided insights into cellular heterogeneity within CHO cultures, enabling the selection of high-performing clones (Clarke et al., 2021). Artificial intelligence (AI) and machine learning (ML) modelling are increasingly utilised in CHO cell-based manufacturing processes to enhance productivity,

process optimisation, and product quality (Zhang et al., 2020). ML algorithms, including support vector machines, random forests, and neural networks are applied to model complex, nonlinear bioprocess parameters, enabling predictive analytics and real-time decision-making (Richter et al., 2025). These models identify critical process parameters and quality attributes, leading to more robust control strategies. These tools are expected to further enhance the efficiency of CHO cell-based manufacturing processes.

1.1.3 CHO culture types

Historically, the preference in the pharmaceutical industry was to culture CHO cells in stainless steel, stirred-tank bioreactors which can reach capacities up to 20,000L (Sieblist et al., 2011) but this preference has shifted to single use vessels which reduce contamination risk (Samaras et al., 2022.). Stirring creates homogeneous conditions, usually achieved using axially pumping impellers (Lemire et al., 2022). The simplest approach to CHO cell culture is batch culture, where all cells are seeded into a fixed amount of medium, and no extra nutrients or cells are provided, with no waste products removed (**Figure 2.A**). The product of interest is harvested when the culture has reached the end of the allocated time (Liang et al., 2024; Tihanyi and Nyitray, 2020). Fed-batch cultures operate in a similar fashion but feed, adapted to provide nutrients required by the specific culture, is added at regular intervals, allowing a greater cell density to develop (Altamirano et al., 2004; Liang et al., 2024) (**Figure 2.B**). Fed-batch culture is the leading culture method used in the production of CHO-derived product, due to the high product titre that can be achieved, exceeding 5g/L (Mahé et al., 2022; Schellenberg et al., 2022a). Continuous culture maintains a constant culture volume, with feed added while spent medium containing cells, product and waste products is being removed, maintaining a steady state (Berrios et al., 2011) (**Figure 2.C**). Continuous culture has been used predominantly to produce unstable products that are liable to degradation with prolonged exposure to bioreactor conditions (Karst et al., 2018; Le et al., 2015). Perfusion culture is also used for the production of labile enzymes or proteins that

are need to be removed as soon as possible from the medium broth (Pollock et al., 2013). In perfusion cultures, spent medium is removed, cells and product are extracted, and the viable cells are returned to the culture, along with fresh medium (Maria et al., 2023) (**Figure 2.D**). In this way, it is possible to achieve a higher cell density than in continuous cultures, increasing productivity and requiring lower bioreactor capacities (Bausch et al., 2019). Perfusion cultures maintain a stable environment which has been shown to lead to more stable cellular events, especially at the protein production level, as well as helping improve the final product quality (Bertrand et al., 2019; Karst et al., 2017). The challenge of perfusion cultures is due to the complex systems needed, using a cell retention device for filtering, and the cost associated with the continuous stream of fresh medium needed (Maria et al., 2023).

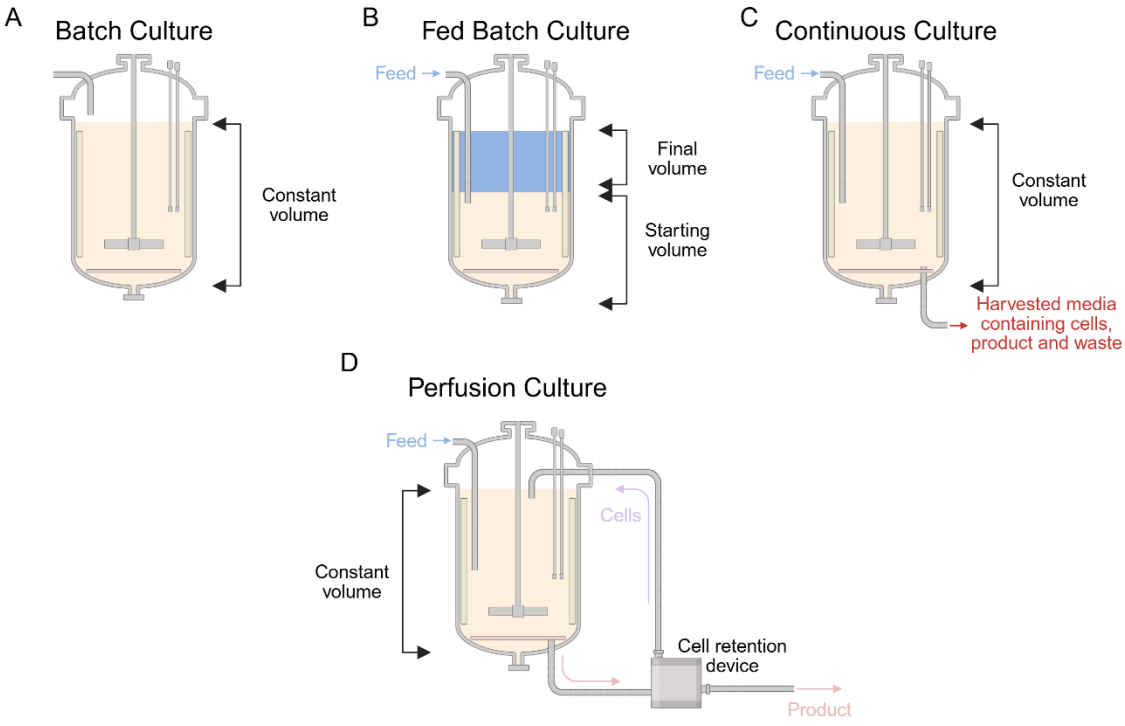


Figure 2: The different types of cultures used in the growing CHO cells for biologic production. A- Batch culture: Cells are grown in a fixed volume of medium without further nutrient addition with no removal of waste products or cells. **B**-Fed-batch culture: Nutrients are incrementally added over time to prolong culture viability and productivity, without removing medium or cells. **C**-Continuous culture: Fresh medium is continuously supplied, and culture medium is simultaneously removed to maintain constant vo lume with cells and waste products also removed. **D**-Perfusion culture: Fresh medium is continuously added, while waste products and spent medium are removed; cells are retained using a cell retention device, allowing for high cell density and extended culture duration. *Figure created in BioRender.com*

1.1.4 Key CHO culture conditions

One of the pivotal developments was the adaptation of CHO cells to serum-free and chemically defined media. This enhanced reproducibility and minimised contamination risks from potential prions or endotoxins associated with animal-derived components (Jayapal et al., 2007). In most formulations, the energy for cells in culture is provided by glucose, which also acts as a carbon source. The CHO cell pathways involved in glucose metabolism are highly efficient, allowing cell specific growth rate, cellular ATP concentration and amino acid metabolic rates to remain stable in cultures with glucose concentration as low as 1.2mM (Lu et al., 2005). High rates of glucose consumption can lead to the accumulation of pyruvate, causing a build-up of lactate in culture media, which can be inhibitory to culture growth and viability (Cruz et al., 2000). Low glucose concentration culture methods have been developed in response to studies showing that high glucose is more likely to lead to lactate accumulation (Chee Furng Wong et al., 2005; Wilkens et al., 2011; Young, 2013).

Amino acid supplementation is also a crucial component of medium formulation, in particular glutamine, which is needed for growth, energy, redox, signalling, and pH regulation (Wahrheit et al., 2014). Glutamine is also necessary for fuelling the tricarboxylic acid (TCA) cycle, as well as acting as a carbon and nitrogen source for pyrimidine production (Becker et al., 2019; Holden et al., 1999). Other amino acids present in CHO cell media include lysine, histidine, leucine, isoleucine, methionine, phenylalanine, threonine, tryptophan, and valine (Fan et al., 2009). Trace metals make up a small, but important percentage of medium formulation. Copper levels in culture can affect lactate consumption and oxidative enzymes, leading to a decrease in cell growth and product titre (Qian et al., 2011).. The addition of zinc in media has been shown to provide a 1.2-fold increase in mAb production (Kim and Park, 2016). Other trace elements include cobalt, manganese, nickel and tin, usually in concentrations under 5 μM, which act as enzyme co-factors for metabolic regulation and cell growth (Ritacco et al., 2018).

1.2 The relationship between culture conditions and CHO productivity

1.2.1 The impact of waste product accumulation on productivity

As a result of glutamine supplementation, ammonia can build up in CHO cell cultures, especially cultures designed to be maintained for prolonged periods. While ammonia will be present in all cultures, compared to other culture conditions, there has been relatively little investigation into its interaction with CHO cells, when compared to other culture conditions. Elevated extracellular ammonia has been shown to reduce viable cell density and overall product yield in CHO cultures. Concentrations above 5 mM are typically considered inhibitory (Yang and Butler, 2000a), but batch cultures can reach concentrations of up to 30 mM (Fan et al., 2015; Xu et al., 2017). Ammonia disrupts intracellular pH regulation, leading to altered homeostasis and impaired mitochondrial function compromising energy metabolism and reducing cell growth rates (Kurano et al., 1990). It has been demonstrated that elevated ammonia reduces productivity of CHO cells (Hansen and Emborg, 1994). Beyond impairing growth and reducing productivity, ammonia alters product quality by disrupting Golgi-mediated glycosylation, particularly galactosylation and sialylation (Gawlitzek et al., 2009). MAbs produced in CHO cells in elevated ammonia cultures display altered glycan profiles, including increased sialylation and reduced galactosylation (Yang and Butler, 2000b). This disruption is linked to perturbed Golgi pH and ionic balance, reduced nucleotide-sugar transport, and the regulation of key enzymes, including increased sialidase activity (Synoground et al., 2021). Strategies to mitigate these effects include optimising glutamine feeding and controlling pH, which fail to fully eliminate the problem, and the use of continuous or perfusion cultures which requires fresh media to replace the ammonia-containing media (Chen and Harcum, 2005; Savizi et al., 2022; Xu et al., 2023). Key gaps remain in our understanding of the effects of elevated ammonia on CHO cultures.

The research presented in **Chapter 4** of this thesis focuses on the impact of elevated ammonia on ER-resident protein expression. Through this investigation, we were able to

distinguish the response of CHO cells to increased ammonia from the response to known ER stress inducers. It was found that proteins within the ER lumen remained relatively unaffected by raised ammonia levels in the early stage of culture, when compared to ER membrane protein expression. Existing knowledge on the effect of ammonia on Golgi-body was supported by the results of this study, which vesicle mediated transport related proteins upregulated in elevated ammonia cultures.

1.2.2 Manipulation of CHO culture temperature to increase productivity

Culture temperature is a carefully monitored variable in bioreactors. CHO cultures, as with most mammalian cell lines, are maintained at 37°C, to mimic body temperature. However, research on several CHO cell lines has shown that lowering culture temperature to between 31-33°C can have favourable effects on productivity. Lower culture temperature is regularly used in industry to improve culture longevity, increase CHO productivity or enhance product quality attributes (Torres et al., 2021). It has been suggested that temperature reduction may provide cleaner downstream feedstocks by reducing the variety of host cell proteins and to stabilise product integrity by reducing the number of proteases (Goey et al., 2017). This method began to gain attention in the early 1990s, with Jenkins and Hovey (1993) being among the first to publish findings on the mild hypothermic conditions. In a batch culture process, they alternated culture temperature between 39°C and 34°C, maintaining high culture viability for 170 h, and resulting in a 35% increase in r-protein yield compared to a culture at 37 °C (Jenkins and Hovey, 1993). Various studies since have demonstrated that lower culture temperatures lead to decreased cell growth and increased productivity in cells producing t-PA, alkaline phosphatase, erythropoietin and anti-4-1BB antibody (Berrios et al., 2009; Kaufmann et al., 1999; Yoon et al., 2005, 2003). The most common practice is to allow a culture to “bulk up” in cell density at 37°C, then lower the temperature at the beginning of the exponential growth phase (Torres et al., 2021).

The lowering of culture temperature arrests cells in the G1 phase of the cell cycle. The productivity increase linked to the arrest of the G1 phase may be associated with the maintenance of cell viability, leading to a delayed start of apoptosis (Moore et al., 1997). This cycle arrest has been linked to the expression of cyclin-dependent kinase inhibitors, including p21 and p27. These inhibit the Cyclin complexes, which are essential for regulating the phase change from G1 to S-phase (Fox et al., 2005). The rate of metabolic activity, glucose consumption, and lactate production is reduced, which also slows progression through the cell cycle (Becerra et al., 2012). It has been found that cells maintained in the G1 phase are larger and more metabolically active than non-arrested cells (Bi et al., 2004). This leads to a reduction in proliferation rate while maintaining cell viability over extended periods (Kaufmann et al., 1999). Temperature reduction also induces significant metabolic reprogramming. CHO cells exposed to mild hypothermia exhibit decreased rates of glucose uptake and lactate production. Amino acid metabolism may be altered, with studies reporting shifts in glutamine consumption and ammonia accumulation (Marchant et al., 2008). These metabolic changes can be associated with a more stable pH environment and reduced accumulation of metabolic waste products, both of which contribute to extended culture durations (Becerra et al., 2012). Transcriptomic and proteomic studies have demonstrated that temperature shift induces a stress-like response in CHO cells, characterised by differential expression of genes involved in protein folding, apoptosis regulation, and cell cycle control (Masterton and Smales, 2014). The expression of ER chaperones and anti-apoptotic factors is upregulated, which may support continued viability and efficient protein folding under sub-optimal conditions (Ji et al., 2022). Additionally, components of the UPR pathway may be modulated (Zhao et al., 2023), although the extent varies in different cell lines.

The increase in cell productivity following temperature downshift has been linked to increased mRNA stability (Oguchi et al., 2006), elevated recombinant gene transcription

(Kou et al., 2011) and increased secretory capacity (Smales et al., 2004). Bedoya-López et al identified 93 genes involved in transcription that were differentially expressed at 48 h after temperature reduction to 30°C, including transcription factors *MYC* and *c-FOS* (Bedoya-López et al., 2016). RT-PCR has been used to prove that mRNA levels increased in proportion to the production of IFN-γ at 32 °C (Fox et al., 2005). Increased mRNA levels are generally by a combination of both enhanced mRNA stability and increased transcriptional activity (Fox et al., 2005; Mason et al., 2014). A reduction of culture temperature to 31°C showed increased PERK expression and phosphorylation of eIF2a, which is linked to an increase in recombinant TNFR-Fc titre (Wang et al., 2018). There was also a reduction in detrimental TNFR-Fc protein aggregation, highlighting how lower culture temperature can increase product quality. The authors linked to phosphorylation of the initiation factor eIF2, which slows global protein synthesis rates, reducing the burden on the ER, allowing for improved protein folding accuracy.

While the benefits of lower culture temperature have been repeatedly proven, there is still no comprehensive explanation as to how this change in culture conditions leads to these benefits. An experiment looking at the ubiquitinated proteome of CHO cells exposed to mild hypothermia forms **Chapter 3** of this thesis. In preliminary work for this study, CHO-SK15-EPO cells were shifted to 31°C culture 48 hours after seeding, with cell samples collected 48- and 120-hours post temperature shift. To explore whether these conditions produced effects in ubiquitination patterns worth investigating, a pan-ubiquitin antibody was used to western blot these samples. The results, shown in **Figure 3**, demonstrate how the lowering of culture temperature leads to an increase in the presence of ubiquitination at the early stages of culture, but by the later stage, ubiquitination levels are similar between temperatures. This indicates that ubiquitination plays an important role in the cellular response to temperature shift. An in-depth proteomic study was then conducted, investigating changes in the whole cell proteome as well as in the ubiquitinated proteome following a reduction in culture

temperature. Including ubiquitination characterisation in this study makes it the first investigation into the ubiquitinated proteome of CHO cells experiencing mild hypothermia. Ubiquitination is a crucial mechanism in the regulation of protein expression, which is explored in detail in chapter 3. To fully understand the stress response of CHO cells, the role of ubiquitination must be included.

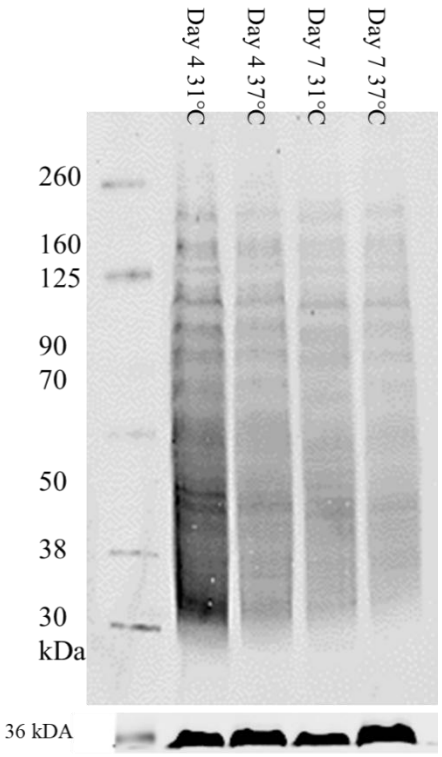


Figure 3: The presence of ubiquitin on various proteins CHO cells experiencing mild hypothermia. Western Blot of ubiquitinated proteins using a pan-ubiquitin antibody to demonstrate levels of ubiquitinated proteins in CHO-SK15-EPO cells in culture at 31°C and 37°C on day 4 and day 7 of culture. GAPDH was used as a loading control.

1.3 Endoplasmic Reticulum (ER) Stress

1.3.1 Cellular function of the ER

The ER is the organelle responsible for protein folding and maturation. There are multiple co-translational and post-translational modifications that occur only in the ER including signal-peptide cleavage, disulphide-bond formation, glycosylation and glycophosphatidylinositol (GPI)-anchor addition. These covalent changes play a vital role in the normal folding of proteins (Ellgaard and Helenius, 2003). Correct folding of the polypeptide chain is crucial for a protein to function properly. In the ER proteins are folded in a different way to the cell cytosol, due to differing topology, chemical composition and in the protein folding machinery present (Schröder and Kaufman, 2005).

The signal recognition particle (SRP) recognises signal sequences on newly synthesised polypeptides present in the cell cytoplasm, and once bound together, protein synthesis slows as translocation takes place. This process is explored in greater detail in **Chapter 4, Figure 9.** The SRP is a crucial element of cellular machinery that is universally conserved, which serves to transport target peptides from ribosomes to cellular membranes, allowing for the typical biogenesis of proteins in the mammalian membrane and secretory pathways (Saraogi and Shan, 2011). Once peptides arrive at the translocon, translation begins once again, as part of a process known as co-translational translocation (Hegde and Kang, 2008; Zimmermann et al., 2011) (**Figure 4.A**) This is how most unfolded polypeptides enter the ER to undergo folding. A smaller subset of proteins enters the ER through posttranslational membrane insertion. Tail-anchored (TA) proteins are translated in the cytosol, and a carboxy-terminal single trans-membrane domain (TMD) is recognised by the cytoplasmic chaperones TRC40, alongside the three-protein complex of BAT3, TRC35, and UBl4A (Borgese and Fasana, 2011). These proteins facilitate the insertion of TA proteins into the ER membrane, assisted by membrane localised GET1/2 receptors (Chartron et al., 2012; Schuldiner et al., 2008) (**Figure 4.B**).

ER-based folding is particularly important for glycoproteins, with 90% of glycoproteins undergoing the covalent addition of N-linked glycans in the ER (Helenius, 1994). N-glycosylation occurs at the consensus sequence Asn-X-Ser/Thr (where X can be any amino acid except Proline), attaching to the nitrogen of Asparagine. There is a core structure to the preassembled oligosaccharide, $Glc_3Man_9GlcNAc_2$, which transfers to the nascent polypeptides. O-linked glycans attach to the oxygen of Ser/Thr, lacking a universal consensus sequence, but often favouring Proline at -1 or +3 positions. More than eight oligosaccharide structures have been linked to O-glycosylation (An and Lebrilla, 2011).. Glucosidases present in the ER continue to process the peptides to generate the monoglucosylated species ($Glc_1Man_9Glc\text{-}NAc_2$) that can be recognised by the membrane-bound lectin-chaperone calnexin (CNX) and its soluble orthologue calreticulin (CRT), both of which are needed to orchestrate the correct folding of the glycosylated polypeptides into glycoproteins (Rutkevich and Williams, 2011) (**Figure 4.C**). The rate-determining step of glycoprotein folding is catalysed by ERP57, forming native intramolecular and intermolecular covalent bonds between cysteine residues (Oliver et al., 1997). When released from calnexin and calreticulin, exo-glucosidase II removes the glucose 1 from the nascent polypeptide, which prevents re-association of the folding polypeptide with glucose-binding lectins, creating the native structure for most glycopeptides at this stage (Soldà et al., 2007). Some polypeptides require further cycles of glucosylation, and this is mediated by the glucosyltransferase UGGT1, which re-glucosylates the mannose residue of peptides to create a peptide that will be recognised by CNX and CRT (Ito et al., 2015). Natively folded, de-mannosylated substrates are glycosylated, then trafficked to the ER membrane and secreted into the cytosol via the ERGIC complex (Ferris et al., 2014) (**Figure 4.D**). Repetitive de-mannosylation from the glycan signals that a polypeptide has not folded correctly. This modification causes terminally misfolded polypeptides to be removed from the CNX cycle and targeted for degradation (Caramelo and Parodi, 2008) (**Figure 4.E**).

BIP (binding immunoglobulin protein) is one of the most abundant ER chaperone proteins and is vital for ER folding, as well as ERAD. BIP contains an ATPase domain and a substrate-binding domain (SBD) which contains a hydrophobic region that binds unfolded proteins (Gething, 1999). Following hydrolysis of ATP, ADP-bound BIP acquires a high affinity for substrates in which the hydrophobic region is closed, preventing unfolded proteins from forming aggregates (Dudek et al., 2009). As such, BiP can recognise unfolded or misfolded regions of the polypeptide, allowing BiP to recognise nascent unfolded polypeptides, as well as misfolded proteins. BiP linkage sends polypeptides to the folded machinery, where BiP should be removed if folding is successful. If BiP is re-attached to an exposed hydrophobic region, the polypeptide is sent to be refolded or targeted for degradation (Adams et al., 2019). There are several DNAJ/HSP40 cochaperones present within the ER which are involved in stimulating ATP hydrolysis of BIP and in regulating its functions in folding and targeting. The ADP-bound BIP is converted to the ATP form by nucleotide exchange factors (NEFs) including GRP170 and SIL1/BAP, which cause BIP to take on an open and accessible form (Kampinga and Craig, 2010).

The formation of stable disulfide bonds on cysteine residues is crucial for protein function and structure, and is mediated by the Protein Disulfide Isomerase (PDI) family of proteins (Kozlov et al., 2010) (**Figure 4.C**). At present, over 20 mammalian PDIs have been identified, each having a different number of thioredoxin-like domains, which are the active elements in the oxidative folding and disulfide bond rearrangement of proteins (Matsusaki et al., 2020). PDIs also exhibit a chaperone function, stabilising unstable conformations of substrates and promoting folding of nascent polypeptide chains (Puig and Gilbert, 1994). PDIs catalyse the isomerisation of substrate proteins through two pathways; direct isomerisation in the substrate molecule and "reduction/oxidation cycle" isomerism mechanism (Wilkinson and Gilbert, 2004). This redox activity is a central part of ER quality

control, as PDIs interact directly with substrates and work with ER oxidoreductases to maintain oxidative protein folding homeostasis (Benham, 2012).

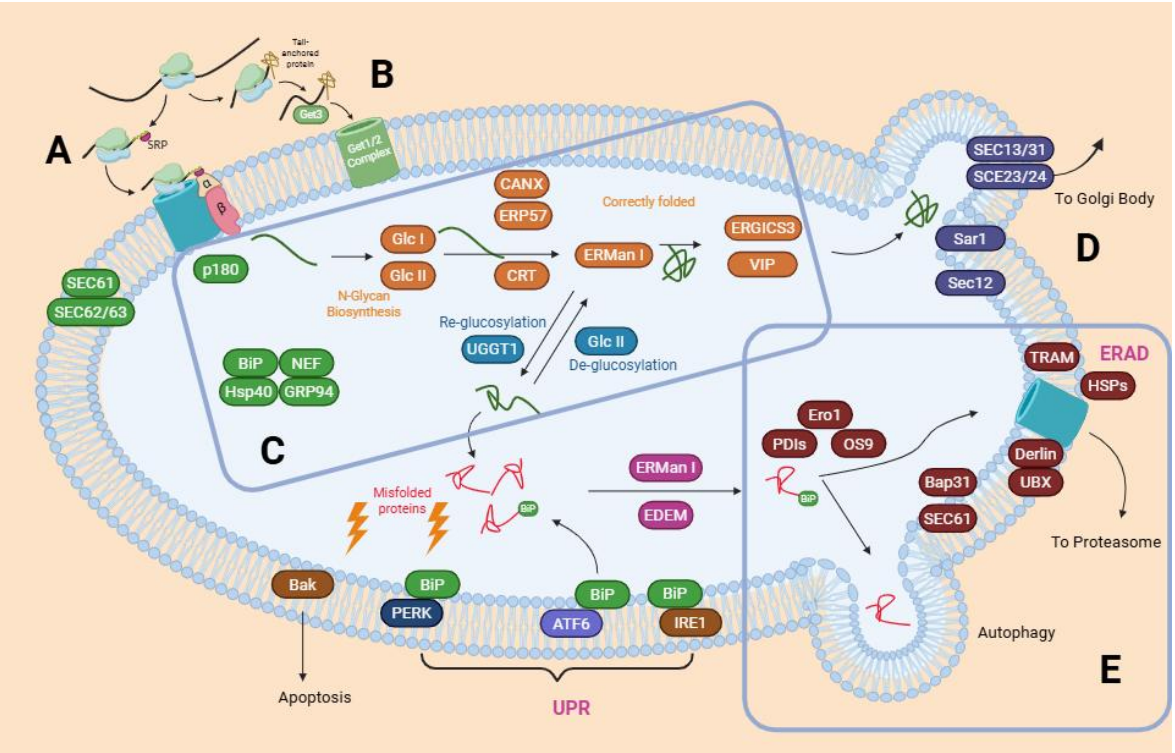


Figure 4: Schematic of major steps in ER protein folding. A-Co-translational translation via Signal Particle receptor pathway. B-Posttranslational membrane insertion through tail-anchored protein recognition by the Bat3 complex. C-The CNX/CRT cycle that allows for proper glycosylation and glucosylation. D-Secretion of natively folded proteins from the ER. E- Degradation of terminally misfolded proteins through ERAD or autophagic degradation.

1.3.2 Unfolded Protein Response (UPR)

One obstacle that still exists in the use of CHO cells is the ER stress response which arises following a disruption to ER homeostasis. This disruption can be caused by interior or exterior factors such as high protein demand, viral infections, environmental toxins, inflammatory cytokines, and mutant protein expression (Oslowski and Urano, 2011). The mechanics of how the disrupted function causes misfolding is not yet fully understood. The complex machinery present in the ER can become exhausted if there is over-expression of wild-type proteins, as shown in studies investigating blood coagulation factor VIII (Dorner et al., 1989) or anti-thrombin III (Schröder and Friedl, 1997), showing how the extra

workload of recombinant CHO cells may lead to an increase in misfolded proteins can induce the conditions of ER stress.

Three ER transmembrane sensors act as the main control of the UPR, inositol-requiring enzyme 1 (IRE1), activating transcription factor 6 (ATF6) and protein kinase R-like ER kinase (PERK) (Qu et al., 2021). When ER function is normal, these proteins are bound to the ER chaperone protein BIP, and are kept inactive, but as misfolded proteins accumulate, BIP will dissociate from these proteins in order to bind to unfolded proteins, thus allowing the activation of the UPR (Gardner and Walter, 2011), as shown in **Figure 5**. The activation of the UPR causes an expansion of the ER membrane to accommodate the misfolded proteins, and then begins to populate the ER lumen with chaperones and protein-folding machinery (Walter and Ron, 2011). The C-terminus of ATF6 extends into the ER lumen, while the N-terminus faces the cytosol, which includes the leucine zipper motif DNA binding region and transactivation domain (Hong et al., 2004). Following the release of BIP, two Golgi localisation sequences (GLS1 and GLS2) become exposed in the ER domain, and ATF6 is translocated to the Golgi body where two enzymes, S1p and S2p cut the lumen domain and transmembrane anchor, allowing the mobilisation of a 50kDa amino terminal cytoplasmic fragment (ATF6f). This transcription factor enters the nucleus, binds to the DNA and activates the transcription of ER stress response elements (Mao et al., 2006; Shen et al., 2005). The transcriptional response induced by ATF6 aims to increase the folding capacity of the ER, with targets including vital ER-resident proteins involved in folding, such as BIP, various protein disulfide isomerases, and the heat shock protein GRP94 (Walter and Ron, 2011).

IRE1 is a transmembrane protein, which senses ER stress through its ER lumen domain and carries out serine/threonine kinase and endoribonuclease activities through its cytosolic domain (Hetz et al., 2011). The dissociation of BIP from the luminal domain allows IRE1 oligomerisation and autophosphorylation, causing the activation of the RNase domain

(Korennykh et al., 2009). When activated, IRE1 cleaves the mRNA that encodes a UPR-specific transcription factor, XBP1 (X-box binding protein 1) at two specific positions, excising an intron (Gardner and Walter, 2011). Removal of this intron from XBP1 leads to the expression of the active form of the protein, which positively regulates ER chaperones, as well as genes that code for proteins in the ERAD pathway (Yoshida et al., 2001). The endo-ribonuclease activity of IRE1 induces activity of Regulated IRE1-dependent mRNA Decay (RIDD), which reduces the folding load of the ER by degrading specific ER-localised mRNAs (Moore and Hollien, 2015). IRE1 also modulates autophagy and apoptosis through an XBP1-independent mechanism, binding to tumour necrosis factor (TNF) receptor-associated factor 2 (TRAF2) and activating JUN amino-terminal kinase (JNK) (Hetz and Glimcher, 2009). These two aspects of IRE1 signalling can influence whether the cell will adapt or undergo apoptosis depending on the severity and duration of ER stress.

Following disassociation of BIP, PERK can dimerise and auto-phosphorylate, activating its kinase function (Harding et al., 2000). The activated PERK is then able to phosphorylate eIF2α (eukaryotic initiation factor 2 alpha), inhibiting global translation initiation, thus reducing the number of nascent polypeptides needing folding (Castellano et al., 2023). This phosphorylation of eIF2α causes it to sequester eIF2β, which normally mediates the GDP-GTP exchange on eIF2α. This exchange is what forms the eIF2α ternary complex which initiates translation (Boye and Grallert, 2020). During this reduction in global translation, specific mRNAs, including ATF4 are preferentially translated. ATF4 activates a panel of genes related to amino acid metabolism, autophagy and ER stress response (Neill and Masson, 2023). One of these proteins is CHOP, which promotes apoptosis under prolonged ER stress, leading to cell death (Li et al., 2014). ATF4 will also induce the expression of GADD34, which recruits phosphatase PP1 to dephosphorylate eIF2α, which will allow translation if ER homeostasis is restored (Novoa et al., 2001).

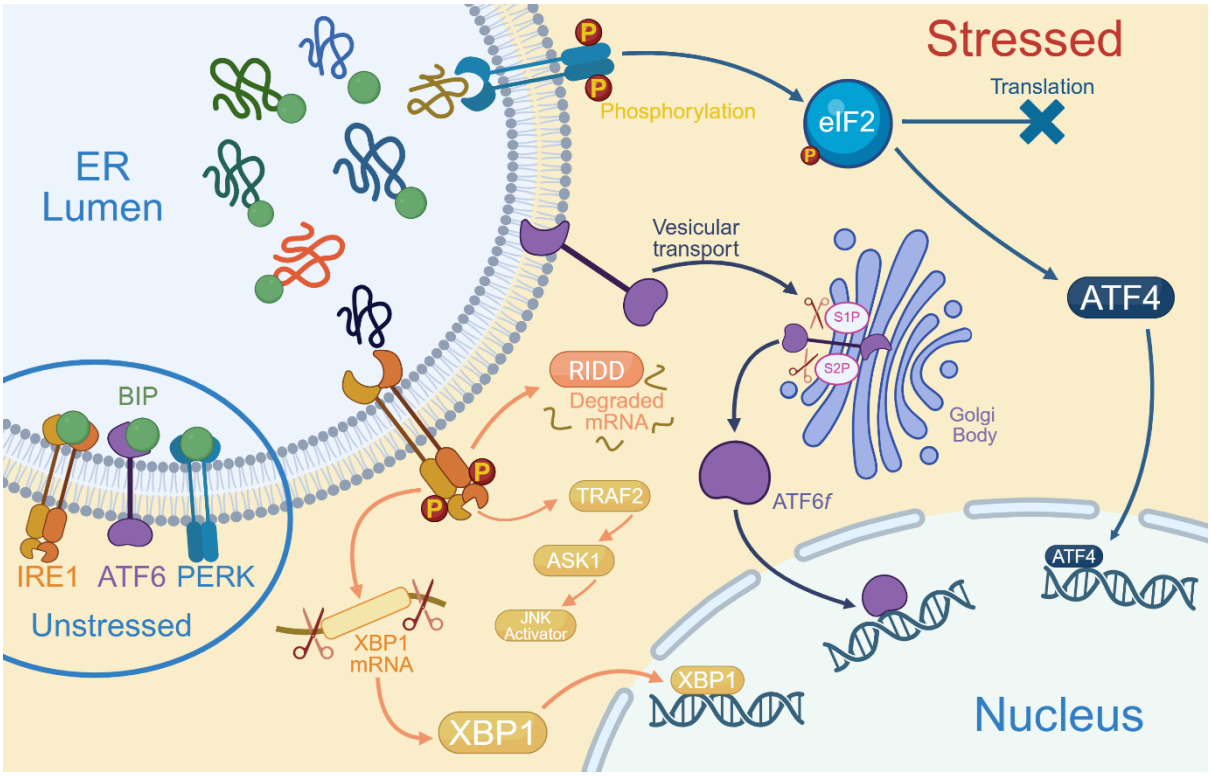


Figure 5: The UPR signalling pathway. This pathway is induced by the separation of BIP from the ER domains of IRE1, ATF6 and PERK as BIP binds to misfolded proteins. The release of BIP allows the N-terminal fragment of ATF6 to be separated so it can travel to the nucleus to activate the transcription of ER stress response proteins. Along with the release of BIP, PERK and IRE1 are activated by phosphorylation, leading to the activation of ATF4 and XBP1, respectively. *Figure created in BioRender.com.*

These signalling pathways culminate in transcriptional and translational regulation, and all three work together to stimulate many adaptive responses that aim to restore ER function and maintain cell survival by regulating genes encoding ER chaperones, ERAD factors, amino acid transport and metabolism proteins and phospholipid biosynthesis enzymes (Hetz et al., 2011; Oslowski and Urano, 2011). If normal ER function is restored, feedback mechanisms switch off the UPR and through the downregulation of stress-induced transcription factors including the dephosphorylation of eIF2α. However, if stress persists, the UPR switches to a pro-apoptotic phase: sustained PERK-eIF2α-ATF4 signalling induces CHOP, which represses anti-apoptotic BCL-2, upregulates pro-apoptotic proteins including BIM, DR5, PUMA, and activates mitochondrial apoptotic pathways (Wang and Kaufman, 2012). Prolonged IRE1α activation promotes RIDD degradation of ER-associated mRNAs, activates TRAF2-ASK1-JNK signalling, and contributes further to apoptosis (Hetz et al.,

2020). The release of calcium from the ER and subsequent mitochondrial Ca^{2+} overload triggers cytochrome c release and activation of the APAF-1/caspase-9 protein complex which amplifies caspase-3/7 activation and cell death (Malhotra and Kaufman, 2011). This transition serves as a quality control measure to prevent damaged or dysfunctional cells from compromising the surrounding tissue. A similar release of calcium is seen in the cold shock cell response. A drop in temperature alters membrane structure and effects the cold-sensitive ion channels, including the inositol 1,4,5-triphospahte receptor, and allows the release of calcium into the cytosol (Wang and Siemens, 2015)

1.3.3 ER-associated degradation (ERAD)

The ERAD pathway manages the misfolded proteins when quality control systems recognise a protein as incorrectly folded and designates the protein for degradation, as shown in **Figure 6**. Proteins that cannot achieve, or lose, their correct folding status (quality control), as well as unnecessary copies of specific proteins (quantity control) are directed for removal from the ER (Hegde and Ploegh, 2010). By binding to exposed mannose residues ER Degradation Enhancing α-Mannosidase-Like proteins (EDEMs) recognise misfolded glycoproteins, identifying these proteins for retro-translocation and degradation (Wang and Hebert, 2003). Other chaperones and lectins, including YOS9p and GRP94 perform a similar role in the detection and trafficking of misfolded substrates to membrane-embedded ERAD adaptors (Bhamidipati et al., 2005; Christianson et al., 2008). The dislocation of ERAD candidates across the ER may be driven by the *cis* to *trans* isomerisation of peptidyl-prolyl bonds, by promoting their detachment from luminal chaperones (Bernasconi et al., 2010) or directing them through a membrane-embedded channel by removing turns in the misfolded polypeptide structure (Ellgaard and Ruddock, 2005). The mechanisms of retro-translocation of misfolded proteins via the ERAD is not yet fully understood, but is likely mediated by the HRD1 complex (composed of HRD1, HRD3, DER1, USA1, and YOS9) (Wu et al., 2020), along with several other key proteins and protein families. Derlin proteins, including

Derlin-1, -2, and -3, assist in extracting misfolded proteins from the ER membrane and delivering them to the cytoplasmic proteasome for degradation, and are also involved in quality control (Kadowaki et al., 2015). The energy needed for this retro-translocation is provided by ATPases, such as p97/VCP (Ahlstedt et al., 2022). Ubiquitin-regulatory X (UBX) proteins assist in directing substrates from the ERAD pathway towards autophagy (Kilgas and Ramadan, 2023), an alternative degradation route that involves the lysosomes.

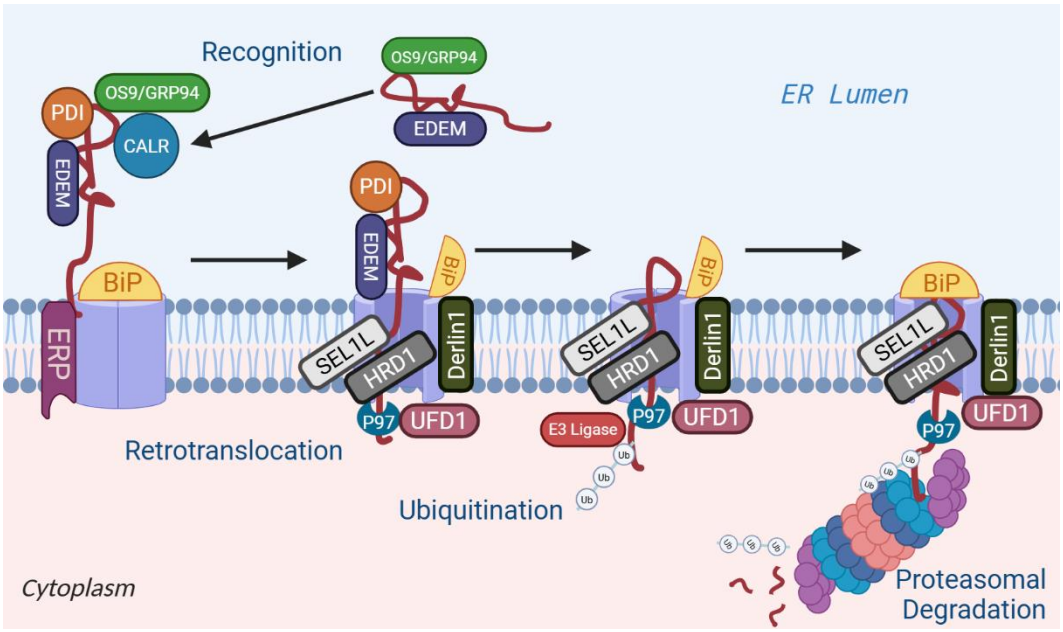


Figure 6: ER-associated degradation (ERAD) pathway. Misfolded proteins in the ER are first recognised by resident chaperones and directed to the retro-translocation machinery at the ER membrane (Recognition). They are then moved across the membrane via the SEL1L-HRD1 complex (Retrotranslocation). In the cytoplasm, the substrate is ubiquitinated (Ubiquitination). Finally, p97 and co-factors extract the substrate and transfer it to the proteasome for degradation (Proteasomal Degradation). The ubiquitin proteins and polypeptides are then recycled by the ER. *Figure created in BioRender.com.*

1.3.4 The relationship between ER stress and CHO cell culture productivity

Several studies have been conducted looking into how the manipulation of ER stress responses can affect CHO cell productivity. These studies can use cell stressors (**Table 1**) or use genetic engineering to alter expression of ER proteins (**Table 2**). The studies indicate that the artificial induction of ER stress leads to a decrease in cell productivity. Exposure to ER stress inducers, thapsigargin and tunicamycin repeatedly limit productivity in CHO cells. These studies directly link the correct function of the ER as vital in the production of recombinant proteins. It has also been found that recombinant protein load can also negatively affect ER function (Prashad and Mehra, 2015). Correct media conditions are also important to correct ER function. High cysteine concentrations have been found to change the redox imbalance in ER and activate the UPR, markedly reducing cell growth (Komuczki et al., 2022). The ER stress placed on CHO cells by mild hypothermia, however, leads to an increase in recombinant protein production. It has been suggested that environmental stressors (e.g. temperature, CO_2, nutrients) cause an increase in cell size, allowing an increase in ER capacity (Kim et al., 2001; Palomares et al., 2023; Schellenberg et al., 2022a; Torres et al., 2021). This increase in cell size has been linked to changes in ion gradients and osmolytes accumulation which draws in more water, cell cycle arrest causing a build-up of cellular mass without division, and the swelling to allow greater space for increased ER capacity (Ingber, 2003; Pozarowski and Darzynkiewicz, 2004). Beneficial adaptation of CHO cells to stress can be achieved without genetic engineering. Chandrawanshi et al. achieved a 1.5-2-fold increase in final titres after adapting CHO cells to ER stress through constant low-level exposure to ER stress inducer tunicamycin (Chandrawanshi et al., 2020). These stress-adapted cells developed an increased ER capacity through the upregulation of chaperones and transcription factors, including BiP and CRT.

Table 1. Literature relating to ER activity in CHO cells and the effects of the various ER stressors on the productivity of CHO cells and the main ER proteins impacted by these stressors.

Stressor	Main ER-related proteins reported	Effect on CHO productivity	Study
Mild hypothermia	Not investigated	Mild hypothermia increased TIMP production 4-fold.	(Jenkins and Hovey, 1993)
Mild hypothermia	Not investigated	Mild hypothermia increased hEPO production 1.85-fold.	(Trummer et al., 2006)
Tunicamycin / recombinant protein load	BIP, CHOP, XBP1 splicing, PERK/IRE1	Tunicamycin exposure and high recombinant load associated with reduced specific productivity.	(Prashad and Mehra, 2015)
Thalidomide - Ubiquitin ligase inhibitor	UBR4 and UBR5	More efficient and stable secretion of recombinant HSA-HG	(Xu et al., 2017)
Tunicamycin/ Thapsigargin	BIP/HSPA5, XBP1s, CHOP	Tunicamycin reduced product titre. Identification of clones with lower ER stress and improved recombinant protein output	(Roy et al., 2017)
BIX (BiP inducer X)	BIP, CHOP	Addition of BIX increased mAb titre and specific productivity.	(Ha et al., 2019)
Mild hypothermia	XBP1s, DDIT3, ATF5 BIP, GRP94, TRIB3	Mild hypothermia increased specific productivity in studied CHO lines.	(Torres et al., 2021)
Thapsigargin/Tunicamycin	GRP94, PDIs, CALR, BIP	Reduction in viable cell number and percentage viability.	(Samy et al., 2021a)
Clone/media combinations	BIP, XBP1, ATF4, CHOP	Clones with lower Hspa5/UPR activation showed higher mAb productivity.	(Tanemura et al., 2022)
CO2 and macronutrient supply	Not investigated	Cell stress increased cell-specific productivity 94.9%-134.4%	(Schellenberg et al., 2022b)
Tunicamycin	BIP, PDIs, ERAD factors, CHOP	Reduction in productivity upon tunicamycin exposure.	(Sulaj et al., 2024)

The genetic manipulation of ER proteins can be incredibly effective at increasing CHO productivity. In 2017, Pieper and colleagues showed that both transient and stable depletion of ATF6β, a repressor of the pro-survival and UPR promotion protein ATF6α, in IgG producing CHO cells allowed for greater specific cell productivity (Pieper et al., 2017). The authors suggest that further ATF6 based cell engineering could create an optimised UPR environment in CHO cells that would allow for increased productivity. Knocking down of ERAD-related E3 ligases was found to increase antibody productivity in CHO cells, possibly by redirecting antibody molecules from degradation to secretion (Tang et al., 2020). The use of a ubiquitin-ligase inhibitor, thalidomide, has been shown to improve the expression of

biologically active fusion protein HSA-HGF (R494E) in CHO cells (Xu et al., 2017). The upregulation of another UPR signalling protein, PERK, has also been shown to reduce cell productivity (Castellano et al., 2023). The findings suggest that attenuation of IRE1a and PDGFRa signalling by PERK reduces ER folding capacity, which reduces specific productivity, in order to mitigate the UPR and prevent apoptosis (Castellano et al., 2023). The increased expression of BIP has been shown to increase the viability of CHO cell cultures when they are exposed to oxidative stress (Jiang et al., 2016). While this was studied in non-producer cells, increasing the stress resilience of CHO cells could allow for increased production or the production of more complex, ER-intense products. Increasing ER chaperone expression and UPR activity helps the cell cope with the folding load imposed by complex proteins (such as bi-specifics or ADCs), reducing misfolding, improving yield, and maintaining product quality. Overexpression of QSOX1b, a catalyst for disulfide bond formation in the ER, in CHO cells extended cell viability, improved folding capacity, and increased antibody productivity by ~45-52 % (Zhang et al., 2024). It has also been shown that in high-producing CHO cell lines, chaperones like CRT, CNX, ERO1α, ERDJ4 and BIP are upregulated over the culture period, correlating with higher secretion rates and more efficient assembly of antibodies (Prashad and Mehra, 2015). These studies show that there is a range of CHO ER stress responses, varying by cell line and product types, which suggests that CHO adaptation may be best utilised on a case-by-case cell line basis. By furthering our understanding of the expression of ER proteins in different environmental conditions, we can more carefully select targets for genetic engineering. Increased knowledge of these mechanisms will allow us to adapt host cell lines to the increased stress put on the ER by complex modalities. These proteins do not operate in isolation, so we must complete expand our understanding of ER pathways to include post-transitional modifications and identify specific proteins that are best suited to specific culture conditions. By understanding the roles of ER proteins, we can select targets for genetic engineering to adapt host cell lines to

the increased stress put on the ER by complex modalities. Studies such as these indicate that proper ER stress mechanisms are crucial to the production of recombinant proteins. This presents significant opportunities for advancing the development of CHO cell lines.

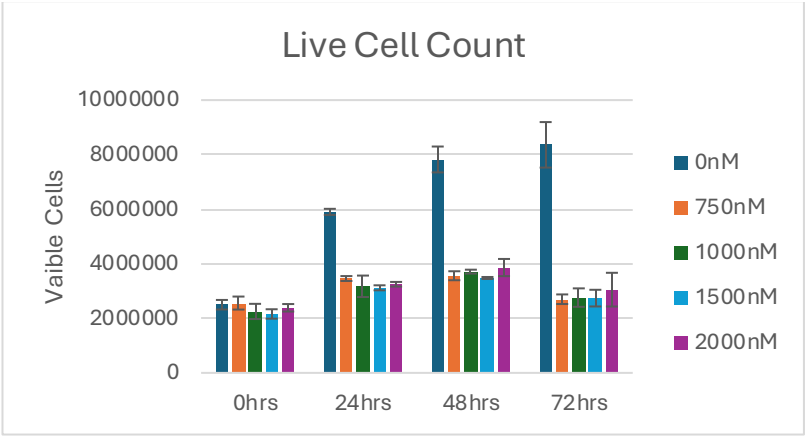
To gain insight into CHO cell ER stress, research was carried out into the effect of exposure to the artificial stress inducer, thapsigargin, explored in Chapter 2 of this thesis (Ryan et al., 2023). In this study, we aimed to get a comprehensive understanding of the ER stress response in CHO cells, so a range of thapsigargin concentrations were tested to find the highest concentration that would not lead to a reduction in culture viability. Thapsigargin disrupts the function of calcium-dependent chaperones, via the non-competitive inhibition of sarco-endoplasmic reticulum Ca^{2+} ATPase (SERCA), leading to an increase in misfolded proteins (Patkar et al., 1979; Xu and Van Remmen, 2021). Thapsigargin binds to a hydrophobic pocket within the transmembrane domain of SERCA, stabilising the protein in a conformation that prevents ATP hydrolysis, blocking the active transport of Ca^{2+} from the cytosol to the ER lumen (Sehgal et al., 2017). **Figure 7** shows the concentrations tested, where 0 hours is thapsigargin exposure, 48 hours after seeding. The concentration range that was explored was based on existing literature, summarised in **Table 3**. From this data, the highest concentration, 2000 nM was chosen for the proteomic study due to it having a similar effect on cell viability as lower concentrations. Studies, such as those in this thesis, that expand our knowledge of ER stress in CHO cells could improve next-generation modality production in CHO cells.

Table 2. Literature relating to ER activity in CHO cells through the genetic engineering of ER-related pathways and the related effects on productivity.

Functional work	Effect on CHO productivity	Study
Over-expression of ATF4	Increased production of recombinant human antithrombin III	(Ohya et al., 2008)
Over-expression of XBP1 and ERO1	Increased MAb yields of between 5.3-6.2-fold	(Cain et al., 2013)
Over-expression of XBP1 and CERT S132A	Increased t-PA production	(Rahimpour et al., 2013)
Over-expression of BIP	Increased cell viability and reduced apoptosis when exposed to stress	(Jiang et al., 2016)
Knock down of UBR4/UBR5	Increased mAb productivity	(Tang et al., 2020)
Knock down of PERK branch of the UPR along with BAX and BAK.	Increased mAb production	(Castellano et al., 2023)
Over-expression of MYC and XBP1s	Increased viable cell density and EPO titre	(Latorre et al., 2023)

Table 3. Literature relating to thapsigargin concentrations used to study ER activity in a range of mammalian cell types. These studies were used to inform the concentrations tested in Figure 6.

Cell type	Thapsigargin concentration	Study
NK1-transfected CHO cells	1000 nM	(Seabrook and Fong, 1993)
Neurotensin receptor-expressing CHO cells	1000 nM	(Gailly, 1998)
Mouse fibroblast	300 nM	(Li et al., 2000)
FLAG-CD36 expressing CHO cells	5000 nM	(Kuda et al., 2011)
Mouse macrophage	2000 nM	(Devarajan et al., 2012)
PDI-overexpressing CHO cells	1000 nM	(Samy et al., 2021b)



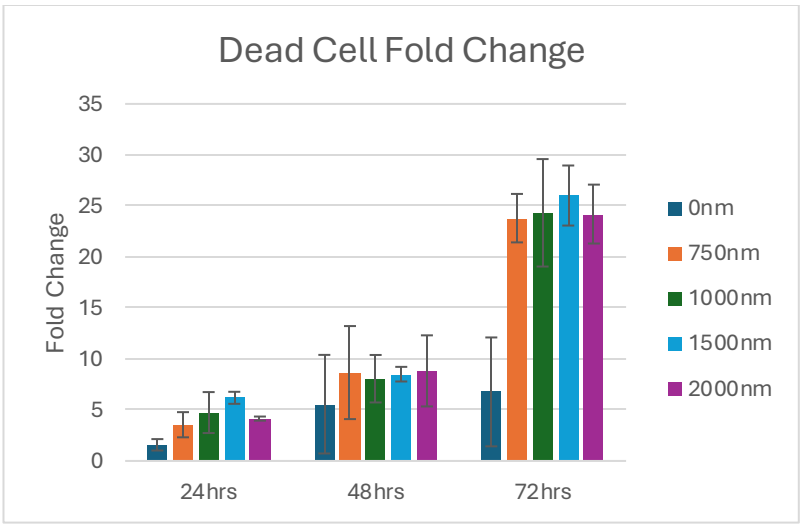


Figure 7: The effects of thapsigargin exposure on CHO culture viability. Viable cell count and fold change in dead cell count in CHO-K1 cells exposed to different concentrations of thapsigargin where 0hrs is thapsigargin exposure, 48 hours after cell seeding. The data lines are the mean value of triplicate results, with error bars representing the standard deviation from the mean value.

1.4 Ubiquitin-a crucial regulator of cellular homeostasis

1.4.1 Biological structure of ubiquitin

The ubiquitination of proteins is a fundamental part of cellular regulation. Over 1000 genes have been identified in the regulation of protein ubiquitination (Clague et al., 2015), indicating how complex and vital this PTM is. In-depth proteomics studies have found that thousands of proteins are ubiquitinated, with most proteins likely ubiquitinated at some stage (Swatek and Komander, 2016), including ubiquitin itself (Komander and Rape, 2012). Ubiquitin is a small protein, only 76 amino acids long with a globular and tightly folded structure. The linkage ability of ubiquitin is dictated by the eight functional sites present, seven Lys (K6, K11, K27, K29, K33, K48, K63) and the N-terminus methionine (Met1), with the different linkage types aligning with the different function played by ubiquitin in that bond (Komander and Rape, 2012), as shown in **Figure 8**. The most biologically abundant of these linkage types is the 48-linkage, which mainly functions to target substrates for 26S proteasome-mediated degradation (Jacobson et al., 2009).

The attachment of ubiquitin to proteins is directed by a three-step cascade, as shown in **Figure 9**. The process is started by E1 activating enzymes, which activates the ubiquitin protein through carboxy-terminal adenylation and thiol transfer, and coordinate the use of ubiquitin in specific downstream pathways by charging E2 conjugating enzymes (Schulman and Harper, 2009). Active E2 enzymes possess a core ubiquitin-conjugating (UBC) domain, which contains the catalytic cysteine residue and interacts with the E1 enzymes (Ye and Rape, 2009). UBC domains from different E2s show a high degree of sequence homology, with similar structures and the active-site Cys is also highly conserved (Eddins et al., 2006), indicating the vital role of ubiquitination in the cellular functions. When E1 becomes charged with ubiquitin a conformational change in E1 exposes E2-binding sites and allows the formation of the E1-E2-ubiquitin complex (Jin et al., 2007). This conjugate is often referred to as E2~Ub. The final step in this cascade is the E3 ligases, which transfer the activated

ubiquitin to the target protein (Yang et al., 2021). E3 ligases are able to selectively attach ubiquitin to the lysine, serine, threonine or cysteine residues of specific substrates (Zheng and Shabek, 2017). This attachment is called ubiquitylation. E3 ubiquitin ligases are split into three types: RING-type (really interesting new gene), RBR-type (RING-in-between-RING) and HECT-type (homologous to the E6AP carboxyl terminus) (Deshaies and Joazeiro, 2009; Dove et al., 2016; Rotin and Kumar, 2009). All types of eukaryotic E3 ligases bind to an E2~Ub thioester and then, in RING-types, catalyse the transfer of ubiquitin from the E2 to a substrate lysine via an amino-lysis reaction, or for HECT and RBR-types, to another cysteine in the E3 via a trans-thio-esterification reaction (Berndsen and Wolberger, 2014).

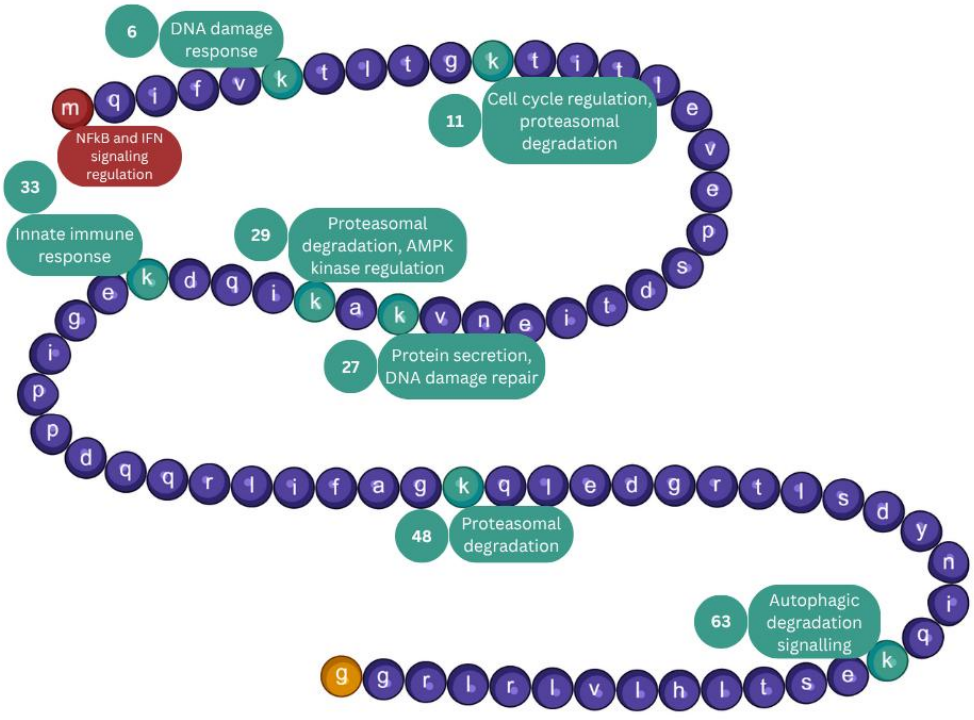


Figure 8: The amino acid structure of ubiquitin. The seven lysine sites (K6, K11, K27, K29, K33, K48, K63) and N-terminal methionine1 (Met1) site which characterise the structure of ubiquitin, as well as acting as functional sites. The different linkage of these sites serves different functions in the role of the modification. *Figure created in BioRender.com*

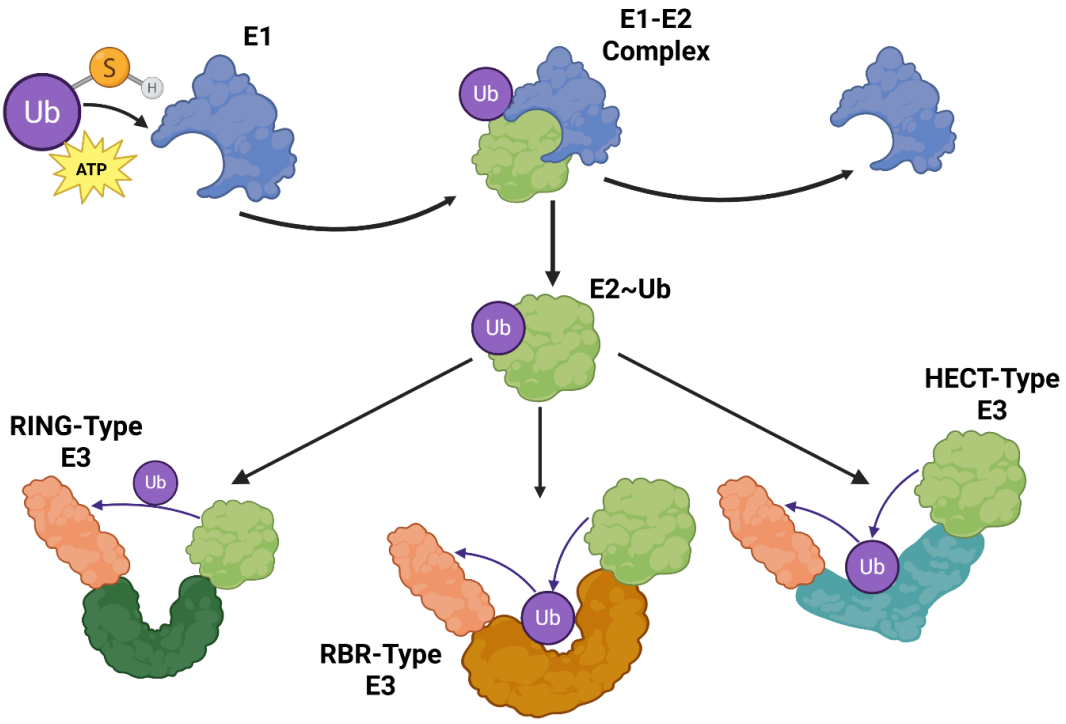


Figure 9. The attachment of ubiquitin to a target substrate. The attachment of a ubiquitin protein to a substate protein involves an enzymatic cascade, using the ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2) and ubiquitin ligases (E3). Firstly, ubiquitin is activated by ATP, and thiol transfer attaches the ubiquitin to E1. E1 then binds to E2, where it passes over the ubiquitin to an active site on the E2. It is the E3 ligases that then co-ordinate the final transfer of ubiquitin onto the target protein. RING-type ligases transfers ubiquitin directly from the E2~Ub conjugate to the substrate protein. In contrast, RBR- and HECT-type ligases transfer ubiquitin to onto an active site cysteine on itself, creating a thioester intermediate before transferring ubiquitin to the target protein. *Figure created in BioRender.com*

RING-type E3 ligases contain a RING domain, which is a monomeric or dimeric conserved cross-brace structure with two structural Zn^{2+} ions (Bentley et al., 2011; Liew et al., 2010), which can bind to the N-terminal of the helix of E2 enzymes (Zheng et al., 2000). This domain interacts with a hydrophobic patch on ubiquitin, which immobilises the donor ubiquitin and positions the C-terminus in a structurally conserved groove in the E2 (Plechanovová et al., 2012). The E3 then binds to the target protein, with most E3 enzymes being able to recognise specific signals in specific target proteins (Jackson et al., 2000). This conformation aligns the ubiquitin in such a way that the relevant thioester between E2 and the ubiquitin C-terminus will be in the correct orientation to react with the substrate lysine

with the E3 acting as a scaffold for ubiquitin attachment via nucleophilic attack (Plechanovová et al., 2012).

HECT E3 ligases contain a HECT domain in their C-terminal region, which consists of an N-lobe that interacts with the E2 enzyme and a C-lobe which contains a catalytic cysteine residue which forms a transient thioester bond with ubiquitin (Bernassola et al., 2019). The target protein is then bound by the E3, bringing the E3~Ub in proximity with the relevant Lys of the target protein. The ε-amino group of the Lys residue then attaches via a nucleophilic attack on the thioester bond of the E3~Ub and an isopeptide bond forms between the C-terminal glycine (G76) of ubiquitin and the Lys residue of the target protein (**Figure 10**). Unlike RING-type ligases, HECT ligases form a covalent intermediate (E3~Ub), providing direct catalytic involvement in ubiquitination (Kamadurai et al., 2013).

RBR-type ligases include features from both RING and HECT type ligases. These ligases contain two RING domains, one which binds to the E2~Ub (RING1) and the other which contains the cysteine residue that forms a thioester intermediate with ubiquitin (RING2), as well as an IBR domain (In-Between-RING) which structurally binds the two RING domains together (Spratt et al., 2014). The E3 ligase then binds to the target protein and the positions it close to the RING2 domain, and then the ubiquitin attaches to the protein through the same mechanism as HECT ligases (Wenzel et al., 2011).

Ubiquitination can occur in several conformations based on the way the modification interacts with the target substrate: mono-ubiquitination, multi-ubiquitination, and polyubiquitination. Mono-ubiquitination involves a single ubiquitin attaching to a protein, multi-ubiquitin is the attachment of ubiquitin to several different residues on a proteins and polyubiquitination is the attachment of several ubiquitin proteins to a single amino acid of a protein, joined in a chain (Senft et al., 2018).

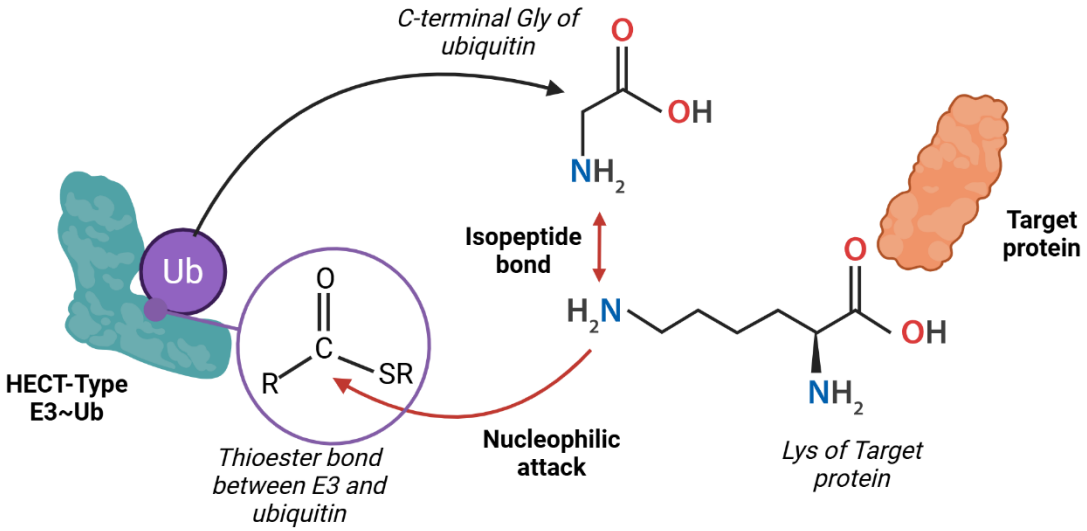


Figure 10: The transfer of ubiquitin via HECT E3 ligases. The ε-amino group of the Lys residue of the target protein attaches, via a nucleophilic attack, to the thioester bond of the E3~Ub and an isopeptide bond forms between the C-terminal glycine (G76) of ubiquitin and the Lys residue of the target protein. *Figure created in BioRender.com*

1.4.2 The role of ubiquitin in cell signalling pathway regulation

Non-proteolytic ubiquitination plays a vital role in multiple signalling cascades in various cellular pathways, including immune regulation. K33, K48 and K63 linkages are associated with the innate immune response, in particular in response to viral infections, particularly polyubiquitin linkages (Tracz and Bialek, 2021; van Huizen and Kikkert, 2020). In the "nuclear factor kappa-light-chain-enhancer of activated B cells" (NF-κB) pathway, K63- and M1-linked ubiquitin chains act as scaffolds to recruit signalling complexes that can activate IκB kinase (IKK), which is crucial to the cellular response to inflammation (Chen, 2012). Pathways dictating the function of tumour necrosis factor receptor (TNFR) and interleukin-1 receptor (IL-1R) rely on E3 ligases and the assembly of K63- and M1-linked polyubiquitin chains to propagate downstream signals (Walczak et al., 2012). K63-linked ubiquitin modifications regulate pattern recognition receptor (PRR) signalling through Toll-like receptors (TLRs) and RIG-I-like receptors (RLRs), orchestrating antiviral responses (Zeng

et al., 2009). K63-linked ubiquitin chains are involved in the assembly of protein complexes in response to extracellular stimuli (Chen and Sun, 2009). Linear (M1-linked) ubiquitination is vital to modulating inflammatory signalling and cell survival (Tokunaga et al., 2009). In these ways, ubiquitin serves as an essential recruitment platform for kinase complexes, enabling signal amplification and specificity.

Ubiquitination plays a critical role in DNA damage response, where it modulates the recruitment and assembly of DNA repair machinery. In response to DNA double-strand breaks, E3 ligases RNF8 and RNF168 catalyse the formation of K63-linked ubiquitin chains on histone H2A and associated proteins, facilitating the recruitment of repair factors including BRCA1 (Doil et al., 2009; Stewart et al., 2009). These modifications alter chromatin structure and promote signal amplification, coordinating efficient DNA repair. The monoubiquitination of histones, particularly H2A and H2B, regulates transcription and chromatin compaction, as well as promoting the methylation of histones, which are epigenetic markers associated with active transcription (Zhou et al., 2008).

Ubiquitin also regulates intracellular trafficking, including endocytosis and lysosomal sorting, through either monoubiquitination or K63-linked polyubiquitin chains. Plasma membrane receptors, such as the epidermal growth factor receptor (EGFR), are monoubiquitinated following ligand binding, which signals their internalisation (Haglund et al., 2003).

1.4.3 The role of ubiquitin in proteasomal degradation targeting

In relation to ER function, the most important role of ubiquitin as a target attached to misfolded or unneeded proteins for proteasomal degradation, most associated with K48- or K11- linkages (Jacobson et al., 2009). However, in cells where the proteasome has been inhibited, K-6, K-23 and K-29 linkages were found to accumulate to varying levels and K-63 ubiquitin chain levels were unaffected (Kim et al., 2011), suggesting that most linkages

play a role in proteasomal degradation, except K-63. In order to be detected by the 26S proteasome, target proteins must be linked to a polyubiquitin chain if at least four ubiquitin molecules, with increased length beyond four molecules adding a moderate increase in proteasome affinity (Lam et al., 2002). A substrate requires a polyubiquitin chain of at least four ubiquitin molecules for proper binding to the 26S proteasome with increased length beyond four molecules adding a moderate increase in proteasome affinity (Lam et al., 2002). This suggests that the level of ubiquitination on a substrate may dictate the speed and efficiency of the degradation (Lederkremer and Glickman, 2005). Efficient degradation of substrates often involves not only ubiquitination, but also an unstructured initiation region on the substrate that is engaged by the proteasome's ATPases for unfolding (Tomita and Matouschek, 2019)

Proteins tagged with ubiquitin are transported for degradation in the 26S proteasome, a large 2.5 MDa, multi-protein, ATP-dependent protease complex with two subunits, the core particle (CP, 20S) and the regulatory particle (RP, 19S) (Bedford et al., 2010). The RP has both base and lid subassemblies that recognise ubiquitinated substrates and prepare them for degradation in the CP (Lander et al., 2012). The RP base contains six AAA ATPases (RPT1-RPT6), which form a ring and function together unfold the substrate and translocate it to the 20S subunit (Smith et al., 2007). Two non-ATPase proteins, RPN10 and RPN13, are also present in the RP subunit which function as receptors for ubiquitinated substrates through ubiquitin-binding domains (Finley, 2009). It is these ubiquitin receptors that recognise substrates marked with K48- or K11-linked polyubiquitin chains (Saha et al., 2020). Three de-ubiquitinases, RPN11, USP14 and UCH37, then release ubiquitin from the substrate and disassemble it, allowing it to be used for ubiquitination of other proteins (Peth et al., 2009). The 20S particle includes three N-terminal threonine-containing β-type subunits in the interior space (Finley, 2009). The subunits perform caspase-like, trypsin-like, and chymotrypsin-like activities, digesting the substrate into peptides of 2-24 amino acids,

destroying the substrate, and providing amino acids for further protein synthesis (Dahlmann, 2016). β1 activity cleaves acidic residues supported by a positively charged Arg45 at the base of its S1 pocket (Harshbarger et al., 2015). β2 has a larger S1 pocket which is negatively charged, so it cleaves after basic residues. β5 cleaves after bulky hydrophobic residues, due to the large hydrophobic S1 pocket shaped by residues including Met45, Ala20 and Cys52 (Fernandes et al., 2024). The Thr1 oxy-nucleophile in each β-subunit stabilises transition states and assists proton transfer during proteolytic cleavage (Fernandes et al., 2024). Degradation continues progressively, as one peptide bond is cleaved, the substrate fragment moves to an adjacent site for further cleavage, with the final oligopeptides that exiting via the axial pore (Tanaka, 2009).

It is predominantly this function of ubiquitin that is of interest in CHO cell adaptation. The comprehensive function of ubiquitin in mitigating ER stress has yet to be fully understood. The ubiquitinated proteome of CHO cells cultured under mild hypothermia was studied in detail in **Chapter 3** of this thesis. The aim of this study was to increase our understanding of the role ubiquitin plays in the ER stress response to mild hypothermia, which has not previously been studied using proteomic methods.

1.5 Proteomic methods for uncovering differential protein expression

1.5.1 Bottom-up proteomics

To prepare biological samples for proteomic mass spectrometry analysis, several steps are necessary, with different options available for most steps (Wojtkiewicz et al., 2021). "Bottom-up proteomics" or "Shotgun proteomics" was used for all proteomic work carried out in this thesis. Cells are lysed open to release proteins, which are then reduced and alkylated to break open disulphide bonds and "cap" cysteine residues to prevent these bonds re-forming. The proteins are then digested into peptides using various enzymes, with trypsin being one of the most common choices. Trypsin cleaves at lysine and arginine, except in the presence of proline, and tends to produce peptides 7-35 amino acids long (Krokhin et al., 2004). The sample is then "cleaned", using C18 columns or filters, where peptides are bound, washed, and then eluted, normally with acetonitrile and formic acid-based solutions. This eluant is then evaporated, and the dried peptides resuspended before loading into a chromatography system attached to the mass spectrometer (Wojtkiewicz et al., 2021). This method results in short peptides which can be separated by chromatography and fragmented by the mass spectrometer to generate amino acid information which can then be aligned with a protein database for identification. This approach requires no prior knowledge of what peptides may be present in a sample, allowing for either data-dependant or independent methods.

1.5.2 The use of mass spectrometry in proteomics

Mass spectrometry plays a vital role in the field of proteomics, revolutionising our understanding of the complexities of proteins and the role they play in biological systems. As technology has advanced, this analytical technique has enabled researchers to identify, quantify, and characterise proteins with ever-increasing precision. Peptides resulting from enzymatic digestion are separated by liquid chromatography systems coupled to mass

spectrometers (LC-MS). During the analysis, eluting peptides from the chromatography are selected to a predefined criteria (e.g. peptide signal rising above the noise in a full scan mass spectrum) and fragmented through conversion of translational kinetic energy to internal vibrational energy, breaking into smaller, charged product ions, producing tandem (MS/MS) mass spectra known as data-dependent acquisition, where the most intense peptide ions from an MS scan is selected and fragmented, based on setting determined by the user (Glish and Vachet, 2003).. As these peptides travel through the mass spectrometer, they are ionised through the application of high voltage to a sample spray, creating charged ions (Awad et al., 2015), and subjected to high-resolution mass analysis. By measuring the mass-to-charge (m/z) ratios of these ions, researchers can determine the precise molecular weight of peptides, allowing for the identification of the corresponding proteins through comparison with extensive protein databases. The development of CHO gene and protein databases, including the sequencing of the CHO genome (Xu et al., 2011), has made it increasingly easy and efficient to identify the large numbers of proteins that are often detected, particularly in whole cell samples. Additionally, mass spectrometry facilitates the study of PTMs (Mnatsakanyan et al., 2018), protein-protein interactions (Wheat et al., 2021), and, using various methods, including label-free quantification via spectral counting or peak intensity, metabolic labelling, chemical tagging, and targeted approaches using isotope-labelled standards, absolute and relative abundance of proteins in various biological samples (Rozanova et al., 2021). The versatility and sensitivity of this method have made mass spectrometry an indispensable tool in understanding the proteome, advancing our understanding of complex biological systems, and facilitating discoveries in fields such as disease biomarker identification (Crutchfield et al., 2016), drug development (Mikami et al., 2012), and personalised medicine (Su et al., 2021). Many types of mass spectrometry techniques exist today, with several regularly used for proteomics research, with differences in ionisation technique and the type of chromatography used to separate peptides before

entry to the mass spectrometer. **Table 4** gives a summary of the differences in ionisation methods and mass analyser type between the most common mass spectrometers.

Table 4: Summary of the ionisation types (ESI and MALDI) and the mass analyser types that differentiate the mass spectrometry methods.

Method	Ionisation/ Mass Analyser Type
Electrospray Ionisation Mass Spectrometry (ESI-MS)	Uses a high-voltage electric field to create charged droplets from a sample solution, which leads to the formation of gas-phase ions.
Matrix-Assisted Laser Desorption/Ionisation Mass Spectrometry (MALDI-MS)	Relies on a laser pulse to ionise analytes that are embedded in a matrix material, producing ions for mass analysis.
Time-of-Flight Mass Spectrometry (TOF-MS)	Measures the time it takes for ions to travel a known distance in an electric field. Heavier ions take longer to reach the detector. Often used in tandem with a MALDI ionisation system.
Quadrupole Mass Spectrometry	Uses electric fields to selectively transmit ions of a specific mass-to-charge ratio.
Ion Trap Mass Spectrometry	Uses a combination of electric and magnetic fields to trap and manipulate ions.
Orbitrap Mass Spectrometry	Ions are trapped in an electrostatic field and oscillation frequencies measured to determine mass-to-charge ratios.
Astral Mass Spectrometry	Accumulates fragment ions, which are then oscillated isochronously in an asymmetric ion mirror.
Inductively Coupled Plasma Mass Spectrometry (ICP-MS)	Ionises samples using a high-temperature plasma source.
Trapped Ion Mobility Mass Spectrometry (timsTOF)	Separates ions by mobility through a trapped ion mobility cell before analyses by time-of-flight MS.

Liquid-Chromatography (LC) is commonly used for the separation of peptide samples before entry to the mass spectrometer. This method works by trapping peptides to an analytical column and then using a gradient of solution to separate the sample based on, most commonly, polarity (Jandera, 2013) or hydrophilic interaction (Buszewski and Noga, 2012). This method can be used in conjunction with many of the listed mass spectrometry methods.

LC provides a crucial sample clean-up and enrichment step, reducing sample complexity and enhancing the resolution and sensitivity of downstream analysis (Wilson et al., 2015). The separation of peptides prior to ionisation also reduces ion suppression effects and improves the dynamic range of detection (Baker et al., 2010). Various LC techniques, such as reversed-phase LC (RP-LC) and hydrophilic interaction liquid chromatography (HILIC), are employed depending on the physicochemical properties of the analytes (Buszewski and Noga, 2012). These approaches enable high-throughput analysis of complex biological mixtures and are widely applied in proteomics. The integration of LC with tandem mass spectrometry (LC-MS/MS) further enhances the ability to identify and quantify peptides with high precision and accuracy, as used in this thesis.

1.5.3 Quantitative Label-free proteomics

Label-free proteomics relies on the direct measurement of peptide ion signals to determine protein abundance across different biological samples. Quantification is achieved by comparing signal intensities, or by counting the number of tandem mass spectra assigned to a specific peptide (peptide-spectral matches). Values are then normalised to account for technical variability and allow relative abundance comparisons between conditions (Zhu et al., 2010). Label-free quantitative proteomics requires high instrument stability and rigorous computational analysis to ensure reproducibility and accuracy, as signal intensity can be influenced by ionisation efficiency, sample preparation, and chromatography performance. Data analysis involves alignment of retention times, feature detection, peptide identification, and statistical modelling to distinguish biologically meaningful changes from noise (Neilson et al., 2011). The published proteomics data included in this thesis is all label-free proteomics, performed on an UltiMate 3000 nanoRSLC system interfaced with an Orbitrap Fusion Tribrid Mass Spectrometer. MS/MS was carried out using an Orbitrap mass analyser for MS1 and an ion trap mass analyser for MS2. An Orbitrap mass analyser traps ions in an electrostatic field around a central spindle-shaped electrode (Zubarev and Makarov, 2013).

Ions oscillate along the axis of this electrode, and the frequencies of oscillation are measured. These frequencies are converted into mass-to-charge (m/z) values. This method provides high-resolution, accurate mass measurements of intact peptide ions before they are selected for fragmentation. A top-speed MS/MS acquisition algorithm is used to determine the number of selected precursor ions to be fragmented (Huang et al., 2021). The resulting fragment ions are then analysed in a linear ion trap mass analyser. Using radiofrequency and direct current fields in a linear electrode setup, ions are isolated, fragmented, and sequentially ejected by m/z ratio (Hopfgartner et al., 2004). In tryptic proteomic studies, the primary types of ions formed are B and Y ions. B-ions form when the charge is retained on the N-terminal of the peptide, whereas Y-ions retain the charge on the C-terminal of the peptide (Bythell et al., 2009).

The mass differences between consecutive peaks of a fragment spectrum correspond to the residue masses of specific amino acids. This allows the amino acid sequence of a peptide to be established (Liu et al., 2007). This sequence is then compared against a protein database to allow for identification. In these studies, results were compared against a *Cricetulus griseus* UniProt database using the SEQUEST HT algorithm in Proteome Discoverer 2.5. For amino acids/peptides with very similar masses, reverse-phase LC can separate isomeric peptides or amino acids based on hydrophobicity, allowing them to enter the mass spectrometer at different times, which aids in distinguishing similar amino acid molecular weights (Winter et al., 2009).

1.5.4 Inclusion of PTMs in proteomic studies

With over 400 different types of protein PTMs identified (Ramazi and Zahiri, 2021), researchers have had to consider how best to include these modifications in proteomics studies. The crucial roles played by PTMs in regulating diverse biological processes make them an important consideration in understanding the proteome (Ali et al., 2019). The

transient nature and structure of many PTMs, however, poses a challenge to study through the usual approaches.

Mass spectrometry has emerged as a powerful tool for studying the ubiquitination of proteins with unparalleled sensitivity and specificity (Hansen et al., 2021; van der Wal et al., 2018; Xu et al., 2014). Ubiquitinated peptides represent only a small fraction of the total peptide population. Therefore, enrichment techniques are employed to selectively isolate ubiquitinated peptides from the mixture. The presence of a remnant ubiquitin double Glycine (diGLY) amino acid residue following tryptic digestion was first noted in 1977 (Goldknopf and Busch, 1977). It is this residue that allows for affinity techniques that isolate ubiquitinated peptides following digestion, with the first attempted study of a ubiquitinated proteome carried out in 2003, characterising ubiquitin conjugates from *Saccharomyces cerevisiae* (Peng et al., 2003). This method relies on the detection of a 114.1 Da increase in the mass of a peptide which is the mass of the di-glycine tag. The methodology used by Peng et al. would provide a general foundation for the large-scale analysis of ubiquitination, which would be adapted and built upon as the field expanded. **Figure 11** details the method of enrichment of peptides containing the ubiquitin di-glycine tag which is formed following tryptic digestion of peptides. This method, adapted and advanced since its original inception, was used for the ubiquitinated proteome study explored in Chapter 3 of this thesis.

Figure 12 demonstrates the distinct signature left by the ubiquitin di-glycine tag in the proteomic data derived from CHO-DP12 cells exposed to mild hypothermia. The same peptide from Gap Junction Protein A1 was detected in both whole cell lysate (**12.1**) and ubiquitin enriched data (**12.2**). This protein has functions linked to protein localisation (Bazzoun et al., 2019), cell-cell signalling and signal transduction (Fishman et al., 1990), and was found to have differential expression of both the ubiquitinated and non-ubiquitinated proteins, discussed in **Chapter 3**. The ubiquitin diglycine (GG) tag is found on K14 of peptide *VAQTDGVNVEMHLKQEIK* in the ubiquitin enriched sample. The

mass difference of these peptides can be seen in the MS spectra from the peptides (**12.3**). Additionally, targeted mass spectrometry approaches enable the quantification of ubiquitination levels in specific proteins of interest based on peptide immunoaffinity enrichment and targeted multiple reaction monitoring mass spectrometry (immuno-MRM) (Boja and Rodriguez, 2012; Mollah et al., 2007). This provides insights into the regulation and functional consequences of ubiquitin signalling pathways.

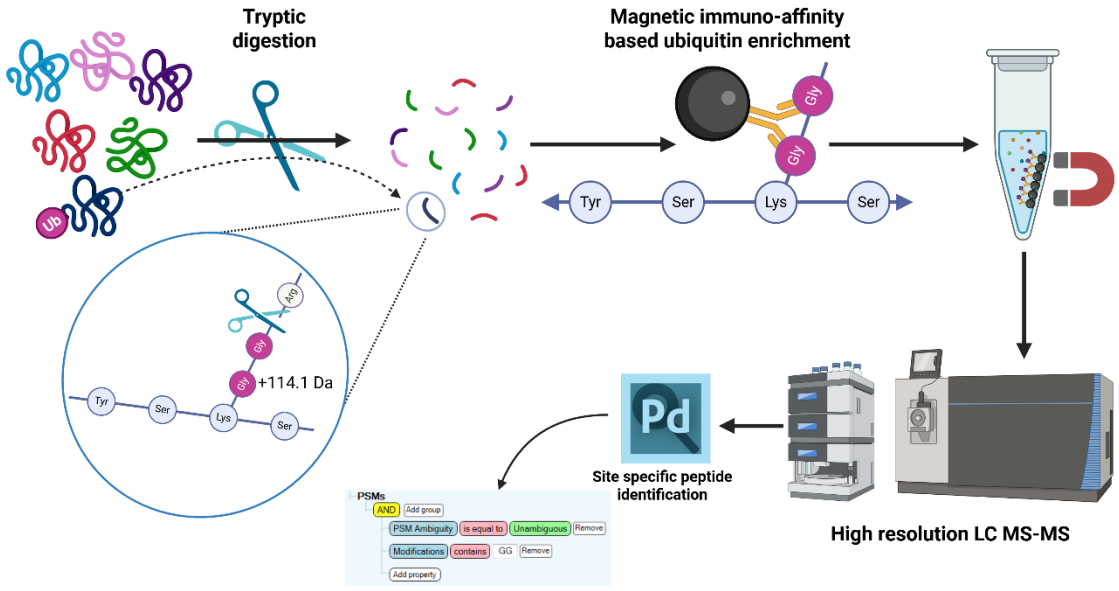


Figure 11. Immuno-affinity enrichment of ubiquitinated peptides. After reduction and alkylation, proteins are digested into peptides using trypsin. This creates a double glycine remanent on lysine residues where a ubiquitin molecule would have been attached. Using magnetic beads conjugated to a diglycine-specific antibody, ubiquitinated peptides can be separated from the non-ubiquitinated peptides. LC-MS/MS then allows for the mass of peptides to be calculated, with the diglycine tag causing a recognisable mass shift of 114.1 Da. Proteome Discoverer is then used to identify these ubiquitinated peptides against a protein database.

	b^{+}	b^{2+}	Seq.	y^{+}	y^{2+}	
1	100.0757	50.54148	V			14
2	171.1128	86.06004	A	1399.663	700.3354	13
3	299.1714	150.0893	Q	1328.626	664.8168	12
4	400.219	200.6132	T	1200.568	600.7875	11
5	515.246	258.1266	D	1099.52	550.2637	10
6	572.2675	286.6374	G	984.4931	492.7502	9
7	643.3046	322.1559	A	927.4717	464.2395	8
8	757.3475	379.1774	N	856.4346	428.7209	7
9	856.4159	428.7116	V	742.3916	371.6995	6
10	971.4429	486.2251	M	643.3323	322.1652	5
11	1102.483	551.7453	D	528.2963	264.6518	4
12	1239.542	620.2748	H	397.2558	199.1315	3
13	1352.626	676.8168	L	260.1969	130.6021	2
14			K	147.1128	74.06004	1

```
1   MGDWSALGKL LDKVQAYSTA GGKVWLSVLF IFRILLLGTA VESAWGDEQS AFRCNTQQPG
                                                          *
61  CENVCYDKSF PISHVRFWVL QIIFVSVPTL LYLAHVFYVM RKEEKLNKKE EELKVAQTDG
              *
121 VNVEMHLKQI EIKKFKYGIE EHGKVKMRGG LLRTYIISIL FKSVFEVAFL LIQWYIYGFS
181 LSAVYTCKRD PCPHQVDCFL SRPTEKTIFI IFMLVVSLVS LALNIIELFY VFFKGVKDRV
241 KGRSDPYHAT TGPLSPSKDC GSPKYAYFNG CSSPTAPLSP MSPPGYKLVT GDRNNSSCRN
301 YNKQASEQNW ANYSAEQNRM GQAGSTISNS HAQPFDFPDD SQNAKKVAAG HELQPLAIVD
361 QRPSSRASSR ASSRPRPDDL EI
```

Figure 12.1: A CHO protein identified using MS without detection of the ubiquitin remanent. GJA1 peptides analysed on the Orbitrap Fusion *VAQTDGVNVEMHLK whole cell lysate of CHO-DP12 cells, cultured at 31°C, 48 hours post temperature downshift. Full amino acid sequence for GJA1, with the peptides detected highlighted in green.

	b^+	b^{2+}	Seq.	y^+	y^{2+}	
1	100.0757	50.54148	V			19
2	171.1128	86.06004	A	2125.071	1063.039	18
3	299.1714	150.0893	Q	2054.034	1027.52	17
4	400.2191	200.6132	T	1925.975	963.4911	16
5	515.246	258.1266	D	1824.927	912.9673	15
6	572.2675	286.6374	G	1709.9	855.4538	14
7	643.3046	322.1559	A	1652.879	826.9431	13
8	757.3475	379.1774	N	1581.842	791.4245	12
9	856.4159	428.7116	V	1467.799	734.403	11
10	971.4429	486.2251	M	1368.73	684.8688	10
11	1102.483	551.7453	D	1253.703	627.3554	9
12	1239.542	620.2748	H	1122.663	563.8351	8
13	1352.626	676.8168	L	985.6041	493.806	7
14	1594.764	797.8858	K-GG	872.52	436.7636	6
15	1722.823	861.915	Q	630.3813	315.6947	5
16	1835.907	918.4571	I	502.3225	251.6654	4
17	1964.949	982.9784	E	389.2385	195.1234	3
18	2078.034	1039.52	I	260.1969	130.6021	2
19			K	147.1128	74.06004	1

```
1   MGDWSALGKL LDKVQAYSTA GGKVWLSVLF IFRILLLGTA VESAWGDEQS AFRCNTQQPG
                                                              *
61  CENVCYDKSF PISHVRFWVL QIIFVSVPTL LYLAHVFYVM RKEEKLNKKE EELKVAQTDG
              GG     *
121 VNVEMHLKQI EIKKFKYGIE EHGKVKMRGG LLRTYIISIL FKSVFEVAFL LIQWYIYGFS
181 LSAVYTCKRD PCPHQVDCFL SRPTEKTIFI IFMLVVSLVS LALNIIELFY VFFKGVKDRV
241 KGRSDPYHAT TGPLSPSKDC GSPKYAYFNG CSSPTAPLSP MSPPGYKLVT GDRNNSSCRN
301 YNKQASEQNW ANYSAEQNRM GQAGSTISNS HAQPFDFPDD SQNAKKVAAG HELQPLAIVD
361 QRPSSRASSR ASSRPRPDDL EI
```

Figure 12.2: A CHO protein identified using MS with detection of the ubiquitin remanent. GJA1 peptides analysed on the Orbitrap Fusion *VAQTDGVNVEMHLKQEIK*, with the ubiquitin diglycine (GG) tag on K14, from ubiquitin enriched lysate. Taken from CHO-DP12 cells, cultured at 31°C, 48 hours post temperature downshift. Full amino acid sequence for GJA1, with the peptides detected highlighted in green.

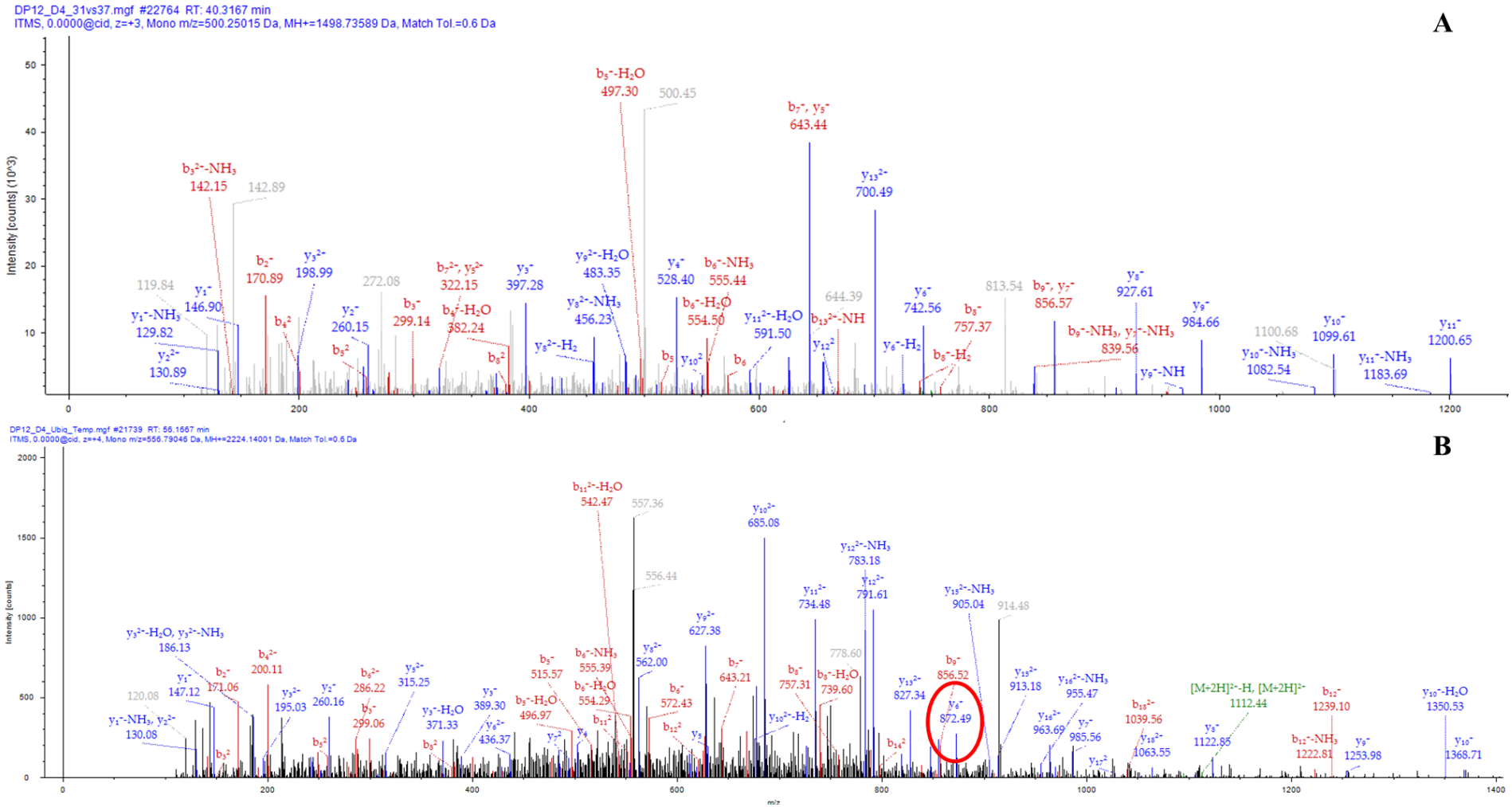


Figure 12.3: MS/MS fragment ion spectra from a GJA1 peptides analysed on the Orbitrap Fusion. (A) VAQTDGVNVEMHLK from whole cell lysate (B) VAQTDGVNVEMHLKQEIK, with the ubiquitin diglycine (GG) tag on K14, with the corresponding ion circled at 872.49 m/z, from ubiquitin enriched lysate. Both samples from CHO-DP12 cells, cultured at 31°C, 48 hours post temperature downshift. Match Tolerance=0.6 D

1.6 Conclusion

CHO cells are the predominant mammalian host for the industrial-scale production of therapeutic proteins due to their robust growth in suspension culture, adaptability to various media formulations, and capacity for complex post-translational modifications. Their long-standing use in biopharmaceutical manufacturing is supported by extensive optimisation of culture conditions, enabling high cell densities and maximised recombinant protein yields. Key environmental factors such as temperature, pH, nutrient availability, and oxygen levels directly influence cellular behaviour, with even slight variations capable of affecting product quality and process consistency. The manipulation of these parameters with precision has been crucial to the development of reliable and scalable production systems.

At the cellular level, the ER plays a central role in the synthesis, folding, and modification of proteins. Under conditions of increased protein synthesis or environmental stress, the ER can become overwhelmed, leading to the accumulation of misfolded proteins, a condition referred to as ER stress. To manage this, cells activate the UPR, a signalling network that temporarily halts protein translation, enhances the production of molecular chaperones, and promotes the degradation of misfolded proteins through ERAD pathways. However, if ER stress persists, these adaptive mechanisms can shift towards apoptosis, reducing cell viability and limiting production efficiency in biomanufacturing processes.

Ubiquitin plays a critical role in regulating these degradation pathways. The attachment of ubiquitin to substrate proteins marks them for recognition by the proteasome, facilitating the selective removal of damaged or misfolded proteins. Beyond degradation, ubiquitination also modulates numerous signalling pathways, acting as a versatile post-translational modifier that impacts cellular stress responses and protein homeostasis.

Proteomic technologies, particularly mass spectrometry, have emerged as essential tools for dissecting the molecular changes associated with CHO cell culture and ER stress. The

inclusion of PTMs in proteomic analyses offers a more comprehensive view of protein function and regulation under industrially relevant conditions. By mapping these modifications, researchers gain insight into stress response pathways and protein quality control mechanisms, informing strategies for improving both cell line robustness and therapeutic protein yields.

In-depth studies of the CHO proteome may hold the key to furthering the potential of these cells as recombinant protein hosts. In this thesis, the impact of both artificial and bioprocess-related stress on the CHO ER are investigated. This was done with the aim of expanding our understanding of the CHO ER stress response and how that response may be utilised to increase productivity. This included studying the role of the PTM ubiquitin in mediating ER stress and maintaining ER homeostasis.

1.7 References

Adams, B.M., Oster, M.E., Hebert, D.N., 2019. Protein quality control in the endoplasmic reticulum. Protein J. 38, 317–329.

Ahlstedt, B.A., Ganji, R., Raman, M., 2022. The functional importance of VCP to maintaining cellular protein homeostasis. Biochem. Soc. Trans. 50, 1457–1469. https://doi.org/10.1042/BST20220648

Ali, F., Dar, J.S., Magray, A.R., Ganai, B.A., Chishti, M.Z., 2019. Posttranslational modifications of proteins and their role in biological processes and associated diseases, in: Protein Modificomics. Elsevier, pp. 1–35.

Altamirano, C., Paredes, C., Illanes, A., Cairó, J.J., Gòdia, F., 2004. Strategies for fed-batch cultivation of t-PA producing CHO cells: substitution of glucose and glutamine and rational design of culture medium. J. Biotechnol. 110, 171–179. https://doi.org/10.1016/j.jbiotec.2004.02.004

An, H.J., Lebrilla, C.B., 2011. Structure elucidation of native N- and O-linked glycans by tandem mass spectrometry (tutorial). Mass Spectrom. Rev. 30, 560–578. https://doi.org/10.1002/mas.20283

Awad, H., Khamis, M.M., El-Aneed, A., 2015. Mass spectrometry, review of the basics: ionization. Appl. Spectrosc. Rev. 50, 158–175.

Baker, E.S., Livesay, E.A., Orton, D.J., Moore, R.J., Danielson III, W.F., Prior, D.C., Ibrahim, Y.M., LaMarche, B.L., Mayampurath, A.M., Schepmoes, A.A., 2010. An LC-IMS-MS platform providing increased dynamic range for high-throughput proteomic studies. J. Proteome Res. 9, 997–1006.

Bandaranayake, A.D., Almo, S.C., 2014. Recent advances in mammalian protein production. FEBS Lett. 588, 253–260. https://doi.org/10.1016/j.febslet.2013.11.035

Barzadd, M.M., Lundqvist, M., Harris, C., Malm, M., Volk, A.-L., Thalén, N., Chotteau, V., Grassi, L., Smith, A., Abadi, M.L., Lambiase, G., Gibson, S., Hatton, D., Rockberg, J., 2022. Autophagy and intracellular product degradation genes identified by systems biology analysis reduce aggregation of bispecific antibody in CHO cells. New Biotechnol. 68, 68–76. https://doi.org/10.1016/j.nbt.2022.01.010

Bausch, M., Schultheiss, C., Sieck, J.B., 2019. Recommendations for comparison of productivity between fed-batch and perfusion processes. Biotechnol. J. 14, 1700721.

Bazzoun, D., Adissu, H.A., Wang, L., Urazaev, A., Tenvooren, I., Fostok, S.F., Chittiboyina, S., Sturgis, J., Hodges, K., Chandramouly, G., Vidi, P.-A., Talhouk, R.S., Lelièvre, S.A., 2019. Connexin 43 maintains tissue polarity and regulates mitotic spindle orientation in the breast epithelium. J. Cell Sci. 132, jcs223313. https://doi.org/10.1242/jcs.223313

Becerra, S., Berrios, J., Osses, N., Altamirano, C., 2012a. Exploring the effect of mild hypothermia on CHO cell productivity. Biochem. Eng. J. 60, 1–8. https://doi.org/10.1016/j.bej.2011.10.003

Becerra, S., Berrios, J., Osses, N., Altamirano, C., 2012b. Exploring the effect of mild hypothermia on CHO cell productivity. Biochem. Eng. J. 60, 1–8. https://doi.org/10.1016/j.bej.2011.10.003

Becker, E., Florin, L., Pfizenmaier, K., Kaufmann, H., 2008. An XBP-1 dependent bottle-neck in production of IgG subtype antibodies in chemically defined serum-free Chinese hamster ovary (CHO) fed-batch processes. J. Biotechnol. 135, 217–223. https://doi.org/10.1016/j.jbiotec.2008.03.008

Becker, S., Feldmann, J., Wiedemann, S., Okamura, H., Schneider, C., Iwan, K., Crisp, A., Rossa, M., Amatov, T., Carell, T., 2019. Unified prebiotically plausible synthesis of pyrimidine and purine RNA ribonucleotides. Science 366, 76–82.

Bedford, L., Paine, S., Sheppard, P.W., Mayer, R.J., Roelofs, J., 2010. Assembly, structure, and function of the 26S proteasome. Trends Cell Biol. 20, 391–401. https://doi.org/10.1016/j.tcb.2010.03.007

Bedoya-López, A., Estrada, K., Sanchez-Flores, A., Ramírez, O.T., Altamirano, C., Segovia, L., Miranda-Ríos, J., Trujillo-Roldán, M.A., Valdez-Cruz, N.A., 2016. Effect of Temperature Downshift on the Transcriptomic Responses of Chinese Hamster Ovary Cells Using Recombinant Human Tissue Plasminogen Activator Production Culture. PLOS ONE 11, e0151529. https://doi.org/10.1371/journal.pone.0151529

Benham, A.M., 2012. The protein disulfide isomerase family: key players in health and disease. Antioxid. Redox Signal. 16, 781–789.

Bentley, M.L., Corn, J.E., Dong, K.C., Phung, Q., Cheung, T.K., Cochran, A.G., 2011. Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex. EMBO J. 30, 3285–3297. https://doi.org/10.1038/emboj.2011.243

Bernasconi, R., Soldà, T., Galli, C., Pertel, T., Luban, J., Molinari, M., 2010. Cyclosporine A-sensitive, cyclophilin B-dependent endoplasmic reticulum-associated degradation. PloS One 5, e13008.

Bernassola, F., Chillemi, G., Melino, G., 2019. HECT-Type E3 Ubiquitin Ligases in Cancer. Trends Biochem. Sci. 44, 1057–1075. https://doi.org/10.1016/j.tibs.2019.08.004

Berndsen, C.E., Wolberger, C., 2014. New insights into ubiquitin E3 ligase mechanism. Nat. Struct. Mol. Biol. 21, 301–307. https://doi.org/10.1038/nsmb.2780

Berrios, J., Altamirano, C., Osses, N., Gonzalez, R., 2011. Continuous CHO cell cultures with improved recombinant protein productivity by using mannose as carbon source: Metabolic analysis and scale-up simulation. Chem. Eng. Sci. 66, 2431–2439. https://doi.org/10.1016/j.ces.2011.03.011

Berrios, J., Díaz-Barrera, A., Bazán, C., Altamirano, C., 2009. Relationship between tissue plasminogen activator production and specific growth rate in Chinese hamster ovary cells cultured in mannose at low temperature. Biotechnol. Lett. 31, 1493–1497.

Bertrand, V., Karst, D.J., Bachmann, A., Cantalupo, K., Soos, M., Morbidelli, M., 2019. Transcriptome and proteome analysis of steady-state in a perfusion CHO cell culture process. Biotechnol. Bioeng. 116, 1959–1972.

Bhamidipati, A., Denic, V., Quan, E.M., Weissman, J.S., 2005. Exploration of the topological requirements of ERAD identifies Yos9p as a lectin sensor of misfolded glycoproteins in the ER lumen. Mol. Cell 19, 741–751. https://doi.org/10.1016/j.molcel.2005.07.027

Bi, J., Shuttleworth, J., Al-Rubeai, M., 2004. Uncoupling of cell growth and proliferation results in enhancement of productivity in p21CIP1-arrested CHO cells. Biotechnol. Bioeng. 85, 741–749.

Boja, E.S., Rodriguez, H., 2012. Mass spectrometry-based targeted quantitative proteomics: Achieving sensitive and reproducible detection of proteins. PROTEOMICS 12, 1093–1110. https://doi.org/10.1002/pmic.201100387

Borgese, N., Fasana, E., 2011. Targeting pathways of C-tail-anchored proteins. Biochim. Biophys. Acta BBA-Biomembr. 1808, 937–946.

Boye, E., Grallert, B., 2020. eIF2α phosphorylation and the regulation of translation. Curr. Genet. 66, 293–297. https://doi.org/10.1007/s00294-019-01026-1

Buszewski, B., Noga, S., 2012. Hydrophilic interaction liquid chromatography (HILIC)—a powerful separation technique. Anal. Bioanal. Chem. 402, 231–247. https://doi.org/10.1007/s00216-011-5308-5

Bythell, B.J., Suhai, S., Somogyi, Á., Paizs, B., 2009. Proton-Driven Amide Bond-Cleavage Pathways of Gas-Phase Peptide Ions Lacking Mobile Protons. J. Am. Chem. Soc. 131, 14057–14065. https://doi.org/10.1021/ja903883z

Cain, K., Peters, S., Hailu, H., Sweeney, B., Stephens, P., Heads, J., Sarkar, K., Ventom, A., Page, C., Dickson, A., 2013. A CHO cell line engineered to express XBP1 and ERO1-Lα has increased levels of transient protein expression. Biotechnol. Prog. 29, 697–706. https://doi.org/10.1002/btpr.1693

Caramelo, J.J., Parodi, A.J., 2008. Getting in and out from calnexin/calreticulin cycles. J. Biol. Chem. 283, 10221–10225.

Carver, J., Ng, D., Zhou, M., Ko, P., Zhan, D., Yim, M., Shaw, D., Snedecor, B., Laird, M.W., Lang, S., Shen, A., Hu, Z., 2020. Maximizing antibody production in a targeted integration host by optimization of subunit gene dosage and position. Biotechnol. Prog. 36, e2967. https://doi.org/10.1002/btpr.2967

Castellano, B.M., Tang, D., Marsters, S., Lam, C., Liu, P., Rose, C.M., Sandoval, W., Ashkenazi, A., Snedecor, B., Misaghi, S., 2023. Activation of the PERK branch of the unfolded protein response during production reduces specific productivity in CHO cells via downregulation of PDGFRa and IRE1a signaling. Biotechnol. Prog. e3354.

Chandrawanshi, V., Kulkarni, R., Prabhu, A., Mehra, S., 2020. Enhancing titers and productivity of rCHO clones with a combination of an optimized fed-batch process and ER-stress adaptation. J. Biotechnol. 311, 49–58. https://doi.org/10.1016/j.jbiotec.2020.02.008

Chartron, J.W., Clemons Jr, W.M., Suloway, C.J., 2012. The complex process of GETting tail-anchored membrane proteins to the ER. Curr. Opin. Struct. Biol. 22, 217–224.

Chee Furng Wong, D., Tin Kam Wong, K., Tang Goh, L., Kiat Heng, C., Gek Sim Yap, M., 2005. Impact of dynamic online fed-batch strategies on metabolism, productivity and N-glycosylation quality in CHO cell cultures. Biotechnol. Bioeng. 89, 164–177. https://doi.org/10.1002/bit.20317

Chen, P., Harcum, S.W., 2005. Effects of amino acid additions on ammonium stressed CHO cells. J. Biotechnol. 117, 277–286. https://doi.org/10.1016/j.jbiotec.2005.02.003

Chen, Z.J., 2012. Ubiquitination in signaling to and activation of IKK. Immunol. Rev. 246, 95–106.

Chen, Z.J., Sun, L.J., 2009. Nonproteolytic functions of ubiquitin in cell signaling. Mol. Cell 33, 275–286.

CHO Cell Culture Media Market Size, Share | Growth Research - 2033 [WWW Document], n.d. URL https://www.marketgrowthreports.com/market-reports/cho-cell-culture-media-market-102962?utm_source=chatgpt.com (accessed 9.10.25).

Christianson, J.C., Shaler, T.A., Tyler, R.E., Kopito, R.R., 2008. OS-9 and GRP94 deliver mutant α1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. Nat. Cell Biol. 10, 272–282. https://doi.org/10.1038/ncb1689

Clague, M.J., Heride, C., Urbé, S., 2015. The demographics of the ubiquitin system. Trends Cell Biol. 25, 417–426. https://doi.org/10.1016/j.tcb.2015.03.002

Crutchfield, C.A., Thomas, S.N., Sokoll, L.J., Chan, D.W., 2016. Advances in mass spectrometry-based clinical biomarker discovery. Clin. Proteomics 13, 1–12.

Cruz, H.J., Freitas, C.M., Alves, P.M., Moreira, J.L., Carrondo, M.J., 2000. Effects of ammonia and lactate on growth, metabolism, and productivity of BHK cells. Enzyme Microb. Technol. 27, 43–52. https://doi.org/10.1016/s0141-0229(00)00151-4

Dahlmann, B., 2016. Mammalian proteasome subtypes: Their diversity in structure and function. Arch. Biochem. Biophys. 591, 132–140.

Deshaies, R.J., Joazeiro, C.A., 2009. RING domain E3 ubiquitin ligases. Annu. Rev. Biochem. 78, 399–434.

Devarajan, A., Grijalva, V.R., Bourquard, N., Meriwether, D., Imaizumi, S., Shin, B.-C., Devaskar, S.U., Reddy, S.T., 2012. Macrophage Paraoxonase 2 regulates calcium homeostasis and cell survival under Endoplasmic Reticulum stress conditions and is sufficient to prevent the development of aggravated atherosclerosis in Paraoxonase 2 deficiency/apoE−/− mice on a Western diet. Mol. Genet. Metab. 107, 416–427. https://doi.org/10.1016/j.ymgme.2012.06.020

Doil, C., Mailand, N., Bekker-Jensen, S., Menard, P., Larsen, D.H., Pepperkok, R., Ellenberg, J., Panier, S., Durocher, D., Bartek, J., Lukas, J., Lukas, C., 2009. RNF168 Binds and Amplifies Ubiquitin Conjugates on Damaged Chromosomes to Allow Accumulation of Repair Proteins. Cell 136, 435–446. https://doi.org/10.1016/j.cell.2008.12.041

Dorner, A.J., Wasley, L.C., Kaufman, R.J., 1989. Increased synthesis of secreted proteins induces expression of glucose-regulated proteins in butyrate-treated Chinese hamster ovary cells. J. Biol. Chem. 264, 20602–20607.

Dove, K.K., Stieglitz, B., Duncan, E.D., Rittinger, K., Klevit, R.E., 2016. Molecular insights into RBR E3 ligase ubiquitin transfer mechanisms. EMBO Rep. 17, 1221–1235. https://doi.org/10.15252/embr.201642641

Dudek, J., Benedix, J., Cappel, S., Greiner, M., Jalal, C., Müller, L., Zimmermann, R., 2009. Functions and pathologies of BiP and its interaction partners. Cell. Mol. Life Sci. 66, 1556–1569.

Eddins, M.J., Carlile, C.M., Gomez, K.M., Pickart, C.M., Wolberger, C., 2006. Mms2–Ubc13 covalently bound to ubiquitin reveals the structural basis of linkage-specific polyubiquitin chain formation. Nat. Struct. Mol. Biol. 13, 915–920.

Ellgaard, L., Helenius, A., 2003. Quality control in the endoplasmic reticulum. Nat. Rev. Mol. Cell Biol. 4, 181–191. https://doi.org/10.1038/nrm1052

Ellgaard, L., Ruddock, L.W., 2005. The human protein disulphide isomerase family: substrate interactions and functional properties. EMBO Rep. 6, 28–32. https://doi.org/10.1038/sj.embor.7400311

Fan, L., Zhao, L., Sun, Y., Kou, T., Zhou, Y., Tan, W.-S., 2009. A high-yielding, generic fed-batch process for recombinant antibody production of GS-engineered cell lines. J. Microbiol. Biotechnol. 19, 1695–1702. https://doi.org/10.4014/jmb.0904.04054

Fernandes, P.M.P., Guedes, R.A., Victor, B.L., Salvador, J.A.R., Guedes, R.C., 2024. Decoding the secrets: how conformational and structural regulators inhibit the human 20S proteasome. Front. Chem. 11. https://doi.org/10.3389/fchem.2023.1322628

Ferris, S.P., Kodali, V.K., Kaufman, R.J., 2014. Glycoprotein folding and quality-control mechanisms in protein-folding diseases. Dis. Model. Mech. 7, 331–341. https://doi.org/10.1242/dmm.014589

Finley, D., 2009. Recognition and processing of ubiquitin-protein conjugates by the proteasome. Annu. Rev. Biochem. 78, 477–513.

Fischer, S., Otte, K., 2019. CHO Cell Engineering for Improved Process Performance and Product Quality, in: Cell Culture Engineering. John Wiley & Sons, Ltd, pp. 207–250. https://doi.org/10.1002/9783527811410.ch9

Fishman, G.I., Spray, D.C., Leinwand, L.A., 1990. Molecular characterization and functional expression of the human cardiac gap junction channel. J. Cell Biol. 111, 589–598. https://doi.org/10.1083/jcb.111.2.589

Fox, S.R., Tan, H.K., Tan, M.C., Wong, S.N.C., Yap, M.G., Wang, D.I., 2005. A detailed understanding of the enhanced hypothermic productivity of interferon-γ by Chinese-hamster ovary cells. Biotechnol. Appl. Biochem. 41, 255–264.

Gailly, P., 1998. Ca2+ entry in CHO cells, after Ca2+ stores depletion, is mediated by arachidonic acid. Cell Calcium 24, 293–304. https://doi.org/10.1016/s0143-4160(98)90053-7

Gardner, B.M., Walter, P., 2011. Unfolded Proteins are Ire1-Activating Ligands that Directly Induce the Unfolded Protein Response. Science 333, 1891–1894. https://doi.org/10.1126/science.1209126

Gawlitzek, M., Estacio, M., Fürch, T., Kiss, R., 2009. Identification of cell culture conditions to control N-glycosylation site-occupancy of recombinant glycoproteins expressed in CHO cells. Biotechnol. Bioeng. 103, 1164–1175. https://doi.org/10.1002/bit.22348

Gething, M.-J., 1999. Role and regulation of the ER chaperone BiP, in: Seminars in Cell & Developmental Biology. Elsevier, pp. 465–472.

Glish, G.L., Vachet, R.W., 2003. The basics of mass spectrometry in the twenty-first century. Nat. Rev. Drug Discov. 2, 140–150. https://doi.org/10.1038/nrd1011

Goey, C.H., Tsang, J.M.H., Bell, D., Kontoravdi, C., 2017. Cascading effect in bioprocessing—The impact of mild hypothermia on CHO cell behavior and host cell protein composition. Biotechnol. Bioeng. 114, 2771–2781. https://doi.org/10.1002/bit.26437

Goldknopf, I.L., Busch, H., 1977. Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. Proc. Natl. Acad. Sci. 74, 864–868.

Ha, T.K., Hansen, A.H., Kildegaard, H.F., Lee, G.M., 2019. BiP inducer X: an ER stress inhibitor for enhancing recombinant antibody production in CHO cell culture. Biotechnol. J. 14, 1900130.

Haglund, K., Sigismund, S., Polo, S., Szymkiewicz, I., Di Fiore, P.P., Dikic, I., 2003. Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. Nat. Cell Biol. 5, 461–466.

Hansen, F.M., Tanzer, M.C., Brüning, F., Bludau, I., Stafford, C., Schulman, B.A., Robles, M.S., Karayel, O., Mann, M., 2021. Data-independent acquisition method for ubiquitinome analysis reveals regulation of circadian biology. Nat. Commun. 12, 254. https://doi.org/10.1038/s41467-020-20509-1

Hansen, H.A., Emborg, C., 1994. Influence of ammonium on growth, metabolism, and productivity of a continuous suspension Chinese hamster ovary cell culture. Biotechnol. Prog. 10, 121–124.

Harding, H.P., Zhang, Y., Bertolotti, A., Zeng, H., Ron, D., 2000. Perk is essential for translational regulation and cell survival during the unfolded protein response. Mol. Cell 5, 897–904.

Harshbarger, W., Miller, C., Diedrich, C., Sacchettini, J., 2015. Crystal Structure of the Human 20S Proteasome in Complex with Carfilzomib. Structure 23, 418–424. https://doi.org/10.1016/j.str.2014.11.017

Hegde, R.S., Kang, S.-W., 2008. The concept of translocational regulation. J. Cell Biol. 182, 225–232. https://doi.org/10.1083/jcb.200804157

Hegde, R.S., Ploegh, H.L., 2010. Quality and quantity control at the endoplasmic reticulum. Curr. Opin. Cell Biol. 22, 437–446. https://doi.org/10.1016/j.ceb.2010.05.005

Helenius, A., 1994. How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. Mol. Biol. Cell 5, 253–265.

Hetz, C., Glimcher, L.H., 2009. Fine-Tuning of the Unfolded Protein Response: Assembling the IRE1α Interactome. Mol. Cell 35, 551–561. https://doi.org/10.1016/j.molcel.2009.08.021

Hetz, C., Martinon, F., Rodriguez, D., Glimcher, L.H., 2011. The unfolded protein response: integrating stress signals through the stress sensor IRE1α. Physiol. Rev. 91, 1219–1243.

Hetz, C., Zhang, K., Kaufman, R.J., 2020. Mechanism, regulation and functions of the unfolded protein response. Nat. Rev. Mol. Cell Biol. 21, 421–438. https://doi.org/10.1038/s41580-020-0250-z

Ho, S.C., Tong, Y.W., Yang, Y., 2013. Generation of monoclonal antibody-producing mammalian cell lines. Pharm. Bioprocess. 1, 71–87.

Holden, H.M., Thoden, J.B., Raushel, F.M., 1999. Carbamoyl phosphate synthetase: an amazing biochemical odyssey from substrate to product. Cell. Mol. Life Sci. CMLS 56, 507–522. https://doi.org/10.1007/s000180050448

Hong, J.K., Lakshmanan, M., Goudar, C., Lee, D.-Y., 2018. Towards next generation CHO cell line development and engineering by systems approaches. Curr. Opin. Chem. Eng. 22, 1–10.

Hong, M., Luo, S., Baumeister, P., Huang, J.-M., Gogia, R.K., Li, M., Lee, A.S., 2004. Underglycosylation of ATF6 as a novel sensing mechanism for activation of the unfolded protein response. J. Biol. Chem. 279, 11354–11363. https://doi.org/10.1074/jbc.M309804200

Hopfgartner, G., Varesio, E., Tschäppät, V., Grivet, C., Bourgogne, E., Leuthold, L.A., 2004. Triple quadrupole linear ion trap mass spectrometer for the analysis of small molecules and macromolecules. J. Mass Spectrom. 39, 845–855.

Huang, P., Liu, C., Gao, W., Chu, B., Cai, Z., Tian, R., 2021. Synergistic optimization of Liquid Chromatography and Mass Spectrometry parameters on Orbitrap Tribrid mass spectrometer for high efficient data-dependent proteomics. J. Mass Spectrom. 56, e4653.

Ingber, D.E., 2003. Tensegrity I. Cell structure and hierarchical systems biology. J. Cell Sci. 116, 1157–1173.

Inniss, M.C., Bandara, K., Jusiak, B., Lu, T.K., Weiss, R., Wroblewska, L., Zhang, L., 2017. A novel Bxb1 integrase RMCE system for high fidelity site-specific integration of mAb expression cassette in CHO cells. Biotechnol. Bioeng. 114, 1837–1846.

Ito, Y., Takeda, Y., Seko, A., Izumi, M., Kajihara, Y., 2015. Functional analysis of endoplasmic reticulum glucosyltransferase (UGGT): Synthetic chemistry's initiative in glycobiology. Semin. Cell Dev. Biol., Dendritic cell development and function & Protein quality control through glycosylation 41, 90–98. https://doi.org/10.1016/j.semcdb.2014.11.011

Jackson, P.K., Eldridge, A.G., Freed, E., Furstenthal, L., Hsu, J.Y., Kaiser, B.K., Reimann, J.D.R., Jackson, P.K., Eldridge, A.G., Freed, E., Furstenthal, L., Hsu, J.Y., Kaiser, B.K., Reimann, J.D.R., 2000. The lore of the RINGs: substrate recognition and catalysis by ubiquitin ligases. Trends Cell Biol. 10, 429–439. https://doi.org/10.1016/S0962-8924(00)01834-1

Jacobson, A.D., Zhang, N.-Y., Xu, P., Han, K.-J., Noone, S., Peng, J., Liu, C.-W., 2009. The lysine 48 and lysine 63 ubiquitin conjugates are processed differently by the 26 s proteasome. J. Biol. Chem. 284, 35485–35494. https://doi.org/10.1074/jbc.M109.052928

Jandera, P., 2013. Liquid chromatography| normal phase.

Jayapal, K.P., Wlaschin, K.F., Hu, W., Yap, M.G., 2007. Recombinant protein therapeutics from CHO cells-20 years and counting. Chem. Eng. Prog. 103, 40.

Jenkins, N., Hovey, A., 1993. Temperature control of growth and productivity in mutant Chinese hamster ovary cells synthesizing a recombinant protein. Biotechnol. Bioeng. 42, 1029–1036.

Ji, X., Lee, Y.J., Eyster, T., Parrillo, A., Galosy, S., Ao, Z., Patel, P., Zhu, Y., 2022. Characterization of cell cycle and apoptosis in Chinese hamster ovary cell culture using flow cytometry for bioprocess monitoring. Biotechnol. Prog. 38, e3211.

Jiang, Q., Sun, Y., Guo, Z., Fu, M., Wang, Q., Zhu, H., Lei, P., Shen, G., 2016. Overexpression of GRP78 enhances survival of CHO cells in response to serum deprivation and oxidative stress. Eng. Life Sci. 17, 107–116. https://doi.org/10.1002/elsc.201500152

Jin, J., Li, X., Gygi, S.P., Harper, J.W., 2007. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. Nature 447, 1135–1138. https://doi.org/10.1038/nature05902

Johari, Y.B., Estes, S.D., Alves, C.S., Sinacore, M.S., James, D.C., 2015. Integrated cell and process engineering for improved transient production of a “difficult-to-express “fusion protein by CHO cells. Biotechnol. Bioeng. 112, 2527–2542.

Kadowaki, H., Nagai, A., Maruyama, T., Takami, Y., Satrimafitrah, P., Kato, H., Honda, A., Hatta, T., Natsume, T., Sato, T., Kai, H., Ichijo, H., Nishitoh, H., 2015. Pre-emptive Quality Control Protects the ER from Protein Overload via the Proximity of ERAD Components and SRP. Cell Rep. 13, 944–956. https://doi.org/10.1016/j.celrep.2015.09.047

Kamadurai, H.B., Qiu, Y., Deng, A., Harrison, J.S., MacDonald, C., Actis, M., Rodrigues, P., Miller, D.J., Souphron, J., Lewis, S.M., 2013. Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. Elife 2, e00828.

Kampinga, H.H., Craig, E.A., 2010. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat. Rev. Mol. Cell Biol. 11, 579–592.

Karst, D.J., Scibona, E., Serra, E., Bielser, J., Souquet, J., Stettler, M., Broly, H., Soos, M., Morbidelli, M., Villiger, T.K., 2017. Modulation and modeling of monoclonal antibody N-linked glycosylation in mammalian cell perfusion reactors. Biotechnol. Bioeng. 114, 1978–1990.

Karst, D.J., Steinebach, F., Morbidelli, M., 2018. Continuous integrated manufacturing of therapeutic proteins. Curr. Opin. Biotechnol. 53, 76–84.

Kaufman, R.J., Sharp, P.A., 1982. Amplification and expression of sequences cotransfected with a modular dihydrofolate reductase complementary dna gene. J. Mol. Biol. 159, 601–621. https://doi.org/10.1016/0022-2836(82)90103-6

Kaufmann, H., Mazur, X., Fussenegger, M., Bailey, J.E., 1999. Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells. Biotechnol. Bioeng. 63, 573–582.

Kilgas, S., Ramadan, K., 2023. Inhibitors of the ATPase p97/VCP: From basic research to clinical applications. Cell Chem. Biol. 30, 3–21. https://doi.org/10.1016/j.chembiol.2022.12.007

Kim, B.G., Park, H.W., 2016. High zinc ion supplementation of more than 30 μM can increase monoclonal antibody production in recombinant Chinese hamster ovary DG44 cell culture. Appl. Microbiol. Biotechnol. 100, 2163–2170. https://doi.org/10.1007/s00253-015-7096-x

Kim, J.Y., Kim, Y.-G., Lee, G.M., 2012a. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl. Microbiol. Biotechnol. 93, 917–930.

Kim, J.Y., Kim, Y.-G., Lee, G.M., 2012b. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl. Microbiol. Biotechnol. 93, 917–930. https://doi.org/10.1007/s00253-011-3758-5

Kim, T.K., Chung, J.Y., Sung, Y.H., Lee, G.M., 2001. Relationship between cell size and specific thrombopoietin productivity in Chinese hamster ovary cells during dihydrofolate reductase-mediated gene amplification. Biotechnol. Bioprocess Eng. 6, 332–336.

Kim, W., Bennett, E.J., Huttlin, E.L., Guo, A., Li, J., Possemato, A., Sowa, M.E., Rad, R., Rush, J., Comb, M.J., Harper, J.W., Gygi, S.P., 2011. Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome. Mol. Cell 44, 325–340. https://doi.org/10.1016/j.molcel.2011.08.025

Komander, D., Rape, M., 2012. The ubiquitin code. Annu. Rev. Biochem. 81, 203–229. https://doi.org/10.1146/annurev-biochem-060310-170328

Komuczki, D., Stadermann, A., Bentele, M., Unsoeld, A., Grillari, J., Mueller, M.M., Paul, A., Fischer, S., 2022. High cysteine concentrations in cell culture media lead to oxidative stress and reduced bioprocess performance of recombinant CHO cells. Biotechnol. J. 17, 2200029. https://doi.org/10.1002/biot.202200029

Korennykh, A.V., Egea, P.F., Korostelev, A.A., Finer-Moore, J., Zhang, C., Shokat, K.M., Stroud, R.M., Walter, P., 2009. The unfolded protein response signals through high-order assembly of Ire1. Nature 457, 687–693.

Kou, T.-C., Fan, L., Zhou, Y., Ye, Z.-Y., Liu, X.-P., Zhao, L., Tan, W.-S., 2011. Detailed understanding of enhanced specific productivity in Chinese hamster ovary cells at low culture temperature. J. Biosci. Bioeng. 111, 365–369.

Kozlov, G., Määttänen, P., Thomas, D.Y., Gehring, K., 2010. A structural overview of the PDI family of proteins. FEBS J. 277, 3924–3936.

Krokhin, O.V., Craig, R., Spicer, V., Ens, W., Standing, K.G., Beavis, R.C., Wilkins, J.A., 2004. An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. Mol. Cell. Proteomics MCP 3, 908–919. https://doi.org/10.1074/mcp.M400031-MCP200

Kuda, O., Jenkins, C.M., Skinner, J.R., Moon, S.H., Su, X., Gross, R.W., Abumrad, N.A., 2011. CD36 Protein Is Involved in Store-operated Calcium Flux, Phospholipase A2 Activation, and Production of Prostaglandin E2*. J. Biol. Chem. 286, 17785–17795. https://doi.org/10.1074/jbc.M111.232975

Kurano, N., Leist, C., Messi, F., Kurano, S., Fiechter, A., 1990. Growth behavior of Chinese hamster ovary cells in a compact loop bioreactor. 2. Effects of medium components and waste products. J. Biotechnol. 15, 113–128. https://doi.org/10.1016/0168-1656(90)90055-g

Kushnir, N., Streatfield, S.J., Yusibov, V., 2012. Virus-like particles as a highly efficient vaccine platform: Diversity of targets and production systems and advances in clinical development. Vaccine 31, 58–83. https://doi.org/10.1016/j.vaccine.2012.10.083

Kuwae, S., Miyakawa, I., Doi, T., 2018. Development of a chemically defined platform fed-batch culture media for monoclonal antibody-producing CHO cell lines with optimized choline content. Cytotechnology 70, 939–948. https://doi.org/10.1007/s10616-017-0185-1

Lam, Y.A., Lawson, T.G., Velayutham, M., Zweier, J.L., Pickart, C.M., 2002. A proteasomal ATPase subunit recognizes the polyubiquitin degradation signal. Nature 416, 763–767. https://doi.org/10.1038/416763a

Lander, G.C., Estrin, E., Matyskiela, M.E., Bashore, C., Nogales, E., Martin, A., 2012. Complete subunit architecture of the proteasome regulatory particle. Nature 482, 186–191.

Latorre, Y., Torres, M., Vergara, M., Berrios, J., Sampayo, M.M., Gödecke, N., Wirth, D., Hauser, H., Dickson, A.J., Altamirano, C., 2023. Engineering of Chinese hamster ovary cells for co-overexpressing MYC and XBP1s increased cell proliferation and recombinant EPO production. Sci. Rep. 13, 1482.

Le, H., Chen, H., Goudar, C.T., 2015. Continuous processing in upstream operations. Chem. Eng. Prog. 111, 32–37.

Lederkremer, G.Z., Glickman, M.H., 2005. A window of opportunity: timing protein degradation by trimming of sugars and ubiquitins. Trends Biochem. Sci. 30, 297–303. https://doi.org/10.1016/j.tibs.2005.04.010

Lemire, L., Pham, P.L., Durocher, Y., Henry, O., 2022. Practical considerations for the scale-up of Chinese hamster ovary (CHO) cell cultures, in: Cell Culture Engineering and Technology: In Appreciation to Professor Mohamed Al-Rubeai. Springer, pp. 367–400.

Li, M., Baumeister, P., Roy, B., Phan, T., Foti, D., Luo, S., Lee, A.S., 2000. ATF6 as a transcription activator of the endoplasmic reticulum stress element: thapsigargin stress-induced changes and synergistic interactions with NF-Y and YY1. Mol. Cell. Biol. 20, 5096–5106.

Li, Y., Guo, Y., Tang, J., Jiang, J., Chen, Z., 2014. New insights into the roles of CHOP-induced apoptosis in ER stress. Acta Biochim Biophys Sin 46, 629–640.

Liang, K., Luo, H., Li, Q., 2024. Optimization of the Process of Chinese Hamster Ovary (CHO) Cell Fed-Batch Culture to Stabilize Monoclonal Antibody Production and Overall Quality: Effect of pH Control Strategies. Fermentation 10, 352. https://doi.org/10.3390/fermentation10070352

Liew, C.W., Sun, H., Hunter, T., Day, C.L., 2010. RING domain dimerization is essential for RNF4 function. Biochem. J. 431, 23–29. https://doi.org/10.1042/BJ20100957

Liu, H., Zhou, C., An, J., Song, Y., Yu, P., Li, J., Gu, C., Hu, D., Jiang, Y., Zhang, L., Huang, C., Zhang, C., Yang, Y., Zhu, Q., Wang, D., Liu, Y., Miao, C., Cao, X., Ding, L., Zhu, Y., Zhu, H., Bao, L., Zhou, L., Yan, H., Fan, J., Xu, J., Hu, Z., Xie, Y., Liu, J., Liu, G., 2021. Development of recombinant COVID-19 vaccine based on CHO-produced, prefusion spike trimer and alum/CpG adjuvants. Vaccine 39, 7001–7011. https://doi.org/10.1016/j.vaccine.2021.10.066

Liu, T., Belov, M.E., Jaitly, N., Qian, W.-J., Smith, R.D., 2007. Accurate mass measurements in proteomics. Chem. Rev. 107, 3621–3653.

Lu, S., Sun, X., Zhang, Y., 2005. Insight into metabolism of CHO cells at low glucose concentration on the basis of the determination of intracellular metabolites. Process Biochem. 40, 1917–1921.

Mahé, A., Martiné, A., Fagète, S., Girod, P.-A., 2022. Exploring the limits of conventional small-scale CHO fed-batch for accelerated on demand monoclonal antibody production. Bioprocess Biosyst. Eng. 45, 297–307.

Malhotra, J.D., Kaufman, R.J., 2011. ER Stress and Its Functional Link to Mitochondria: Role in Cell Survival and Death. Cold Spring Harb. Perspect. Biol. 3, a004424. https://doi.org/10.1101/cshperspect.a004424

Mao, C., Tai, W.-C., Bai, Y., Poizat, C., Lee, A.S., 2006. In vivo regulation of Grp78/BiP transcription in the embryonic heart: role of the endoplasmic reticulum stress response element and GATA-4. J. Biol. Chem. 281, 8877–8887. https://doi.org/10.1074/jbc.M505784200

Marchant, R.J., Al-Fageeh, M.B., Underhill, M.F., Racher, A.J., Smales, C.M., 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, 69–77. https://doi.org/10.1007/s12033-008-9032-0

Maria, S., Bonneau, L., Fould, B., Ferry, G., Boutin, J.A., Cabanne, C., Santarelli, X., Joucla, G., 2023. Perfusion process for CHO cell producing monoclonal antibody: Comparison of methods to determine optimum cell specific perfusion rate. Biochem. Eng. J. 191, 108779. https://doi.org/10.1016/j.bej.2022.108779

Mason, M., Sweeney, B., Cain, K., Stephens, P., Sharfstein, S.T., 2014. Reduced Culture Temperature Differentially Affects Expression and Biophysical Properties of Monoclonal Antibody Variants. Antibodies Basel Switz. 3, 253–271. https://doi.org/10.3390/antib3030253

Masterton, R.J., Smales, C.M., 2014. The impact of process temperature on mammalian cell lines and the implications for the production of recombinant proteins in CHO cells. Pharm. Bioprocess. 2, 49–61.

Mathias, S., Wippermann, A., Raab, N., Zeh, N., Handrick, R., Gorr, I., Schulz, P., Fischer, S., Gamer, M., Otte, K., 2020. Unraveling what makes a monoclonal antibody difficult-to-express: From intracellular accumulation to incomplete folding and degradation via ERAD. Biotechnol. Bioeng. 117, 5–16.

Matsusaki, M., Kanemura, S., Kinoshita, M., Lee, Y.-H., Inaba, K., Okumura, M., 2020. The Protein Disulfide Isomerase Family: from proteostasis to pathogenesis. Biochim. Biophys. Acta Gen. Subj. 1864, 129338. https://doi.org/10.1016/j.bbagen.2019.04.003

Mayer, M., Frey, S., Koivunen, P., Myllyharju, J., Buchner, J., 2004. Influence of the Oxidoreductase ERp57 on the Folding of an Antibody Fab Fragment. J. Mol. Biol. 341, 1077–1084. https://doi.org/10.1016/j.jmb.2004.06.068

Mikami, T., Aoki, M., Kimura, T., 2012. The application of mass spectrometry to proteomics and metabolomics in biomarker discovery and drug development. Curr. Mol. Pharmacol. 5, 301–316.

Min, H., Kim, S.M., Kim, D., Lee, Solhwi, Lee, Sumin, Lee, J.S., 2022. Hybrid cell line development system utilizing site-specific integration and methotrexate-mediated gene amplification in Chinese hamster ovary cells. Front. Bioeng. Biotechnol. 10. https://doi.org/10.3389/fbioe.2022.977193

Mnatsakanyan, R., Shema, G., Basik, M., Batist, G., Borchers, C.H., Sickmann, A., Zahedi, R.P., 2018. Detecting post-translational modification signatures as potential biomarkers in clinical mass spectrometry. Expert Rev. Proteomics 15, 515–535.

Mollah, S., Wertz, I.E., Phung, Q., Arnott, D., Dixit, V.M., Lill, J.R., 2007. Targeted mass spectrometric strategy for global mapping of ubiquitination on proteins. Rapid Commun. Mass Spectrom. Int. J. Devoted Rapid Dissem. Up-to--Minute Res. Mass Spectrom. 21, 3357–3364.

Moore, A., Mercer, J., Dutina, G., Donahue, C.J., Bauer, K.D., Mather, J.P., Etcheverry, T., Ryll, T., 1997. Effects of temperature shift on cell cycle, apoptosis and nucleotide pools in CHO cell batch cultues. Cytotechnology 23, 47–54.

Moore, K., Hollien, J., 2015. Ire1-mediated decay in mammalian cells relies on mRNA sequence, structure, and translational status. Mol. Biol. Cell 26, 2873–2884. https://doi.org/10.1091/mbc.E15-02-0074

Neill, G., Masson, G.R., 2023. A stay of execution: ATF4 regulation and potential outcomes for the integrated stress response. Front. Mol. Neurosci. 16. https://doi.org/10.3389/fnmol.2023.1112253

Neilson, K.A., Ali, N.A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., Lee, A., Van Sluyter, S.C., Haynes, P.A., 2011. Less label, more free: approaches in label-free quantitative mass spectrometry. Proteomics 11, 535–553.

Novoa, I., Zeng, H., Harding, H.P., Ron, D., 2001. Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2α. J. Cell Biol. 153, 1011–1022.

Oguchi, S., Saito, H., Tsukahara, M., Tsumura, H., 2006. pH Condition in temperature shift cultivation enhances cell longevity and specific hMab productivity in CHO culture. Cytotechnology 52, 199–207.

Ohya, T., Hayashi, T., Kiyama, E., Nishii, H., Miki, H., Kobayashi, K., Honda, K., Omasa, T., Ohtake, H., 2008. Improved production of recombinant human antithrombin III in Chinese hamster ovary cells by ATF4 overexpression. Biotechnol. Bioeng. 100, 317–324. https://doi.org/10.1002/bit.21758

Oliver, J.D., Van Der Wal, F.J., Bulleid, N.J., High, S., 1997. Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. Science 275, 86–88.

Oslowski, C.M., Urano, F., 2011. Measuring ER stress and the unfolded protein response using mammalian tissue culture system. Methods Enzymol. 490, 71–92. https://doi.org/10.1016/B978-0-12-385114-7.00004-0

Palomares, L.A., Goudar, C., Wang, T., 2023. Conference program-Cell Culture Engineering XVIII.

Pan, X., Streefland, M., Dalm, C., Wijffels, R.H., Martens, D.E., 2017. Selection of chemically defined media for CHO cell fed-batch culture processes. Cytotechnology 69, 39–56. https://doi.org/10.1007/s10616-016-0036-5

Patel, Y.D., Brown, A.J., Zhu, J., Rosignoli, G., Gibson, S.J., Hatton, D., James, D.C., 2021. Control of multigene expression stoichiometry in mammalian cells using synthetic promoters. ACS Synth. Biol. 10, 1155–1165.

Patkar, S.A., Rasmussen, U., Diamant, B., 1979. On the mechanism of histamine release induced by thapsigargin from Thapsia garganica L. Agents Actions 9, 53–57.

Peng, J., Schwartz, D., Elias, J.E., Thoreen, C.C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., Gygi, S.P., 2003. A proteomics approach to understanding protein ubiquitination. Nat. Biotechnol. 21, 921–926.

Peth, A., Besche, H.C., Goldberg, A.L., 2009. Ubiquitinated proteins activate the proteasome by binding to Usp14/Ubp6, which causes 20S gate opening. Mol. Cell 36, 794–804.

Pieper, L.A., Strotbek, M., Wenger, T., Olayioye, M.A., Hausser, A., 2017. ATF6β-based fine-tuning of the unfolded protein response enhances therapeutic antibody productivity of Chinese hamster ovary cells. Biotechnol. Bioeng. 114, 1310–1318.

Plechanovová, A., Jaffray, E.G., Tatham, M.H., Naismith, J.H., Hay, R.T., 2012. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. Nature 489, 115–120. https://doi.org/10.1038/nature11376

Pollock, J., Ho, S.V., Farid, S.S., 2013. Fed-batch and perfusion culture processes: Economic, environmental, and operational feasibility under uncertainty. Biotechnol. Bioeng. 110, 206–219.

Pozarowski, P., Darzynkiewicz, Z., 2004. Analysis of cell cycle by flow cytometry, in: Checkpoint Controls and Cancer: Volume 2: Activation and Regulation Protocols. Springer, pp. 301–311.

Prashad, K., Mehra, S., 2015. Dynamics of unfolded protein response in recombinant CHO cells. Cytotechnology 67, 237–254. https://doi.org/10.1007/s10616-013-9678-8

Puck, T.T., 1972. The mammalian cell as a microorganism. Holden-Day San Francisco.

Puig, A., Gilbert, H.F., 1994. Protein disulfide isomerase exhibits chaperone and anti-chaperone activity in the oxidative refolding of lysozyme. J. Biol. Chem. 269, 7764–7771.

Qian, Y., Khattak, S.F., Xing, Z., He, A., Kayne, P.S., Qian, N.-X., Pan, S.-H., Li, Z.J., 2011. Cell culture and gene transcription effects of copper sulfate on Chinese hamster ovary cells. Biotechnol. Prog. 27, 1190–1194. https://doi.org/10.1002/btpr.630

Qu, J., Zou, T., Lin, Z., 2021. The Roles of the Ubiquitin–Proteasome System in the Endoplasmic Reticulum Stress Pathway. Int. J. Mol. Sci. 22, 1526. https://doi.org/10.3390/ijms22041526

Rahimpour, A., Vaziri, B., Moazzami, R., Nematollahi, L., Barkhordari, F., Kokabee, L., Adeli, A., Mahboudi, F., 2013. Engineering the cellular protein secretory pathway for enhancement of recombinant tissue plasminogen activator expression in Chinese hamster ovary cells: effects of CERT and XBP1s genes. J. Microbiol. Biotechnol. 23, 1116–1122. https://doi.org/10.4014/jmb.1302.02035

Ramazi, S., Zahiri, J., 2021. Post-translational modifications in proteins: resources, tools and prediction methods. Database 2021, baab012. https://doi.org/10.1093/database/baab012

Richter, J., Wang, Q., Lange, F., Thiel, P., Yilmaz, N., Solle, D., Zhuang, X., Beutel, S., 2025. Machine Learning-Powered Optimization of a CHO Cell Cultivation Process. Biotechnol. Bioeng. 122, 1153–1164. https://doi.org/10.1002/bit.28943

Ritacco, F.V., Wu, Y., Khetan, A., 2018. Cell culture media for recombinant protein expression in Chinese hamster ovary (CHO) cells: History, key components, and optimization strategies. Biotechnol. Prog. 34, 1407–1426. https://doi.org/10.1002/btpr.2706

Rotin, D., Kumar, S., 2009. Physiological functions of the HECT family of ubiquitin ligases. Nat. Rev. Mol. Cell Biol. 10, 398–409.

Roy, G., Zhang, S., Li, L., Higham, E., Wu, H., Marelli, M., Bowen, M.A., 2017. Development of a fluorescent reporter system for monitoring ER stress in Chinese hamster ovary cells and its application for therapeutic protein production. PLoS ONE 12, e0183694. https://doi.org/10.1371/journal.pone.0183694

Rozanova, S., Barkovits, K., Nikolov, M., Schmidt, C., Urlaub, H., Marcus, K., 2021. Quantitative Mass Spectrometry-Based Proteomics: An Overview, in: Marcus, K., Eisenacher, M., Sitek, B. (Eds.), Quantitative Methods in Proteomics, Methods in Molecular Biology. Springer US, New York, NY, pp. 85–116. https://doi.org/10.1007/978-1-0716-1024-4_8

Rutkevich, L.A., Williams, D.B., 2011. Participation of lectin chaperones and thiol oxidoreductases in protein folding within the endoplasmic reticulum. Curr. Opin. Cell Biol., Cell regulation 23, 157–166. https://doi.org/10.1016/j.ceb.2010.10.011

Ryan, D., Sideri, C.-K., Henry, M., Efeoglu, E., Meleady, P., 2023. Label-Free Quantitative Proteomics Analysis of Producer and Non-producer Chinese Hamsters Ovary (CHO) Cells under ER Stress Conditions. Curr. Res. Biotechnol. 100141.

Saha, A., Oanca, G., Mondal, D., Warshel, A., 2020. Exploring the proteolysis mechanism of the proteasomes. J. Phys. Chem. B 124, 5626–5635.

Samaras, J.J., Micheletti, M., Ding, W., 2022. Transformation of biopharmaceutical manufacturing through single-use technologies: current state, remaining challenges, and future development. Annu. Rev. Chem. Biomol. Eng. 13, 73–97.

Samy, A., Yamano-Adachi, N., Koga, Y., Omasa, T., 2021a. Secretion of a low-molecular-weight species of endogenous GRP94 devoid of the KDEL motif during endoplasmic reticulum stress in Chinese hamster ovary cells. Traffic 22, 425–438. https://doi.org/10.1111/tra.12818

Samy, A., Yamano-Adachi, N., Koga, Y., Omasa, T., 2021b. Secretion of a low-molecular-weight species of endogenous GRP94 devoid of the KDEL motif during endoplasmic reticulum stress in Chinese hamster ovary cells. Traffic Cph. Den. 22, 425–438. https://doi.org/10.1111/tra.12818

Sanchez-Martinez, Z.V., Alpuche-Lazcano, S.P., Stuible, M., Durocher, Y., 2024. CHO cells for virus-like particle and subunit vaccine manufacturing. Vaccine 42, 2530–2542. https://doi.org/10.1016/j.vaccine.2024.03.034

Saraogi, I., Shan, S., 2011. Molecular Mechanism of Co-translational Protein Targeting by the Signal Recognition Particle. Traffic 12, 535–542. https://doi.org/10.1111/j.1600-0854.2011.01171.x

Savizi, I.S.P., Maghsoudi, N., Motamedian, E., Lewis, N.E., Shojaosadati, S.A., 2022. Valine feeding reduces ammonia production through rearrangement of metabolic fluxes in central carbon metabolism of CHO cells. Appl. Microbiol. Biotechnol. 106, 1113–1126. https://doi.org/10.1007/s00253-021-11755-4

Schellenberg, J., Nagraik, T., Wohlenberg, O.J., Ruhl, S., Bahnemann, J., Scheper, T., Solle, D., 2022a. Stress-induced increase of monoclonal antibody production in CHO cells. Eng. Life Sci. 22, 427–436.

Schellenberg, J., Nagraik, T., Wohlenberg, O.J., Ruhl, S., Bahnemann, J., Scheper, T., Solle, D., 2022b. Stress-induced increase of monoclonal antibody production in CHO cells. Eng. Life Sci. 22, 427–436. https://doi.org/10.1002/elsc.202100062

Schlatter, S., Stansfield, S.H., Dinnis, D.M., Racher, A.J., Birch, J.R., James, D.C., 2005. On the optimal ratio of heavy to light chain genes for efficient recombinant antibody production by CHO cells. Biotechnol. Prog. 21, 122–133.

Schröder, M., Friedl, P., 1997. Overexpression of recombinant human antithrombin III in Chinese hamster ovary cells results in malformation and decreased secretion of recombinant protein. Biotechnol. Bioeng. 53, 547–559.

Schröder, M., Kaufman, R.J., 2005. ER stress and the unfolded protein response. Mutat. Res. Mol. Mech. Mutagen. 569, 29–63.

Schuldiner, M., Metz, J., Schmid, V., Denic, V., Rakwalska, M., Schmitt, H.D., Schwappach, B., Weissman, J.S., 2008. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. Cell 134, 634–645.

Schulman, B.A., Harper, J.W., 2009. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nat. Rev. Mol. Cell Biol. 10, 319–331. https://doi.org/10.1038/nrm2673

Seabrook, G.R., Fong, T.M., 1993. Thapsigargin blocks the mobilisation of intracellular calcium caused by activation of human NK1 (long) receptors expressed in Chinese hamster ovary cells. Neurosci. Lett. 152, 9–12. https://doi.org/10.1016/0304-3940(93)90470-6

Sehgal, P., Szalai, P., Olesen, C., Praetorius, H.A., Nissen, P., Christensen, S.B., Engedal, N., Møller, J.V., 2017. Inhibition of the sarco/endoplasmic reticulum (ER) Ca2+-ATPase by thapsigargin analogs induces cell death via ER Ca2+ depletion and the unfolded protein response. J. Biol. Chem. 292, 19656–19673.

Senft, D., Qi, J., Ronai, Z.A., 2018. Ubiquitin ligases in oncogenic transformation and cancer therapy. Nat. Rev. Cancer 18, 69–88. https://doi.org/10.1038/nrc.2017.105

Shen, J., Snapp, E.L., Lippincott-Schwartz, J., Prywes, R., 2005. Stable binding of ATF6 to BiP in the endoplasmic reticulum stress response. Mol. Cell. Biol. 25, 921–932. https://doi.org/10.1128/MCB.25.3.921-932.2005

Shin, S.W., Lee, J.S., 2020. CHO cell line development and engineering via site-specific integration: challenges and opportunities. Biotechnol. Bioprocess Eng. 25, 633–645.

Sieblist, C., Jenzsch, M., Pohlscheidt, M., Lübbert, A., 2011. Insights into large-scale cell-culture reactors: I. Liquid mixing and oxygen supply. Biotechnol. J. 6, 1532–1546. https://doi.org/10.1002/biot.201000408

Sinharoy, P., Aziz, A.H., Majewska, N.I., Ahuja, S., Handlogten, M.W., 2020. Perfusion reduces bispecific antibody aggregation via mitigating mitochondrial dysfunction-induced glutathione oxidation and ER stress in CHO cells. Sci. Rep. 10, 16620. https://doi.org/10.1038/s41598-020-73573-4

Smales, C.M., Dinnis, D.M., Stansfield, S.H., Alete, D., Sage, E.A., Birch, J.R., Racher, A.J., Marshall, C.T., James, D.C., 2004. Comparative proteomic analysis of GS-NS0 murine myeloma cell lines with varying recombinant monoclonal antibody production rate. Biotechnol. Bioeng. 88, 474–488.

Smith, D.M., Chang, S.-C., Park, S., Finley, D., Cheng, Y., Goldberg, A.L., 2007. Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's α ring opens the gate for substrate entry. Mol. Cell 27, 731–744.

Soldà, T., Galli, C., Kaufman, R.J., Molinari, M., 2007. Substrate-specific requirements for UGT1-dependent release from calnexin. Mol. Cell 27, 238–249.

Spratt, D.E., Walden, H., Shaw, G.S., 2014. RBR E3 ubiquitin ligases: new structures, new insights, new questions. Biochem. J. 458, 421–437. https://doi.org/10.1042/BJ20140006

Stewart, G.S., Panier, S., Townsend, K., Al-Hakim, A.K., Kolas, N.K., Miller, E.S., Nakada, S., Ylanko, J., Olivarius, S., Mendez, M., 2009. The RIDDLE syndrome protein mediates a ubiquitin-dependent signaling cascade at sites of DNA damage. Cell 136, 420–434.

Strugnell, R., Zepp, F., Cunningham, A., Tantawichien, T., 2011. Vaccine antigens. Perspect. Vaccinol. 1, 61–88. https://doi.org/10.1016/j.pervac.2011.05.003

Su, M., Zhang, Z., Zhou, L., Han, C., Huang, C., Nice, E.C., 2021. Proteomics, Personalized Medicine and Cancer. Cancers 13, 2512. https://doi.org/10.3390/cancers13112512

Sulaj, E., Schwaigerlehner, L., Sandell, F.L., Dohm, J.C., Marzban, G., Kunert, R., 2024. Quantitative proteomics reveals cellular responses to individual mAb expression and tunicamycin in CHO cells. Appl. Microbiol. Biotechnol. 108, 1–19. https://doi.org/10.1007/s00253-024-13223-1

Swatek, K.N., Komander, D., 2016. Ubiquitin modifications. Cell Res. 26, 399–422. https://doi.org/10.1038/cr.2016.39

Synoground, B.F., McGraw, C.E., Elliott, K.S., Leuze, C., Roth, J.R., Harcum, S.W., Sandoval, N.R., 2021. Transient ammonia stress on Chinese hamster ovary (CHO) cells yield alterations to alanine metabolism and IgG glycosylation profiles. Biotechnol. J. 16, 2100098. https://doi.org/10.1002/biot.202100098

Takagi, Y., Kikuchi, T., Wada, R., Omasa, T., 2017. The enhancement of antibody concentration and achievement of high cell density CHO cell cultivation by adding nucleoside. Cytotechnology 69, 511–521. https://doi.org/10.1007/s10616-017-0066-7

Tanaka, K., 2009. The proteasome: Overview of structure and functions. Proc. Jpn. Acad. Ser. B Phys. Biol. Sci. 85, 12–36. https://doi.org/10.2183/pjab.85.12

Tanemura, H., Masuda, K., Okumura, T., Takagi, E., Kajihara, D., Kakihara, H., Nonaka, K., Ushioda, R., 2022. Development of a stable antibody production system utilizing an Hspa5 promoter in CHO cells. Sci. Rep. 12, 7239.

Tang, D., Sandoval, W., Lam, C., Haley, B., Liu, P., Xue, D., Roy, D., Patapoff, T., Louie, S., Snedecor, B., 2020. UBR E3 ligases and the PDIA3 protease control degradation of unfolded antibody heavy chain by ERAD. J. Cell Biol. 219, e201908087.

Tihanyi, B., Nyitray, L., 2020. Recent advances in CHO cell line development for recombinant protein production. Drug Discov. Today Technol. 38, 25–34.

Tokunaga, F., Sakata, S., Saeki, Y., Satomi, Y., Kirisako, T., Kamei, K., Nakagawa, T., Kato, M., Murata, S., Yamaoka, S., Yamamoto, M., Akira, S., Takao, T., Tanaka, K., Iwai, K., 2009. Involvement of linear polyubiquitylation of NEMO in NF-κB activation. Nat. Cell Biol. 11, 123–132. https://doi.org/10.1038/ncb1821

Tomita, T., Matouschek, A., 2019. Substrate selection by the proteasome through initiation regions. Protein Sci. Publ. Protein Soc. 28, 1222–1232. https://doi.org/10.1002/pro.3642

Tong, X., Kawabata, H., Koeffler, H.P., 2002. Iron deficiency can upregulate expression of transferrin receptor at both the mRNA and protein level. Br. J. Haematol. 116, 458–464.

Torres, M., Akhtar, S., McKenzie, E.A., Dickson, A.J., 2021. Temperature Down-Shift Modifies Expression of UPR-/ERAD-Related Genes and Enhances Production of a Chimeric Fusion Protein in CHO Cells. Biotechnol. J. 16, 2000081. https://doi.org/10.1002/biot.202000081

Tracz, M., Bialek, W., 2021. Beyond K48 and K63: non-canonical protein ubiquitination. Cell. Mol. Biol. Lett. 26, 1–17. https://doi.org/10.1186/s11658-020-00245-6

Trummer, E., Fauland, K., Seidinger, S., Schriebl, K., Lattenmayer, C., Kunert, R., Vorauer-Uhl, K., Weik, R., Borth, N., Katinger, H., 2006. Process parameter shifting: Part II. Biphasic cultivation—A tool for enhancing the volumetric productivity of batch processes using Epo-Fc expressing CHO cells. Biotechnol. Bioeng. 94, 1045–1052.

Valerio, L.G., 2007. Mammalian iron metabolism. Toxicol. Mech. Methods 17, 497–517. https://doi.org/10.1080/15376510701556690

van der Wal, L., Bezstarosti, K., Sap, K.A., Dekkers, D.H.W., Rijkers, E., Mientjes, E., Elgersma, Y., Demmers, J.A.A., 2018. Improvement of ubiquitylation site detection by Orbitrap mass spectrometry. J. Proteomics 172, 49–56. https://doi.org/10.1016/j.jprot.2017.10.014

van Huizen, M., Kikkert, M., 2020. The Role of Atypical Ubiquitin Chains in the Regulation of the Antiviral Innate Immune Response. Front. Cell Dev. Biol. 7. https://doi.org/10.3389/fcell.2019.00392

Vanhove, M., Usherwood, Y.-K., Hendershot, L.M., 2001. Unassembled Ig heavy chains do not cycle from BiP in vivo but require light chains to trigger their release. Immunity 15, 105–114.

Wahrheit, J., Nicolae, A., Heinzle, E., 2014. Dynamics of growth and metabolism controlled by glutamine availability in Chinese hamster ovary cells. Appl. Microbiol. Biotechnol. 98, 1771–1783. https://doi.org/10.1007/s00253-013-5452-2

Walczak, H., Iwai, K., Dikic, I., 2012. Generation and physiological roles of linear ubiquitin chains. BMC Biol. 10, 23. https://doi.org/10.1186/1741-7007-10-23

Walsh, G., 2018. Biopharmaceutical benchmarks 2018. Nat. Biotechnol. 36, 1136–1145. https://doi.org/10.1038/nbt.4305

Walter, P., Ron, D., 2011. The Unfolded Protein Response: From Stress Pathway to Homeostatic Regulation. Science 334, 1081–1086. https://doi.org/10.1126/science.1209038

Wang, H., Siemens, J., 2015. TRP ion channels in thermosensation, thermoregulation and metabolism. Temperature 2, 178–187.

Wang, K., Zhang, T., Chen, J., Liu, C., Tang, J., Xie, Q., 2018. The Effect of Culture Temperature on the Aggregation of Recombinant TNFR-Fc is Regulated by the PERK-eIF2a Pathway in CHO Cells. Protein Pept. Lett. 25, 570–579. https://doi.org/10.2174/0929866525666180530121317

Wang, S., Kaufman, R.J., 2012. The impact of the unfolded protein response on human disease. J. Cell Biol. 197, 857–867. https://doi.org/10.1083/jcb.201110131

Wang, T., Hebert, D.N., 2003. EDEM an ER quality control receptor. Nat. Struct. Mol. Biol. 10, 319–321. https://doi.org/10.1038/nsb0503-319

Weng, Z., Jin, J., Shao, C., Li, H., 2020. Reduction of charge variants by CHO cell culture process optimization. Cytotechnology 72, 259–269. https://doi.org/10.1007/s10616-020-00375-x

Wenzel, D.M., Lissounov, A., Brzovic, P.S., Klevit, R.E., 2011. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. Nature 474, 105–108.

Wheat, A., Yu, C., Wang, X., Burke, A.M., Chemmama, I.E., Kaake, R.M., Baker, P., Rychnovsky, S.D., Yang, J., Huang, L., 2021. Protein interaction landscapes revealed by advanced in vivo cross-linking–mass spectrometry. Proc. Natl. Acad. Sci. 118, e2023360118.

Wilkens, C.A., Altamirano, C., Gerdtzen, Z.P., 2011. Comparative metabolic analysis of lactate for CHO cells in glucose and galactose. Biotechnol. Bioprocess Eng. 16, 714–724.

Wilkinson, B., Gilbert, H.F., 2004. Protein disulfide isomerase. Biochim. Biophys. Acta BBA-Proteins Proteomics 1699, 35–44.

Wilson, S.R., Vehus, T., Berg, H.S., Lundanes, E., 2015. Nano-LC in proteomics: recent advances and approaches. Bioanalysis 7, 1799–1815.

Winter, D., Pipkorn, R., Lehmann, W.D., 2009. Separation of peptide isomers and conformers by ultra performance liquid chromatography. J. Sep. Sci. 32, 1111–1119.

Wojtkiewicz, M., Berg Luecke, L., Kelly, M.I., Gundry, R.L., 2021. Facile Preparation of Peptides for Mass Spectrometry Analysis in Bottom-Up Proteomics Workflows. Curr. Protoc. 1, e85. https://doi.org/10.1002/cpz1.85

Wu, X., Siggel, M., Ovchinnikov, S., Mi, W., Svetlov, V., Nudler, E., Liao, M., Hummer, G., Rapoport, T.A., 2020. Structural basis of ER-associated protein degradation mediated by the Hrd1 ubiquitin ligase complex. Science 368, eaaz2449. https://doi.org/10.1126/science.aaz2449

Xu, D., Wan, A., Zhang, J., Peng, L., Chen, Y., He, Y., Yang, J., Jin, J., 2017. Inhibition of the ubiquitin ligase activity improves the production of biologically active fusion protein HSA-HGF in Chinese hamster ovary cells. Bioengineered 8, 256–264.

Xu, G., Deglincerti, A., Paige, J.S., Jaffrey, S.R., 2014. Profiling lysine ubiquitination by selective enrichment of ubiquitin remnant-containing peptides. Methods Mol. Biol. Clifton NJ 1174, 57–71. https://doi.org/10.1007/978-1-4939-0944-5_4

Xu, H., Van Remmen, H., 2021. The Sarco Endoplasmic Reticulum Calcium ATPase (SERCA) pump: a potential target for intervention in aging and skeletal muscle pathologies. Skelet. Muscle 11, 25.

Xu, W.-J., Lin, Y., Mi, C.-L., Pang, J.-Y., Wang, T.-Y., 2023. Progress in fed-batch culture for recombinant protein production in CHO cells. Appl. Microbiol. Biotechnol. 107, 1063–1075. https://doi.org/10.1007/s00253-022-12342-x

Xu, X., Nagarajan, H., Lewis, N.E., 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, 735–741.

Yang, M., Butler, M., 2000a. Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation. Biotechnol. Bioeng. 68, 370–380.

Yang, M., Butler, M., 2000b. Effect of ammonia on the glycosylation of human recombinant erythropoietin in culture. Biotechnol. Prog. 16, 751–759. https://doi.org/10.1021/bp000090b

Yang, Q., Zhao, J., Chen, D., Wang, Y., 2021. E3 ubiquitin ligases: styles, structures and functions. Mol. Biomed. 2, 1–17. https://doi.org/10.1186/s43556-021-00043-2

Ye, Y., Rape, M., 2009. Building ubiquitin chains: E2 enzymes at work. Nat. Rev. Mol. Cell Biol. 10, 755–764. https://doi.org/10.1038/nrm2780

Yoon, S.K., Choi, S.L., Song, J.Y., Lee, G.M., 2005. Effect of culture pH on erythropoietin production by Chinese hamster ovary cells grown in suspension at 32.5 and 37.0°C. Biotechnol. Bioeng. 89, 345–356. https://doi.org/10.1002/bit.20353

Yoon, S.K., Kim, S.H., Lee, G.M., 2003. Effect of low culture temperature on specific productivity and transcription level of anti-4–1BB antibody in recombinant Chinese hamster ovary cells. Biotechnol. Prog. 19, 1383–1386.

Yoshida, H., Matsui, T., Yamamoto, A., Okada, T., Mori, K., 2001. XBP1 mRNA Is Induced by ATF6 and Spliced by IRE1 in Response to ER Stress to Produce a Highly Active Transcription Factor. Cell 107, 881–891. https://doi.org/10.1016/S0092-8674(01)00611-0

Young, J.D., 2013. Metabolic flux rewiring in mammalian cell cultures. Curr. Opin. Biotechnol. 24, 1108–1115. https://doi.org/10.1016/j.copbio.2013.04.016

Zhang, C., Fu, Y., Zheng, W., Chang, F., Shen, Y., Niu, J., Wang, Y., Ma, X., 2024. Enhancing the Antibody Production Efficiency of Chinese Hamster Ovary Cells through Improvement of Disulfide

Bond Folding Ability and Apoptosis Resistance. Cells 13, 1481. https://doi.org/10.3390/cells13171481

Zhang, M., Liu, L., Lin, X., Wang, Y., Li, Y., Guo, Q., Li, S., Sun, Y., Tao, X., Zhang, D., Lv, X., Zheng, L., Ge, L., 2020. A Translocation Pathway for Vesicle-Mediated Unconventional Protein Secretion. Cell 181, 637-652.e15. https://doi.org/10.1016/j.cell.2020.03.031

Zhang, Y., Baig, E., Williams, D.B., 2006. Functions of ERp57 in the Folding and Assembly of Major Histocompatibility Complex Class I Molecules *. J. Biol. Chem. 281, 14622–14631. https://doi.org/10.1074/jbc.M512073200

Zhao, J., Xia, C., Tang, Y., Wan, H., 2023. Role of PERK-mediated pathway in the effect of mild hypothermia after cerebral ischaemia/reperfusion. Eur. J. Clin. Invest. 53, e14040.

Zheng, N., Shabek, N., 2017. Ubiquitin Ligases: Structure, Function, and Regulation. Annu. Rev. Biochem. 86, 129–157. https://doi.org/10.1146/annurev-biochem-060815-014922

Zheng, N., Wang, P., Jeffrey, P.D., Pavletich, N.P., 2000. Structure of a c-Cbl-UbcH7 complex: RING domain function in ubiquitin-protein ligases. Cell 102, 533–539. https://doi.org/10.1016/s0092-8674(00)00057-x

Zhou, W., Zhu, P., Wang, J., Pascual, G., Ohgi, K.A., Lozach, J., Glass, C.K., Rosenfeld, M.G., 2008. Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation. Mol. Cell 29, 69–80. https://doi.org/10.1016/j.molcel.2007.11.002

Zhu, W., Smith, J.W., Huang, C.-M., 2010. Mass spectrometry-based label-free quantitative proteomics. BioMed Res. Int. 2010, 840518.

Zimmermann, R., 2016. Components and Mechanisms of Import, Modification, Folding, and Assembly of Immunoglobulins in the Endoplasmic Reticulum. J. Clin. Immunol. 36, 5–11. https://doi.org/10.1007/s10875-016-0250-0

Zimmermann, R., Eyrisch, S., Ahmad, M., Helms, V., 2011. Protein translocation across the ER membrane. Biochim. Biophys. Acta BBA-Biomembr. 1808, 912–924.

Zubarev, R.A., Makarov, A., 2013. Orbitrap mass spectrometry. ACS Publications.

Research Aims

The overall objective of this thesis is to characterise the changing proteomic landscape of CHO cells under ER stress. Through artificial stress induction and bioprocess-relevant conditions, the work explores the relationship between culture performance metrics: viability, growth, and productivity, with proteomic changes. Emphasis is placed on identifying the role of ubiquitin in the adaptation to ER stress, with the long-term goal of informing strategies for cell line engineering and process optimisation in biopharmaceutical production. This was achieved through the completion of the following specific aims:

- Induce ER stress in producer and non-producer CHO cells with thapsigargin, quantifying effects on viable cell density, culture percentage viability, and productivity.
- Validate ER stress induction by detecting established markers via western blotting.
- Apply the bioprocess-relevant stressors, temperature downshift and elevated culture ammonia, to explore the proteomic changes related to changes in the culture environment.
- Enrich ubiquitinated peptides by immuno-affinity precipitation to examine ubiquitin's regulatory role in the CHO Cell response to mild hypothermia.
- Perform LC-MS/MS on digested peptides, using Proteome Discoverer and Progenesis QI for quantitative proteomic profiling of differentially expressed proteins.
- Map proteomic changes to biological pathways using gene ontology and KEGG, focusing on "Protein Processing in the Endoplasmic Reticulum."
- Apply regression and machine learning methods to ER stress mass spectrometry data to identify ER stress biomarkers with potential as targets for future genetic engineering and the creation of ER stress monitoring assays.

Chapter 2: Paper 1- Label-Free Quantitative Proteomics Analysis of Producer and Non-producer Chinese Hamsters Ovary (CHO) Cells under ER Stress Conditions

Ryan, D., Sideri, C.-K., Henry, M., Efeoglu, E., & Meleady, P. (2023). Label-Free Quantitative Proteomics Analysis of Producer and Non-producer Chinese Hamsters Ovary (CHO) Cells under ER Stress Conditions. Current Research in Biotechnology, 100141. DOI-10.1016/j.crbiot.2023.100141.

This study uses quantitative label-free LC-MS/MS to investigate proteomic changes the ER stress response CHO cell lines. It focuses on understanding the molecular mechanisms vital to protein production and stress resilience in both recombinant protein-producing and non-producing cells. Thapsigargin, an inhibitor of the sarco-endoplasmic reticulum Ca^{2+}-ATPase (SERCA), was used to induce ER stress by depleting ER calcium stores, which disrupts the function of protein folding chaperones. Three CHO cell lines were analysed: a non-producer-CHO-K1, and two producer lines-CHO DP-12 (IgG)and CHO SK15-EPO (EPO).

The ER stress response represents a potential bottleneck to productivity of recombinant CHO cells in culture. Increased levels of ER stress can lead to cell death via apoptosis (S. Zhang et al., 2016), which can impact the product titre of a culture. In a comparison of high and low producing CHO cells, Prashad & Mehra, 2015, demonstrated that a high producing line had much higher induction of UPR related proteins, linking recombinant protein production to ER stress (Prashad & Mehra, 2015). Previous studies have shown how ER stress proteins interfere with product quality, linking the upregulation of ER stress-response proteins to disulfide bond reduction in mAbs, reducing product quality (Park et al., 2021), as well as ERAD-related ubiquitin ligases negatively affecting recombinant fusion-protein structure (Xu et al., 2017). A complete understanding of the biological processes involved in the ER stress response could be crucial to the optimisation of CHO genetic engineering, allowing for greater productivity, enhanced product quality and resilience to stress. The increased workload on the ER that is caused by recombinant protein production often involves the

induction of some ER stress/UPR responses (Du et al., 2013; Prasad and Mehra, 2015). While thapsigargin exposure does not create a direct mirror of recombinant protein production in CHO cells, it provides a benchmark for understanding the proteomic changes involved in CHO ER stress/UPR responses. It can be utilised as a baseline case for further stress studies, relating to more realistic bioprocessing stress conditions, as a case of the most extreme version of ER stress. 2000 nM thapsigargin was chosen as the experimental dose in this study in pursuit of this aim, of building a clear view of a strong ER stress response. This limits the interpretation of the results in this study directly to bioprocessing conditions but allows for a greater experimental understanding of ER stress. This greater clarity on the proteomic changes during periods of ER stress can be used as the foundation for future studies in adapting the stress response for increasing stress resilience.

To examine the cellular response to ER stress, label-free quantitative liquid chromatography-tandem mass spectrometry (LC-MS/MS) was employed. This high-resolution proteomic analysis enabled the identification of hundreds of proteins with significantly changed expression in response to thapsigargin treatment. The analysis revealed 607 differentially expressed proteins in CHO-K1, 322 in CHO DP-12, and 900 in CHO SK15-EPO, highlighting both shared and cell line-specific stress response mechanisms. These numbers are reflective of DE proteins found in other CHO proteomic studies using similar methods, with several hundred proteins typically identified (Xu et al., 2017, Pérez-Rodriguez et al., 2021)

Across all three cell lines, known markers of ER stress were observed. These included upregulation of protein disulphide isomerases (PDIs), chaperones, and enzymes involved in glycoprotein processing. ERAD pathways were activated, with elevated expression of components of the ubiquitin-proteasome system, particularly in the recombinant protein-producing cell lines. This enrichment suggests that producer cell lines possess an enhanced

capacity to identify, target and degrade misfolded proteins, possibly an adaptative response to the increased folding load created by recombinant gene expression.

Gene ontology analysis identified wider cellular pathways which are affected by thapsigargin exposure. These included RNA transport, aminoacyl-tRNA biosynthesis, and amino acid metabolism, indicating that the response to ER stress involves the elements from across the cellular environment. This widespread reprogramming of cellular processes reflects a coordinated effort to manage proteotoxic stress and maintain homeostasis under unfavourable conditions. Other studies looking at ER stress in CHO cells have identified changes in mitochondrial UPR components, highlighting the complexity of the stress response in these cells, and the need for further investigation of these mechanisms (Sulaj et al., 2024).

This study highlights key molecular targets that could be manipulated to enhance CHO cell performance in industrial bioproduction. Engineering efforts aimed at enhancing ER folding capacity and protein production could improve cell line robustness and productivity through upregulation of UPR mediators, ERAD components, and folding chaperones. The enhanced stress response capacity observed in the producer lines highlights the value of comparative analysis between producer and non-producer lines and identifying elements of the stress response which could be used to optimise biopharmaceutical manufacturing.

In this work, I was responsible for the culturing of CHO-K1 and CHO-SK15-EPO cells, including stressing the cells, viability and productivity assays and cell pellet collection. I prepared samples from these cell lines for proteomic analysis and performed the differential expression study, through MS analysis. In the manuscript preparation, I prepared the methodology section, and contributed extensively to the discussion of viability, productivity and DE results, in particular for the CHO-K1 and CHO-SK15-EPO cell lines.

Label-Free Quantitative Proteomics Analysis of Producer and Non-producer Chinese Hamsters Ovary (CHO) Cells under ER Stress Conditions

David Ryan[1,2,¶], Christiana-Kondylo Sideri[1,2,¶], Michael Henry[1], Esen Efeoglu[1], Paula Meleady[1,2,3]

[1]National Institute for Cellular Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland.

[2]School of Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland.

[3]SSPC, The Science Foundation Ireland Research Centre for Pharmaceuticals, Limerick, Ireland

*Corresponding author

Tel: +353-1-7005910; Email: paula.meleady@dcu.ie

¶ David Ryan and Christiana-Kondylo Sideri contributed equally to this work.

2.1 Abstract

Chinese Hamster Ovary (CHO) cells are the most commonly used host cell line used in the biopharmaceutical industry for the production of biotherapeutics. However, the lack of knowledge of endoplasmic reticulum (ER) stress mechanisms involved in protein production, folding and secretion remains a bottleneck preventing improved production titres and product quality. This study aimed to uncover differentially expressed proteins in ER stress pathways in a non-producer (CHO-K1) and two producer (CHO-DP12 and CHO-SK15-EPO) cell lines following exposure to the known stress inducer, thapsigargin using label-free quantitative LC-MS/MS proteomic analysis. This resulted in the identification of 607, 322 and 900 differentially expressed proteins in CHO-K1, CHO DP-12 and CHO SK15-EPO cells, respectively, compared to untreated controls. We identified key pathways related to protein folding, the unfolded protein response (UPR), ER-associated degradation (ERAD), RNA transport, aminoacyl-tRNA biosynthesis and amino acid biosynthesis following induction of ER stress conditions, with expression changes observed in several glycosyltransferases, protein disulfide isomerases and molecular chaperones. Proteins

targeting to the ubiquitin-ligase complex and ERAD were observed to be more enriched in the producer cells following induction of ER stress. This study extends existing knowledge of proteomic changes that govern the cellular stress response potentially identifying targets that can facilitate targeted cell line engineering to obtain production cell lines that demonstrate a higher stress tolerance and improved productivity.

Keywords: Chinese Hamster Ovary (CHO) cells, endoplasmic reticulum stress, unfolded protein response (UPR), endoplasmic reticulum associated degradation (ERAD), thapsigargin

2.2 Introduction

Chinese hamster ovary (CHO) cells are the most common mammalian host cell line for the industrial production of recombinant protein therapeutics, due to their ability to grow in suspension culture at high cell densities, to carry out correct protein folding and post-translational modifications and their adaptability to a variety of growth conditions (Tihanyi and Nyitray, 2020). Protein therapeutics, including monoclonal antibodies (Mabs) and recombinant proteins, remain the largest group of new products in development by the biopharmaceutical industry, and studies have been focusing on enhancing the productivity of biotherapeutics (Dahodwala and Sharfstein, 2017; Dhara et al., 2018; Fischer et al., 2015). Meeting the increasing demand for biopharmaceutical products remains a challenge and drives constant innovation in the manufacturing processes in order to establish a highly efficient process with a stable product and low production costs (Fischer et al., 2015; Tihanyi and Nyitray, 2020). With the purpose of facilitating targeted engineering strategies to further optimise current CHO cell systems, the understanding of their cellular and molecular conformation is imperative.

The endoplasmic reticulum (ER) is the largest organelle in the eukaryotic cell and it is responsible for protein synthesis and transport, protein folding, lipid and steroid synthesis,

carbohydrate metabolism and calcium storage (Schwarz and Blower, 2016). In the secretory pathway, all secretory proteins enter through the ER, where they fold into their native conformation, and if folded correctly they continue to the Golgi apparatus to be released to the cell membrane and/or extracellular space. Misfolded or incompletely folded proteins remain in the ER to be targeted for degradation or to complete the folding process (Ellgaard et al., 1999; Lin et al., 2008). The accumulation of misfolded proteins in the ER can be caused by multiple conditions that interfere with the ER quality control and disrupt ER homeostasis, leading to 'ER stress'. The unfolded protein response (UPR) maintains the protein-folding homeostasis within the ER, preventing the accumulation of folded proteins (Hetz et al., 2020). To support proper protein folding and prevent aggregation in the ER lumen, numerous ER-resident chaperones are involved, with BiP/GRP78 being the most abundant (Kaufman, 2002). Rough ER plays a crucial role in protein synthesis and aids distribution of the newly synthesized proteins to the cell membrane or extracellular region. The high recombinant protein synthesis capability of CHO cells can place extra pressure on the ER machinery combined with environmental factors such as pH and temperature fluctuations, prolonged culturing times and feeding frequencies in industrial settings (Li et al., 2022; Lin et al., 2008; Prashad and Mehra, 2015).

The production of recombinant proteins is a coordinated process of transcription, translation, folding, post-translational modification, and secretion. Each of these steps may act as a bottleneck in the production of correctly assembled proteins (Reinhart et al., 2014). In particular, cellular processes such as folding and degradation in the ER have attracted significant attention in the field of recombinant protein production. In mammalian production cell cultures, the expression of high levels of recombinant proteins is often linked to enhanced UPR levels to maintain cell viability and productivity (Du et al. 2013; Prasad and Mehra, 2015). During fed-batch culture, continuous high productivity and growth rate was found to increase the ER load on CHO cells, ultimately inducing UPR (Prashad and

Mehra, 2015). The activation of the UPR involves three UPR sensors, inositol-requiring enzyme 1 (IRE1), protein kinase R-like ER kinase (PERK) and activating transcription factor 6 (ATF6), which all play vital roles in returning protein homeostasis to levels seen in non-stressed cells (Read & Schröder, 2021). Artificial induction of the UPR, and by using both stable and transient transfection approaches, has been shown to improve the production of difficult-to-express (DTE) fusion proteins and Mabs (Johari et al., 2015; Latorre et al., 2023; Malm et al., 2022; Pieper et al., 2017; Pybus et al., 2014a). Controlling and fine-tuning the UPR to improve production processes is therefore a current challenge in CHO cell engineering (Latorre et al., 2023; Pieper et al., 2017). An example of this 'fine-tuning' has recently been shown where knocking out the PERK pathway, in conjunction with the deletion of components of the apoptosis pathway, i.e., Bax and Bak, resulted in increased cell viability and higher specific productivity in Mab producing CHO cells (Castellano et al., 2023). There is also evidence from several CHO cell line engineering studies that the ubiquitin proteasome system (UPS) pathway itself plays an important role in regulating growth, productivity and product quality of rCHO cells. Studies have shown altered levels of ubiquitin-proteasome related proteins as direct targets of miRNAs found to have phenotypic effects on growth (Costello et al. 2018) and productivity (Fischer et al. 2015). Although rCHO cells are the most commonly used host for the production of high cost recombinant biopharmaceuticals, ER stress mechanisms remain poorly understood creating a major bottleneck in improving the efficiency of production (Komatsu et al., 2020a; Kyeong and Lee, 2022; Pérez-Rodriguez et al., 2021; Saghaleyni et al., 2022).

In this study, the role of ER stress on a non-producer CHO cell line, CHO-K1, and two producer cell lines, an IgG producer CHO DP-12 and EPO producer CHO SK15-EPO were investigated using state-of-the-art mass spectrometry. In order to induce ER stress on these three CHO cell lines, thapsigargin, one of the most common artificial ER stressors, was used. Thapsigargin is a non-competitive inhibitor of the sarco(endo)plasmic reticulum calcium

ATPase (SERCA) (Lytton et al., 1991), and was originally isolated as a sesquiterpene lactone from the plant *Thapsia garganica* (Rasmussen et al., 1978). This inhibition of the ER Ca^{2+}-dependent ATPase depletes Ca+ storage in the ER, which, in turn, decreases the activity of Ca-dependent chaperones leading to an increase in unfolded proteins, thereby causing ER stress (Kaufman, 2002; Lindner et al., 2020; Oslowski and Urano, 2011; Roy et al., 2017). To obtain a pronounced ER stress response, high levels of ER stressor and prolonged exposure times (up to 72 hours) were applied. Differential expression (DE) analysis of the whole cell proteome (Lakshmanan et al., 2019; Park et al., 2017; Xu et al., 2017a) following exposure to thapsigargin was carried out, followed by Gene Ontology (GO) analysis on the DE data to determine enriched biological processes, molecular functions as well as KEGG pathways. With this study, we aim to improve our understanding of ER stress mechanisms in CHO cells, and to uncover potential targets for cell line engineering of CHO cells to increase the productivity of biopharmaceuticals.

2.3 Materials and Methods

2.3.1 Cell Culture

CHO-K1 (non-producer, ATCC CCL-61), CHO DP-12 (a recombinant human anti-IL-8 producer, ATCC CRL-12445 clone #1934) and CHO SK15-EPO (an in-house CHO-K1 derived cell line, expressing recombinant human erythropoietin in a pcDNA3.1 vector (Invitrogen) modified with puromycin resistance as selection system) (Costello et al., 2019) were grown in BalanCD CHO Growth A (FujiFilm, NL) supplemented with 2 mM L-glutamine (Sigma Aldrich, IE) and maintained in a Climo-Shaker ISF1-X (Kühner, CH) at 37°C under 5% CO_2 and 80% humidity with constant shaking (170 rpm). Cells were subcultured every 4 days. Every 2-3 weeks, the anti-IL-8 and EPO producer CHO cells were pulsed with 400 nM methotrexate (Sigma-Aldrich, IE) and 10 mg/mL puromycin (Sigma-Aldrich, IE), respectively. The cells were routinely tested in-house for *Mycoplasma* contamination and tested negative.

2.3.2 Exposure to ER Stressor Thapsigargin

Cells were seeded at $8x10^5$ cells per mL in a 5 mL final volume using 50 mL TPP® TubeSpin Bioreactor tubes (Merck, IE). The cells were incubated for 24 h, and total cell count and viability were determined via Guava flow cytometer using ViaCount™ reagent (LuminexCorp, UK), giving the zero-hour time point results (where zero-hour is the time just before exposure to ER stressor). Thapsigargin (Sigma Aldrich, IE), resuspended in dimethyl sulfoxide (DMSO) (BioSciences, IE), was added to the bioreactor tubes at 2 µM concentration. The concentration was chosen based on two criteria, adequate induction of ER stress and minimal cell death over a period of 72 hours **(Supplementary Information Figure 1)**. Following thapsigargin exposure, cellular viability, cell death, and apoptosis rates were measured at 24, 48 and 72 hours. Cells were also exposed to 10 µL DMSO which corresponds to the amount of DMSO in thapsigargin-containing samples, and no effect was observed in viability (see Supplementary Fig. 2). Pellets from the cells grown with and without thapsigargin exposure were collected at 0, 24, 48 and 72-hours by centrifugation and washed three times with ice-cold PBS. The frozen cell pellets were kept at -80 °C for further use.

2.3.3 Protein Extraction and Western Blotting

Three hundred microliters of urea lysis buffer with 1X protease inhibitor (Thermo Scientific, IE) were added onto cell pellets, and short pulses of sonication were employed for mechanical disruption. The disrupted cell pellets were placed on a rotator for 1 h to achieve complete lysis. Following 1 h incubation, the lysate was centrifuged for 15 min at 220 ×g and supernatant containing extracted proteins was collected. Protein levels in the supernatant were quantified by Bradford assay (BioRad, IE). Twenty microgrammes of protein sample was mixed with Laemmli buffer (1:1 ratio) and incubated at 70°C for 10 min. Protein samples were loaded onto a 12% SDS-polyacrylamide gel (Invitrogen, IE) and run at 130V for 60 min. One microlitre of Chameleon® Duo Pre-stained Protein Ladder (Li-Cor, UK)

was used in all blots. The Pierce Power Blot Cassette and Power Blotter Select Transfer Stacks (Invitrogen, IE) were used for the transfer of proteins from gel to membrane. The blots were blocked for 1 h using 10 mL of 1X Pierce Clear Milk Blocking Buffer (Biosciences, IE) and incubated overnight at 4°C with primary antibody for ATF6 (Cell Signalling, IE), BIP (Cell Signalling, IE), Ubiquitin (Cell Signalling, IE) or β-actin (Cell Signalling, IE). All primary antibodies were used within the manufacturers' recommended concentrations. Blots were then incubated with relevant secondary antibodies (Li-Cor, 1:15000 dilution) for 1 h and then rinsed with 1x TBS Tween (Thermo Fisher Scientific, IE). Membranes were imaged using the Odyssey Scanner (Li-Cor, UK).

2.3.4 ELISA

EPO titre was quantified by enzyme linked immunosorbent assay (ELISA) of the cell culture supernatant. EPO supernatant was harvested by centrifugation of the cell suspension at 91 ×g for 5 min. Nunc-Immuno™ MicroWell™ 96 well solid plates (Sigma Aldrich, IE) were coated with 100 µL of capture antibody MAB-287 (R&D Systems Eire, UK) at a concentration of 2 µL/mL in coating buffer (C3041-50CAP) (Sigma-Aldrich, IE) overnight at 4 °C. Wells were washed three times with 100 µL/well of ELISA wash buffer (Sigma Aldrich, IE). 200 µL of Blocking Buffer (Sigma-Aldrich, IE) was added to each well before being incubated for 2 h in the dark at room temperature. Standards of recombinant human EPO (Merck-Millipore, IE) were diluted to range from 0 to 8000 pg/mL in sample dilution buffer (Sigma-Aldrich, IE) supplemented with 0.05% Tween 20. Samples were diluted to fall within the range of the standard curve and 100 µL of sample was incubated on the plate at 37 °C for 1 h at 170 RPM. Detection was achieved using Rabbit Immunoglobulin-HRP (Agilent, IE) diluted 1:2000 with 100 µL added to each well and incubated for 1 h. 100 µL of development solution; 1.2 mL of 3,3′,5,5′-Tetramethylbenzidine at 1 mg/mL in dimethyl sulfoxide, 2.4 µL 30% (v/v) hydrogen peroxide and 10.8 mL Phosphate-Citrate buffer (Merck, IE), was added to each well and incubated for 30 min. The reaction was stopped by

adding 100 µL 0.18 M H_2SO_4 to each well. The protein concentration was determined by reading the plate on a Multi-scan Go (Thermo Fisher Scientific, IE).

IgG ELISA assay was performed using the Human IgG ELISA kit (Merck, IE) with supernatant harvested from control and treated cultures for each time point (24h, 48h, 72h) by centrifugation of the cell suspension at 91 ×g for 5 min. Following the Sandwich Assay Procedure described by the manufacturer, 100 µL of each standard and sample were added into appropriate wells and incubated overnight at 4°C with gentle shaking. This solution was then removed, and the wells were washed (300 µL wash buffer per well) four times before incubation with the detection antibody (100 µL per well) for 1 h at room temperature with gentle shaking. The antibody was discarded, and wells washed four times followed by incubation with Streptavidin (100 µL per well) for 45 min at room temperature with gentle shaking. Again, wells were washed four times, followed by incubation with the substrate (100 µL per well) for 30 min at room temperature with gentle shaking in the dark. The stop solution was added (50 µL per well) and the absorbance was read at 450 nm immediately.

Specific cell productivity (Qp), which measures the rate of protein expression per cell and time unit, is calculated as pg/cell/day for both EPO and IgG producer cell lines as described in Supplementary information 1.1. Qp is an average value for the amount of the specific protein produced by an individual cell across the time in culture. Due to the decline in CHO-DP12 live cells at 72 h, we cannot calculate Qp at 72 h for this cell line, hence the value is calculated at 48 h. Where product titre gives an indication of the experimental effect on the culture as a whole, Qp allows for insight into the effect on individual cell productivity.

2.3.5 Protein extraction and preparation for Mass Spectrometry

Cells were collected and pelleted at $3x10^6$ cells in each pellet and stored at -80 °C. Each protein sample was digested and prepared for LC-MS/MS using the PreOmics iST kit following manufacturer's instructions (PreOmics, DE). The kit includes lyse buffer,

digestion mix, stop and wash solutions, elution buffer and LC-LOAD buffer. Briefly, the lyse buffer was used to lyse the cell pellets at 95°C for 10 min at 1000 RPM. Following this, the Trypsin/LysC mix was added, and the sample incubated at 37°C for 3 h at 500 RPM to allow for proteolytic digestion. Once digestion was stopped, the peptide samples were washed and filtered through the cartridge. Once filtered, the peptide samples were placed in the vacuum evaporator at 45°C until fully dry. Samples were then stored at -80°C.

2.3.6 LC-MS/MS and label-free quantitative differential analysis

For LC-MS/MS profiling, an UltiMate 3000 nanoRSLC (Thermo Scientific, UK) system interfaced with an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific, UK) was used. Using nano-flow HPLC separation, 1 µg of tryptic peptide from each sample was loaded onto the trapping column (Acclaim PepMap100, C18, 300 µm × 5 mm, Thermo Fisher Scientific) at a flow rate of 25 µL/ min with 2% (v/v) acetonitrile (ACN), 0.1% (v/v) trifluoroacetic acid (TFA) for a 3 min desalting and concentration step. Each sample was then resolved onto an analytical column (Acclaim PepMap 100, 75 µm × 50 cm, 3 µm bead diameter column, Thermo Fisher Scientific, UK) into the mass spectrometer. A binary gradient of: solvent A (0.1% (v/v) formic acid in LC-MS grade water) and solvent B (80% (v/v) ACN, 0.08% (v/v) formic acid in LC-MS grade water) using 2-32% B for 75 min, 32-90% B in 5 min and holding at 90% for 5 min at a flow rate of 300 nL/min was used to elute peptides.

For MS and MS/MS analysis, a temperature of 320 °C and a voltage of 2.0 kV was used for peptide ionization. The Orbitrap mass analyser with a resolution of 120,000 (at m/z 200), a maximum injection time of 50 ms and an automatic gain control (AGC) value of 4×10^5 was used to perform full MS scans for peptide mass information. Data-dependent acquisition was performed using a full scan range of 380-1500 m/z. A top-speed MS/MS acquisition algorithm was used to determine the number of selected precursor ions for fragmentation. A

dynamic exclusion was applied to analysed peptides after 60 s to avoid repeat fragmentation of the same peptide and only peptides with a charge state between 2+ and 7+ were analysed. Peptides were fragmented using higher energy collision-induced dissociation (HCD) with a normalised collision energy of 28%. The resulting MS/MS fragment ions were measured in the low-resolution linear ion trap.

Peptides were identified using Proteome Discoverer 2.4 (Thermo Scientific, UK) against a *Cricetulus griseus* UniProt database (downloaded September 2021) using the SEQUEST HT algorithm. A 20-ppm mass tolerance and a 0.6 Da fragment mass tolerance were selected for all searches in the Proteome discoverer. Methionine oxidation was set as a dynamic modification, whereas carbamidomethylation was set as a static modification. Trypsin was selected as a digestion enzyme and a maximum of two missed cleavages was allowed. FDR was set as < 5% using Percolator. Differential expression analysis is carried out using Progenesis QI for Proteomics (NonLinear Dynamics, a Waters Company, UK). Quantitative label-free analysis was performed using Progenesis QI for Proteomics to compare the thapsigargin-exposed cells to untreated (control) cells at the 24, 48 and 72 h time points for each of the three cell lines. Only peptide ions with charge state +2 and +3 were allowed and normalised. Peptides with a one-way ANOVA p-value ≤.05 between experimental groups were exported and identified using Proteome Discoverer. Protein identifications were imported into Progenesis and differentially expressed proteins were accepted if they passed the following criteria: (i) a protein level one-way ANOVA p-value ≤.05 between experimental groups, (ii) a minimum of 1 peptide contributing to protein identification and (iii) a minimum fold change in abundance of ± 1.5 fold between untreated and thapsigargin-treated cells.

2.3.7 KEGG pathway and STRING association analysis

Proteins which were differentially expressed between control and test samples were gathered and uploaded into STRING (https://www.string-db.org) (Szklarczyk et al., 2021), which was

used to group the proteins based on protein-protein interactions and to provide groupings within KEGG biological pathways (https://www.genome.jp/kegg) (Kanehisa et al., 2023).

2.3.8 Statistical Analysis

Statistical analysis of differences was performed in OriginPro 2022b SR1 9.9.5.171 (Academic) by one-way analysis of variance (ANOVA). The null hypothesis was rejected at the 0.05 level (*). Error bars in all figures indicate standard deviation of the arithmetic mean of three independent experiments. Viability was expressed as fold-change when compared to values at the zero-hour condition.

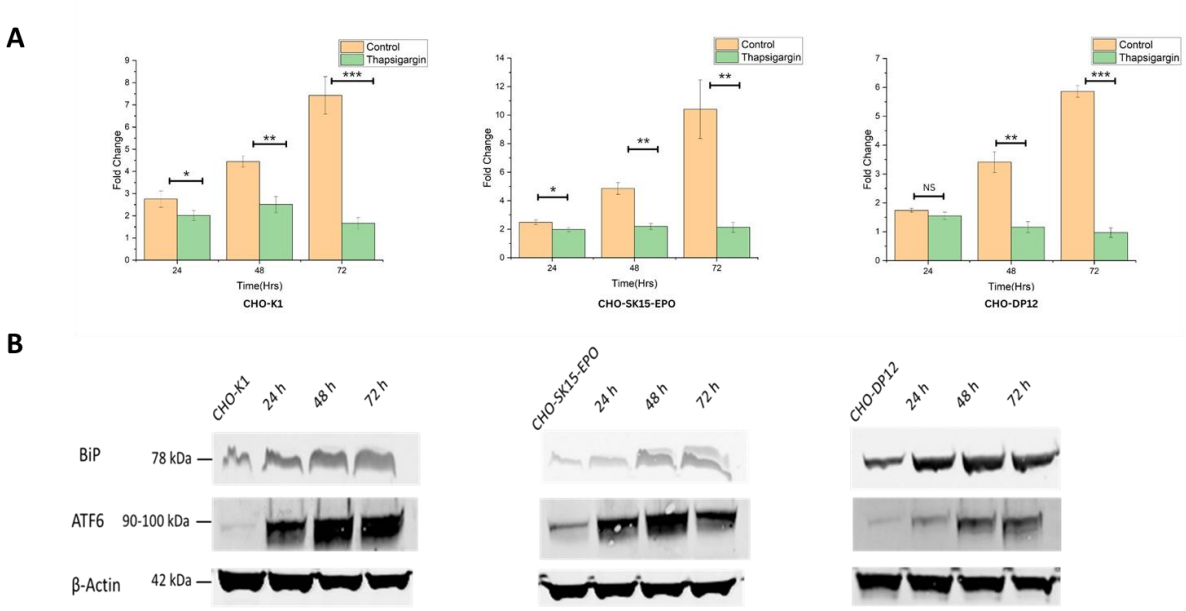
2.4 Results and Discussion

2.4.1 CHO cell survival, productivity and expression of stress related proteins under ER stress conditions

In this study, non-producer (CHO-K1) and producer (CHO-SK15 EPO and CHO DP-12) CHO cells were investigated under ER stress conditions. Figure 1(A) shows fold change in viable cell count as a function of time (24 -72 h), using the 0 h measurements as a baseline, following exposure to 2 μM Thapsigargin. With the exception of the 24 h time point from CHO-DP12 cells treated with thapsigargin, all treated cells showed significant decreases in growth compared to untreated cells.

To determine if ER stress was induced following treatment with thapsigargin, western blotting was carried out using antibodies for the known ER stress marker proteins, BiP and ATF6, with β-actin as a loading control. Differential expression proteomic data confirmed that there was no change in the expression level of β-actin. All three cell lines showed increased expression of BiP and ATF6 following thapsigargin treatment, as shown in Figure 1(B).

The specific cell productivity (Qp) of the two producer cell lines was decreased following exposure to thapsigargin (Figure 1(C)). CHO DP-12 Qp was measured at the 48h time point as the number of viable cells at the 72h time point was less than the 0h count, so Qp was not calculable for 72h. Although the measured Qp values were lower compared to industrial standard CHO cells, both the CHO DP-12 and CHO SK15-EPO cell lines still have value as model cell lines to understand the mechanisms of ER stress formation. The Qp value for unexposed CHO SK15-EPO is in line with the Qp values obtained by the research group responsible for creation of this line (Costello et al., 2019).



C

	Qp Control	Qp Thapsigargin Exposed
CHO SK15-EPO (72 hours)	1.59	1.20
CHO DP-12 (48 hours)	0.53	0.18
Qp=pg/cell/day		

Figure 1. Viable cell number, stress indicators and productivity in Thapsigargin exposed CHO cells. Fold change in viable cell number based on zero-hour measurement of CHO-K1, CHO SK15-EPO, CHO DP-12 including one-way ANOVA p-values comparing fold change means between treated and untreated conditions at each time point (A). *represents significance at the 0.05 level, ** at 0.01 and *** at 0.001. Western blotting of known ER stress antibodies BiP and ATF6 as well as the loading control β-Actin (B).

Comparison of specific cell productivity (Qp) of the thapsigargin-treated versus untreated cells in the two producer cell lines (C).

2.4.2 Proteomic investigation of producer and non-producer CHO cells under ER stress conditions

Differential expression analysis of the untreated and thapsigargin-treated CHO cell lines (at 24, 48 and 72 h) showed 607, 900 and 322 proteins being differentially expressed for CHO-K1, CHO SK15-EPO and CHO DP-12 cells, respectively (ANOVA $p<0.05$, minimum of 2 unique peptides). Table 1 shows the number of upregulated and downregulated proteins for each CHO cell line for each chosen time point. The number of downregulated proteins was observed to increase with increasing exposure time for CHO-K1 cells, whereas the CHO DP-12 and CHO SK15-EPO cell lines did not follow a trend (see Supplementary Table 1, 2 and 3 for DE lists for each cell line).

Table 1. Number of differentially expressed proteins upon thapsigargin exposure in producer and non-producer CHO cell lines.

Exposure Time	**24h**		**48h**		**72h**	
Cell line/Expression level	**UpR**	**DownR**	**UpR**	**DownR**	**UpR**	**DownR**
CHO-K1	352	255	314	293	297	310
CHO DP-12	139	183	139	183	139	183
CHO SK15-EPO	409	491	379	521	417	483

STRING gene ontology analysis of the differentially expressed proteome of the non-producer, EPO producer and IgG producer CHO cells showed significant common

enrichments in KEGG Pathways such as Carbon metabolism (cge01200), Protein processing in endoplasmic reticulum (cge04141), Citrate cycle (TCA cycle) (cge00020), RNA transport (cge03013), Aminoacyl-tRNA biosynthesis (cge00970), Spliceosome (cge03040) and Biosynthesis of amino acids (cge01230) (Table 2).

Table 2. Gene ontology analysis of differential expression data and enriched KEGG pathways common for CHO-K1, CHO DP-12 and CHO SK-15 EPO cell lines.

Enriched pathways	**CHO-K1**		**CHO SK15-EPO**		**CHO DP-12**	
	Strength	**FDR**	**Strength**	**FDR**	**Strength**	**FDR**
Protein processing in endoplasmic reticulum	0.64	$6.03x10^{-10}$	0.56	$7.84x10^{-9}$	0.66	$7.25x10^{-7}$
Aminoacyl-tRNA biosynthesis	0.79	$3.55x10^{-5}$	0.7	$8.08x10^{-5}$	0.79	0.0027
RNA transport	0.76	$4.40x10^{-17}$	0.6	$9.99x10^{-11}$	0.76	$1.44x10^{-10}$
Biosynthesis of amino acids	0.67	$1.33x10^{-5}$	0.57	$8.08x10^{-5}$	0.66	0.0017
Valine, leucine and isoleucine degradation	0.65	0.0014	0.82	$9.74x10^{-9}$	0.65	0.0274
Cysteine and methionine metabolism	0.71	0.00025	0.55	0.0049	0.79	0.0013
Glutathione metabolism	0.65	0.00013	0.58	0.00015	0.72	0.00079
Ribosome biogenesis in eukaryotes	0.62	$9.00x10^{-5}$	0.66	$2.08x10^{-7}$	0.85	$1.15x10^{-7}$
Pyrimidine metabolism	0.57	0.0127	0.49	0.0134	0.66	0.0261
Fatty acid metabolism	0.68	0.00022	0.68	$1.97x10^{-5}$	0.62	0.0369
Fatty acid degradation	0.7	0.00063	0.82	$1.65x10^{-7}$	0.71	0.0178
Lysosome	0.49	0.00069	0.48	$8.08x10^{-5}$	0.64	$7.71x10^{-5}$
Spliceosome	0.84	$3.64x10^{-18}$	0.72	$3.92x10^{-14}$	0.66	$2.64x10^{-5}$
Pyruvate metabolism	0.76	0.00022	0.75	$3.67x10^{-5}$	0.71	0.0325
Carbon metabolism	0.8	$3.32x10^{-14}$	0.71	$1.54x10^{-12}$	0.63	0.00034
Galactose metabolism	0.62	0.0363	0.62	0.0083	0.74	0.0478
Drug metabolism	0.42	0.0358	0.46	0.003	0.66	0.0011

For in-depth proteomic investigation of the mechanisms involved in ER stress in producer and non-producer CHO cells, the ‘protein processing in ER’ pathway was investigated further by also including proteins identified with one unique peptide. Table 1 shows the

comparison of the number of proteins identified from the network of proteins related to 'protein processing in ER' for non-producer, EPO producer and IgG producer CHO cells as well as their strength and FDR value of the KEGG pathway enrichment. (Supplementary Tables 4, 5 and 6). 'Count in network' represents the number of identified proteins based on MS analysis of CHO samples and their thapsigargin exposed counterparts and compares this to the total number of proteins which have been determined to be involved in the pathway of interest based on active interaction sources such as databases, text mining and experimental data. Strength indicates how large the enrichment effect is by calculation of log10(observed/expected). False discovery rate (FDR) shows how significant enrichment is using the Benjamini-Hochberg procedure.

Table 3. Comparison of the number of proteins which contributes to the enrichment of the KEGG Pathway: 'Processing in Endoplasmic Reticulum' (cge04141).

GO Analysis Results			
	CHOK1	**CHO-DP12**	**CHO-SK15-EPO**
Count in	35/161	23/161	39/161
Strength	0.64	0.66	0.56
FDR	6.03×10^{-10}	7.25×10^{-7}	7.84×10^{-9}
Identified DE proteins			
CHO-K1	PDIA3, CANX, STUB1, PDIA6, GANAB, PLAA, SEC31A, HERPUD1, SAR1A, SEC24D, UBE2D3, SEC61B, HSP90B1, DNAJC3, SEC24C, RPN1, DAD1, STT3A, NSFL1C, PRKCSH, PDIA4, DNAJA1, ATF6, MOGS, DDOST, SSR4, VCP, UGGT1, SEC13, ERP29, CUL1, EIF2S1, DNAJB11, LMAN2. CKAP4		
CHO DP-12	PDIA3, CANX, PDIA6, GANAB, DNAJC10, PLAA, SAR1A, HSPA1L, HSP90B1, RPN1, NSFL1C, PRKCSH, PDIA4, ATF6, DDOST, UGGT1, ERO1LB, ERP29, EIF2S1, DNAJB11, LMAN2, ERLEC1, CKAP4		
CHO SK15-EPO	PDIA3, CANX, STUB1, STT3B, PDIA6, GANAB, DNAJC10, PLAA, SEC31A, SKP1, SAR1A, HSPA2, UFD1L, HSP90B1, DNAJC3, UBE4B, SEC61A1, RPN1, RPN2, DAD1, STT3A, NSFL1C, PRKCSH, PDIA4, UBE2D4, DNAJA1, ATF6, DDOST, SSR4, VCP, UGGT1, ERO1LB, ERP29, CUL1, EIF2S1, DNAJB11, UBXN4, LMAN2, CKAP4		

ER and lumenal chaperones aid the folding of newly synthesized polypeptide chains and take part in the fate of folded proteins by retaining them in the ER, sending them to the Golgi apparatus to be secreted or placed in the cell membrane, or directing them to proteasomes for degradation. Briefly, newly synthesized polypeptide chains are recognised on the ER membrane via their signal peptide and are taken up in the ER lumen via translocons. Following cleavage of the signal peptide, the polypeptide chain is positioned for folding with the help of lumenal chaperones and undergoes glycosylation (Cooper, 2000). Correctly folded proteins are packed and transported into small vesicles which shuttle them to the Golgi apparatus. Misfolded proteins are retained in the ER lumen by BiP and re-glycosylated (Gething, 1999). If re-glycosylation is unsuccessful after a couple of rounds, they are

classified as terminally misfolded and directed to proteasomes for ubiquitin-mediated degradation (Ellgaard and Helenius, 2003; Ferris et al., 2014). The accumulation of misfolded proteins in the ER lumen results in ER stress and triggers activation of the UPR to recover from stress. If the stress conditions are severe, UPR mechanism cannot bring back the normal cellular function and cells go to an apoptotic state (Hetz, 2012). As illustrated in the String Gene Ontology database, this process, 'Protein Processing in Endoplasmic Reticulum' pathway, includes 116 proteins annotated to it (Supplementary Figure 1, https://www.kegg.jp/kegg-bin/show_pathway?cge04141).

Our study showed that 35, 23 and 39 proteins associated with protein processing in ER are upregulated or downregulated upon thapsigargin exposure in CHO-K1 cells, CHO-DP12 and CHO-SK15-EPO cells, respectively (Table 2, Supplementary Tables 4, 5 and 6). For all the CHO cell lines, proteins associated with glycosylation, UPR and ERAD pathways are observed to be significantly enriched. The network of differentially expressed proteins involved in protein processing in the ER and their associated fold-changes showed that the most significant changes are observed in the expression levels of ATF6, GANAB, UGGT1, PRKCSH, ERP29, DDOST, CKAP4 and LMAN2 (> 2.5 fold change for each cell line compared to untreated cells), and these proteins were shared between producer and non-producer cell lines **(Figure 3)**. Even though expression levels differed between the cell lines, ATF6 was highly upregulated in all cell lines, whereas GANAB, UGGT1, PRKCSH, ERP29, DDOST, CKAP4 and LMAN2 were downregulated significantly. All downregulated differentially expressed proteins are associated with protein transport into the ER lumen, folding and N-glycosylation. The upregulated ATF6 levels indicate the activation of UPR in thapsigargin-exposed CHO cells.

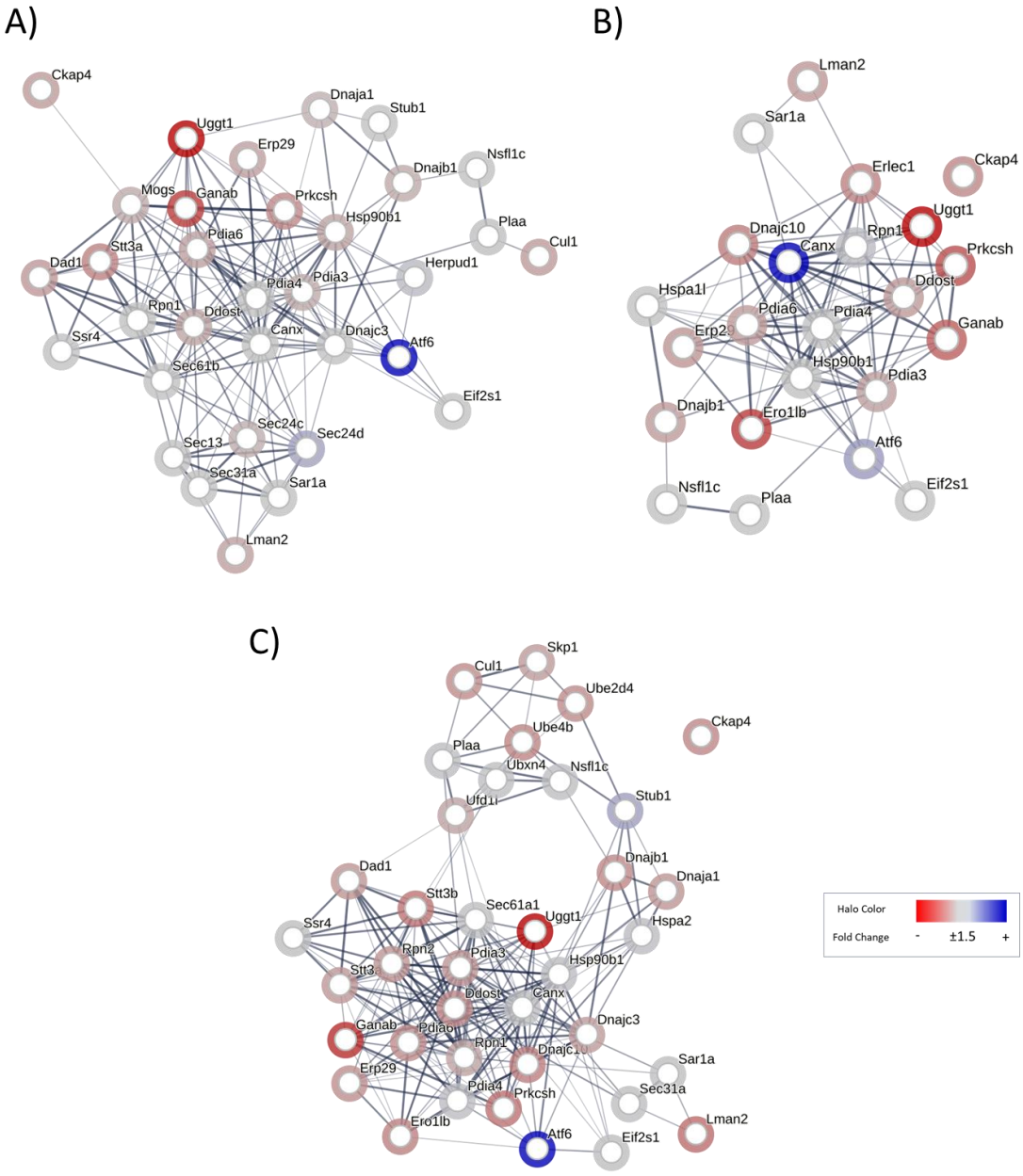


Figure 2. Protein-protein association networks in CHO cells experiencing ER stress. Network of proteins involved in protein processing in ER and related fold changes: CHO-K1 (A), CHO DP-12 (B), CHO SK15-EPO (C). From grey to blue halo around the protein node indicates the upregulation of protein with at least 1.5-fold change in its expression with an increase of upregulation with intensifying blue colour, whereas an intensifying red halo indicates higher degrees of downregulation. Protein-protein associations are indicated with grey lines with minimum of 0.4 to maximum of 0.9 confidence.

In total, 19 of these network proteins are identified to be shared among non-producer and producer cell lines. In addition to the above mentioned 19 proteins within all cell lines, there were no differentially expressed proteins shared between CHO-K1 and CHO DP-12 cells. However, between CHO-K1 and CHO SK15-EPO cells, nine proteins, DAD1, SEC31A, SSR4, DNAJA1 CUL1, DNAJC3, VCP, STUB1 and STT3A were identified to be shared (not identified to be differentially expressed in CHO DP-12 cells). STT3A, SSR4 and DAD1 form the part of oligosaccharyltransferase complex which conducts the transfer of nascent proteins from ER membrane bound ribosomes to ER lumen. The remaining five proteins are involved in protein folding, control of folding efficiency and ubiquitin-ligase binding. The observation of additional differentially expressed shared proteins between CHO-K1 and CHO SK15-EPO, but not with CHO DP-12, can be attributed to variation of ER stress mechanisms even between the producer cell lines. The type of recombinant protein, complexity of the protein, and production amounts can lead to variation upon ER stress response mechanisms. CHO DP-12 cells stably produce recombinant human IgG1 antibodies against IL-8. CHO SK15-EPO cells produce recombinant human erythropoietin (EPO). Human IgG1 is composed of 4 polypeptide chains (2 identical heavy chains and 2 light chains) and has a molecular weight of 150 kDa, whereas human EPO is a small protein, formed of 4 α-helices, two β-sheets and has a molecular weight of 30.4 kDa. The changes in molecular weight and amount of folding to establish a functional protein can burden host cells, making CHO DP-12 cells adapt different mechanisms to cope with the misfolded protein accumulation compared to CHO SK15-EPO cells. The similarity between the ER stress response mechanism between CHO-K1 and CHO SK15-EPO could also be related to lineage of these cell lines. CHO SK15-EPO is an in-house derived cell line from CHO-K1 lineage via insertion of rEPO expressing pcDNA3.1 vector.

The two producer cell lines, CHO SK15-EPO and CHO DP-12 were identified to share two differentially expressed proteins, ERO1β and DNAJC10 (HSP40), and these proteins were absent in the differential expression list of non-producer CHO-K1 cells. ERO1β is an isoform of the oxidoreductase ERO1, and is involved in the formation of native disulphide bonds that help maintain optimal redox conditions, which is upregulated in the UPR (Hansen et al., 2014). DNAJC10 is an endoplasmic reticulum disulphide reductase involved both in the correct folding of proteins and degradation of misfolded proteins that aids protein folding by catalysing the removal of non-native disulphide bonds formed during the folding of proteins (Cunnea et al., 2003). The significant changes in ER associated degradation proteins shared between producer cells without showing changes in expression in non-producer CHO cells possibly indicates the more regulated ER associated degradation is in producer cells upon ER stress.

The remaining differentially expressed proteins are either shared between two of the cell lines or observed to be unique to one of the cell lines. Seven proteins (HERPUD1, SEC13, SEC24C, SEC24D, UBE2D3, MOGS and SEC61B) were identified to be specific to changes in the cellular proteome of CHO-K1 cells upon thapsigargin exposure. Only two of these proteins, ERLEC1 (XTP3B) and HSPA1L (HSP70), were identified to be specific to changes in the proteome upon thapsigargin exposure in CHO DP-12 cells. In total, nine proteins (UBXN4, RPN2, UFD11, HSPA2, UBE4B, SEC61A1, STT3B, UBE2D4 and SKP1) were identified to be unique to changes in the proteome upon ER stress in CHO SK15-EPO cells. Similar to the IgG producer CHO cell line, proteins identified to be unique to the EPO producer cells are determined to play a role in ubiquitin-mediated proteolysis. The differences in biological function of the identified unique proteins indicates activation of ubiquitin mediated proteolysis and the ERAD pathway in producer cells, whereas non-producer CHO-K1 cells potentially use protein exit/transport mechanism as a way to manage ER stress. This also correlates with significant changes observed in the expression levels of

ERO1β and DNAJC10 which play an active role in the ERAD pathway and are shared between producer CHO cells.

In order to investigate in detail, the mechanisms of ER stress, UPR and ERAD activation, and the role of differentially expressed proteins in these mechanisms in producer and non-producer cells, the 'Protein processing in the ER' (cge04141) pathway is simplified as shown as Figure 3 which includes only the differentially expressed proteins identified in our study. Briefly, the pathway is divided into 3 sections and differentially expressed proteins and their roles are discussed. The first section which is indicated with the orange colour represents the protein processing starting with transport of nascent proteins into ER, folding, N-glycosylation, folding checkpoints, and secretory pathway. As misfolded proteins accumulate in ER lumen, this activates the ER stress response, and the identified proteins which are involved in this process are encapsulated in a yellow box (Figure 3). Finally, differentially expressed proteins which are involved in the activation of ERAD are indicated within the blue box. Each section of the 'Protein processing in ER' pathway as outlined in Figure 3 is investigated in the following sections in order to improve our understanding of the ER stress mechanism of CHO cells and to identify common and differing pathways between producer and non-producer cell lines as well as a comparison of IgG1 and EPO producing CHO cells. Please note that Figure 3 includes only the proteins from this pathway that were observed in this study, and these are the proteins focused on in this analysis. In total 1323, 834 and 1740 differentially expressed proteins were observed in CHO-K1, CHO DP-12 and CHO SK15-EPO cells, respectively.

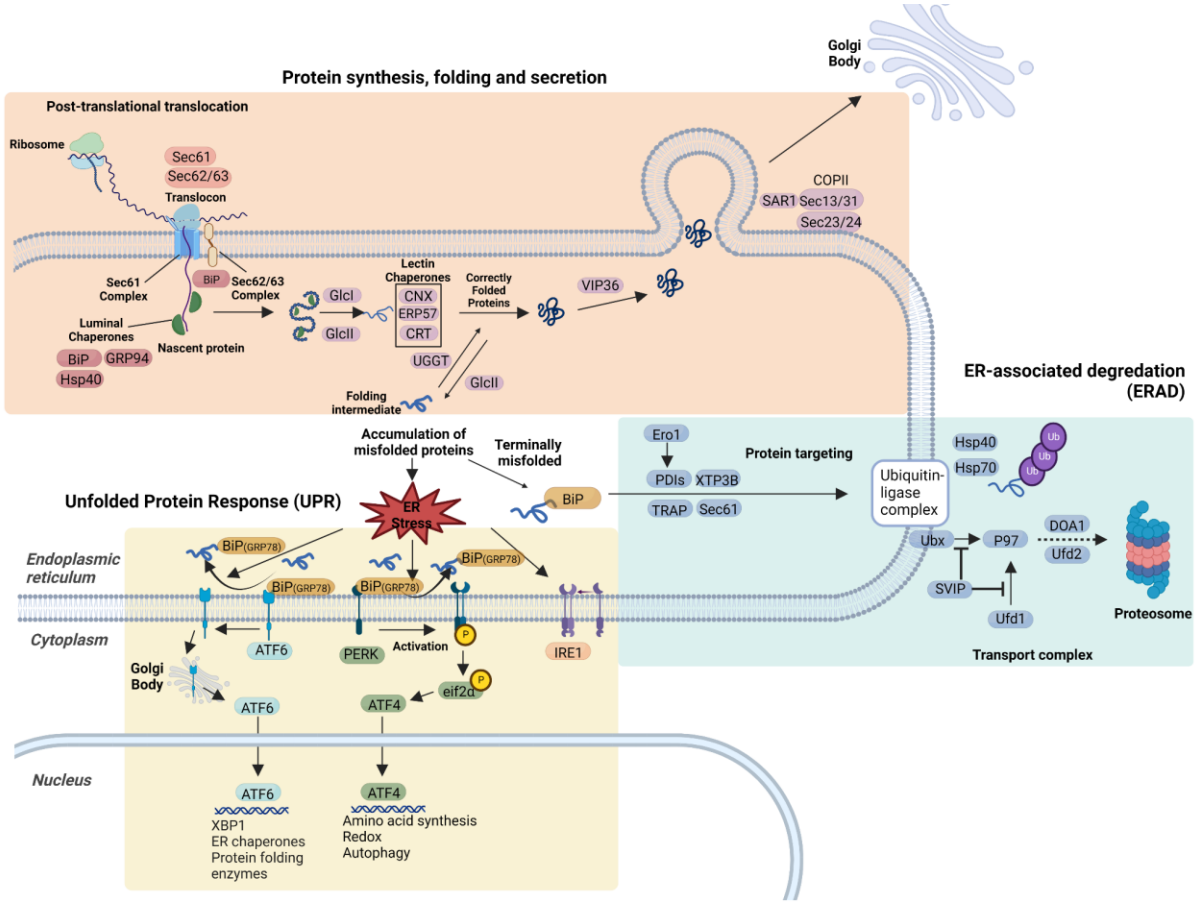


Figure 3. Graphical depiction of proteins investigated in this study in the Protein Processing pathway. "Protein Processing in ER" under physiological conditions and ER stress. This figure is designed to feature only proteins from this pathway that were observed to be differentially expressed in this study.

2.4.3 Protein synthesis, folding and secretion under ER stress conditions in producer and non-producer CHO cells.

The rough ER plays a crucial role in synthesis, folding and secretion of newly synthesized proteins in the cell. Nascent proteins (polypeptide chains) are produced by ER membrane-bound ribosomes and they are taken into the ER lumen through translocon channels via recognition of the signal sequence on their amino terminus (Saraogi and Shan, 2011). Following uptake into the ER lumen, the signal sequence gets cleaved and nascent proteins are positioned with lumenal chaperones for their correct folding (Saraogi and Shan, 2011). Lumenal chaperones such as BiP (Gething, 1999) and GRP94 (Melnick et al., 1992), also

form the first step of quality control of the polypeptides after they are inserted into the ER lumen, allowing sorting of proteins (correctly folded, misfolded, underglycosylated and unassembled). Polypeptide chains oriented by lumenal chaperones proceed to lectin chaperones (CNX and CRT with their co-chaperone ERP29), which also allows folding as well as quality checks (Ellgaard and Helenius, 2003; Nakao et al., 2017). Oligosaccharyltransferases (OST) transfers carbohydrates onto the peptide from lipid-linked precursors (Silberstein and Gilmore, 1996). GLCI removes the terminal glucose shortly after the glycan addition, exposing the second glucose and giving malectin a site where it can associate with the peptide (Trombetta et al., 1989). Simultaneously, GLCII (GANAB) removes the exposed second glucose (Trombetta et al., 1989). When all glucose has been removed, the protein cannot bind to lectin chaperones and folds to its native state. The folding sensor UGGT1 docks to proteins that have taken on a non-native conformation and re-glucosylates the glycan, allowing another round of lectin chaperone binding for additional folding or retention (Helenius et al., 1997). After folding is completed, correctly folded proteins are transported towards the exterior of the lumen, where the proteins exit to the Golgi apparatus (Alberts et al., 2002) . The five-protein COPII complex creates transport vesicles which transfer proteins from the ER to the Golgi apparatus (Salama and Schekman, 1995).

In our study, differential expression and subsequent gene ontology analysis of the producer and non-producer CHO cells upon thapsigargin treatment showed significant changes in expression of proteins involved in transport into the ER lumen, the N-glycosylation pathway and diversion of correctly folded proteins to the Golgi apparatus. Two proteins, GANAB (16-fold decrease) and UGGT1 (11-fold decrease) are found to be highly downregulated in CHO-K1 cells upon thapsigargin exposure. The glycosylation mechanism in CHO cells ensures the refolding, stability and bioactivity of recombinant proteins and various enzymes (Cherepanova et al., 2016). Both GANAB and UGGT1 are involved in this regulation and

quality control system in the ER. GANAB plays a crucial role in the folding and quality control of nascent glycoproteins by cleavage of alpha-1,3-linked glucose residues from the monoglucosylate glycoforms ($Glc_2Man_9GlcNAc_2$) (Burda and Aebi, 1999). UGGT1 also works as a quality control checkpoint for protein folding by detecting minor folding defects in glycoproteins and re-glucosylating the N-glycans near the misfolded region allowing proteins to be recognised by lectin-like molecular chaperones CNX and CRT, and get directed to either re-folding or degradation (Ito et al., 2015). Xu et al. showed the expression levels of GANAB1 and UGGT1 in various sublines of CHO cells and speculated that the CHO S subline, which has the fastest growth, also shows higher expressions of proteins related to glycosylation of proteins (Xu et al., 2017b).

In our study, induction of ER stress resulted in highest downregulation of GANAB and UGGT1 in CHO-K1 cell line compared to producer cell lines. In addition to significant changes in the expression levels of GANAB and UGGT1, the other glycosyltransferases such as PRKCSH, MOGS, RPN1, RPN2, STT3A, STT3B and DDOST, showed variation in expression levels (mainly downregulation upon thapsigargin exposure except RPN1 for CHO-K1 and CHO DP-12) within producer and non-producer CHO cells (Figure 4, Supplementary Table 7).

Our study also showed significant changes in the expression levels of the calcium dependent lectin chaperone, CNX and its co-factor ERP29. CNX was found to show increased levels in thapsigargin treated cells. ERP29 has been shown to interact with CNX, similarly to ERP57, with some data suggesting ERP29 and EPR57 recognise the same domain of CNX but with possibly differing modes of interaction (Nakao et al., 2017). The changes in the CNX/ERP system detected in all chosen (producer and non-producer) CHO cell lines may be attributed to the mechanism of action of thapsigargin and its effects on the Ca-dependent lumenal chaperones. LMAN2, which is downregulated in all three cell lines, plays a role in the early secretory pathway, and is needed for transport and sorting of glycoproteins carrying

high mannose-type glycans. It acts as an intracellular transmembrane lectin that shuttles between the ER, Golgi apparatus and plasma membrane (Shimada et al., 2003).

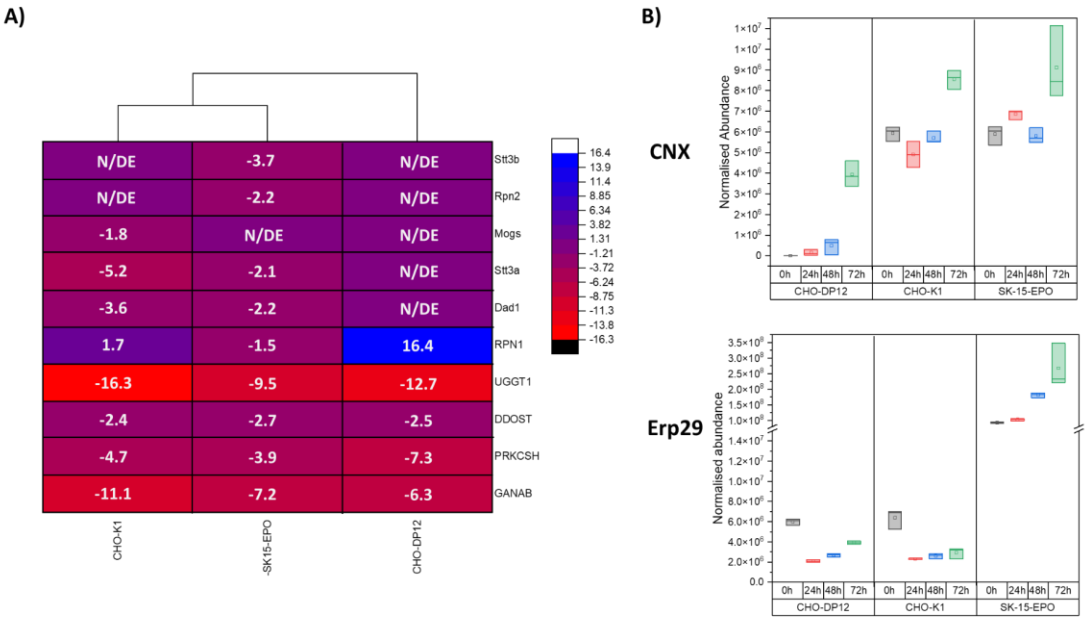


Figure 4. Comparison of glycosylation-related protein fold changes and normalised abundance. Proteins involved in post-glycosylation with changes in their expression profile upon induction of ER stress. A) Heatmap and fold changes of proteins involved in post-glycosylation of proteins in CHO-K1, CHO DP-12 and CHO SK15-EPO lines. B) Comparison of abundances of CNX and ERP29 upon thapsigargin exposure from 24 to 72 hours in CHO-K1, CHO DP-12 and CHO SK15-EPO cell lines. (Data sets are normalised within the group for each cell line).

2.4.4 Activation of the Unfolded Protein Response (UPR) under ER stress conditions in producer and non-producer cells

If there is an accumulation of misfolded proteins, forming faster than the ER can remove them, the unfolded protein response (UPR) is triggered, to return to homeostasis. A cell where the UPR is active is considered to be experiencing ER stress. The UPR represents several pathways within the cell, and in particular in the ER. Activation of the UPR in mammalian cell lines is facilitated by three ER transmembrane stress sensors: PKR-like ER kinase (PERK), inositol requiring enzyme 1 (IRE-1), and activating transcription factor 6 (ATF6) (Shaheen, 2018). These proteins initiate signalling pathways that lead to an

increased rate of protein folding and degradation and reduce protein translation. These proteins begin a complex transcriptional cascade, which causes the transcription of genes involved with ERAD, apoptosis, protein secretion and protein folding. These factors also lead to the upregulation of molecular chaperones, which aid the necessary genes in regaining homeostasis, or inducing cell death pathways (Senft and Ze'ev, 2015).

In physiological conditions, BiP (GRP78) binds to ATF6 which stabilises its location on the ER as a transmembrane protein. Accumulation of misfolded proteins in ER results in cleavage of BiP (GRP78) from ATF6 which leads to the translocation of ATF6 to the Golgi apparatus where it gets activated by cleavage of site-1-protease (S1P) and site-2 protease (S2P) (Shen et al., 2005). This cleavage liberates the ATF6 and translocates it to the nucleus where it activates the production of ER chaperones and protein-folding enzymes (Ye et al., 2000). Therefore, proteolytic cleavage of ATF6 (activation) and translocation to the nucleus provides recovery from ER stress. Ye et al. (2000) have shown that S2P lacking mutant CHO cell lines failed to form an active form of ATF6 and induction of ER chaperone, BiP, was completely abolished (Ye et al., 2000). Wang et al. and Yoshida et al. have also shown the role of ATF6 in ER stress response (Wang et al., 2000; Yoshida et al., 2000). Our study shows increased levels of ATF6 upon exposure to thapsigargin. The fold change was determined to be 945-fold for non-producer CHO-K1 cells, whereas it was 110-fold for both producer cell lines; CHO DP-12 and CHO SK15-EPO (Supplementary Table 8). The higher levels of ATF6 in non-producer cells could be related to the retained form of ATF6 (without proteolytic cleavage) in non-producer cells and higher detection via mass spectrometry, whereas active UPR mechanisms of producer cells to recover from ER stress may lead to proteolytic cleavage of ATF6; however, this requires further investigation.

2.4.5 Activation of ER associated degradation (ERAD) under ER stress conditions in producer and non-producer cells

If a protein does not fold properly after going through several cycles of re-folding with the aid of UGGT1, lectin chaperones and glucosidases, it remains tagged with BiP and it accumulates in the ER lumen and is classified as a misfolded protein. The misfolded proteins are checked via ER quality checkpoints such as MANI/MAN1B1, EDEM1-3 and if they cannot pass the quality check, they are tagged as terminally misfolded proteins (Hebert et al., 2014). The terminally misfolded proteins are directed to the ubiquitin-ligase complex with the aid of the protein targeting systems (e.g. OS-9, GRP94) and exported from the ER (Christianson et al., 2008). This pathway is known as ERAD and plays an important role in cellular homeostasis by reducing ER stress. Differential expression analysis of non-producer and producer cell lines under ER stress conditions showed significant changes in the transport of misfolded proteins to the proteasome, ERAD-associated heat shock proteins, targeting of misfolded protein to the ER membrane and the ubiquitin-ligase complex. These proteins, along with their expression changes in each cell line are shown in Figure 5 and Supplementary Table 9**.**

Following the exit of terminally misfolded protein from the ER, PLAA (DOA1) takes part in transport of these proteins in the cytoplasm to the proteasome where their ubiquitin dependant proteolysis occurs (Papadopoulos et al., 2017). Our study showed upregulation of PLAA upon thapsigargin exposure in both non-producer and producer cell lines compared to control (non-treated) cells. UFD1L is a part of the ubiquitin-dependent proteolytic pathway responsible for degrading ubiquitin fusion proteins (Johnson et al., 1995). There is downregulation of UFD1L following thapsigargin exposure. UFD2, or UBE4B, is an additional conjugation factor that is necessary for the efficient multiubiquitination of substrates required for proteasomal targeting (Koegl et al., 1999), and is downregulated in CHO SK15-EPO cells upon thapsigargin exposure.

Misfolded proteins are channelled out of the ER through a dislocon channel, where proteins in the cytoplasm take over transport to the proteasome. Members of the Heat Shock Protein (HSP) family, HSP70 and HSP90, act as chaperones outside of the ER and determine whether the protein can be repaired or degraded by the proteasomal system (Wegele et al., 2004). They are assisted in this process by several co-chaperones including CHIP (carboxyl terminus of HSP70-interacting protein), BAG-1 (Bcl-2-associated athanogene-1), HIP/p48 (heat shock cognate 70-interacting protein), and HOP/STI1 (HSP-organizing protein/stress-inducible protein), all of which influence the binding affinity of HSP70 and HSP90 (Kästle and Grune, 2012). HSPA1L is part of the HSP70 family and acts as a molecular chaperone related to several cellular processes including the protein quality control system, including re-folding of misfolded proteins and targeting of proteins for degradation (Radons, 2016). There is upregulation of HSPA1L in thapsigargin exposed cells. HSPA2 is a molecular chaperone associated with several cellular processes, including protection of the proteome from stress, folding and transport of newly synthesized polypeptides, proteolysis of misfolded proteins and the formation and dissociation of protein complexes (Radons, 2016). It is upregulated in thapsigargin treated cells. DNAJA1 and DNAJB1 showed greater downregulation in thapsigargin exposed cells. DNAJ family members function as molecular chaperones, which stimulate the ATPase activity of HSP70 heat-shock proteins, to prevent the accumulation of misfolded proteins and promote folding (Qiu et al., 2006).

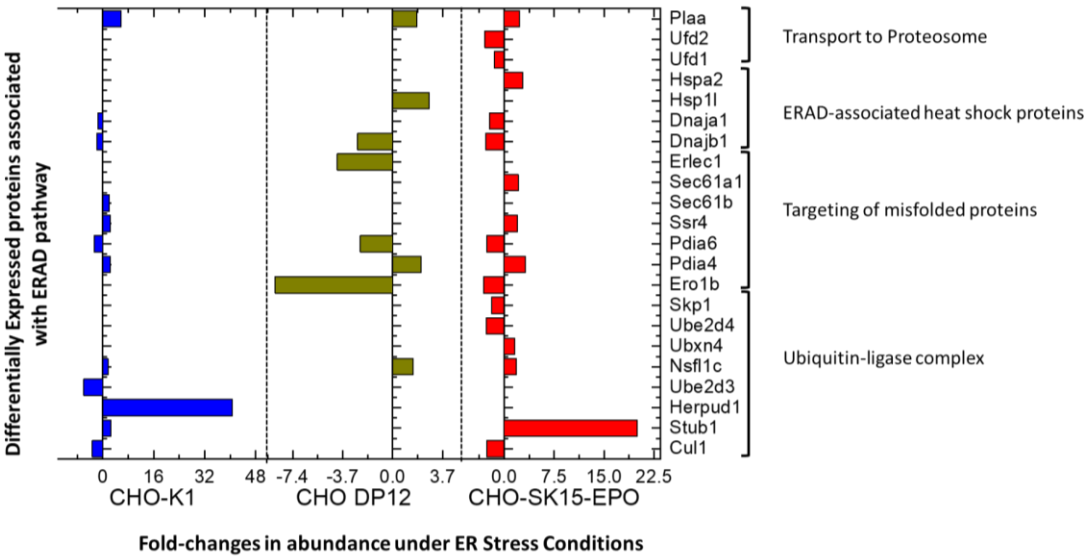


Figure 5. Fold changes of ERAD associated proteins in CHO-K1, CHO-DP12 and CHO-SK15-EPO. Proteins involved in ERAD pathway, associated fold-changes under ER stress conditions and biological function of the identified proteins in CHO-K1, CHO DP-12 and CHO SK15-EPO cells.

The protein disulfide isomerase (PDI) family of proteins play an important role in the quality control mechanisms within the ER. PDIs are versatile proteins that could potentially oxidize, reduce, or isomerize disulfide bonds in misfolded proteins to promote proper folding (Okumura et al., 2015), and also to promote the degradation of misfolded substrates (Forster et al., 2006). PDIA4 shows similar levels of up-regulation between all lines whereas PDIA6 shows a similar level of down-regulation. This increase in PDIA4 expression is consistent with the increase seen in CHO-K1 and CHO-HcD6 cells exposed to thapsigargin (Komatsu et al., 2020); however, this study did not see a decrease in PDIA6 expression. ERO1 works with PDIs to keep them in an oxidative state, allowing for the necessary redox reactions to occur (Shergalis et al., 2020) and is also part of chaperone cofactor-dependent protein refolding (Zito, 2015). ERO1β was found to be down-regulated in both producer cell lines CHO DP-12 and CHO SK15-EPO. SSR4 is responsible for the delta subunit of the translocon-associated protein complex needed for the transport of proteins across the ER

membrane (Hartmann et al., 1993). It is upregulated following thapsigargin exposure. SEC61A1 is part of the channel-forming translocon complex that facilitates transport of nascent polypeptides into the ER (Meacock et al., 2002), and is upregulated upon thapsigargin exposure. SEC61B is the beta subunit of the SEC61 complex, which forms a central part of the ER membrane protein translocation apparatus (Meacock et al., 2002). ERLEC1 is a resident ER lectin that may selectively bind to misfolded proteins in the ER (Cruciat et al., 2006), and is downregulated upon thapsigargin exposure.

Several proteins are necessary for the organisation and transportation of ubiquitinated peptides to the ubiquitin-ligase complex for degradation. The presence of ubiquitin was analysed by Western Blotting, with a pan-ubiquitin antibody (Figure 6). In all three cell lines studied, an increase in ubiquitination was observed upon exposure to thapsigargin over time. The smear-like appearance of the blots is an indication of proteins of varied molecular weights being ubiquitinated upon thapsigargin exposure. STUB1 is an E3 ubiquitin-protein ligase which directs chaperoned misfolded polypeptides to the proteasome (Ballinger et al., 1999). STUB1 shows greater upregulation in the EPO producer CHO SK15-EPO (+20 fold-change) than the non-producer CHO-K1 (+2.5-fold-change). UBE2D3 is part of the E2 ubiquitin-conjugating enzyme family, and functions by accepting ubiquitin from the E1 complex and catalyses the covalent attachment of ubiquitin to target proteins (Shi et al., 2017). HERPUD1 binds to ubiquilins, allowing for efficient degradation of CD3D via the ERAD pathway (Kim et al., 2008) and it is upregulated upon thapsigargin treatment in CHO-K1 cells (Figure 5, Supplementary Table 9).

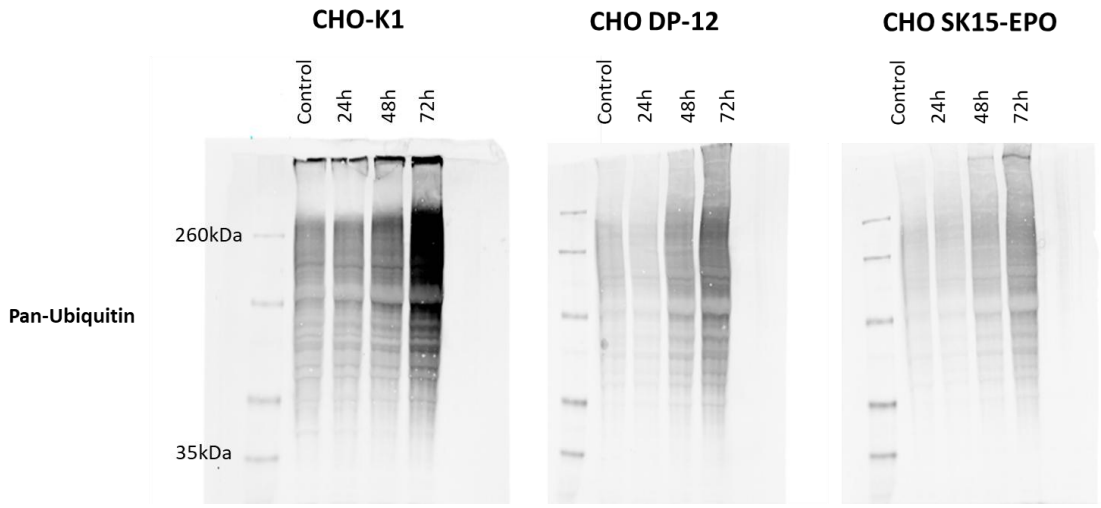


Figure 6. Western Blots using a Pan-Ubiquitin antibody on CHO Cells. Ubiquitination levels of CHO cell lines upon induction of ER stress.

SUGT1 may be involved in ubiquitination and proteasomal degradation of target proteins, but mechanisms of action are not yet fully understood, but it may suppress the expression of SKP1 (Kitagawa et al., 1999). SUGT1 is upregulated with thapsigargin exposure. UBXN4 is an integral membrane protein that allows for the recruitment of UBQLN1 (involved in proteasomal targeting of misfolded proteins for degradation) and VCP (needed for export of misfolded proteins from ER to cytoplasm) proteins to the ER (Lim et al., 2009). UBXN4 is up-regulated following thapsigargin exposure.

2.5 Conclusion

In-depth label free proteomic analysis of recombinant Chinese Hamster Ovary (CHO) cells exposed to thapsigargin allows for greater insight into the effects of ER stress on the protein processing pathways. Proteins involved in the unfolded protein response (UPR), ER associated degradation (ERAD) and the lectin chaperones involved in glycosylation mechanisms had changed expression levels between control and thapsigargin exposed producer (hEPO, IgG) and non-producer cell lines. Glycosylation cycle proteins GANAB and UGGT1 are significantly downregulated in producer and non-producer cells upon

exposure to thapsigargin, whereas the transcription factor ATF6 is upregulated. Producer cell lines showed a greater level of expression of proteins linked to protein targeting to ubiquitin-ligase complex and ERAD following ER stress induction when compared to the non-producer cell line, suggesting adapted mechanisms for the restoration of homeostasis within the ER. Non-producer CHO-K1 showed a greater number of common differentially expressed proteins with the producer CHO SK15-EPO than either cell line shared with the producer CHO DP-12, suggesting that the complexity of the protein produced may influence stress response mechanisms. The proteins differentially expressed in the producer cells may be potential targets for functional studies into engineered cell resistance to stress and improved productivity.

Data Availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE (Perez-Riverol et al., 2022) partner repository with the dataset identifier PXD042614 and 10.6019/PXD042614.

Author Contributions

David Ryan and Christiana-Kondylo Sideri contributed to sample preparation, acquisition, and interpretation of the data, they also drafted the manuscript. Michael Henry contributed to the data acquisition. Esen Efeoglu contributed to design of the work, data analysis, interpretation of the data and revised the work. Paula Meleady contributed to the conception of the work, interpretation of the data and revised the work. All authors have read and approved the submitted version.

Acknowledgement

Figure 3 was created with BioRender.com. Origin(Pro) Version 2022 was used for plotting and one-way ANOVA analysis.

Funding

This work was supported and funded by Science Foundation Ireland (SFI) Frontiers for the Future Award (grant no. 19/FPP/6759). The Orbitrap Fusion Tribrid mass spectrometer was funded under a Science Foundation Ireland Infrastructure Award to Dublin City University (SFI 16/RI/3701).

Conflict of Interest

The authors declare no conflict of interest.

2.6 References

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P., 2002. Transport from the ER through the Golgi Apparatus, in: Molecular Biology of the Cell. 4th Edition. Garland Science.

Ballinger, C.A., Connell, P., Wu, Y., Hu, Z., Thompson, L.J., Yin, L.Y., Patterson, C., 1999. Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions. Mol. Cell. Biol. 19, 4535-4545. https://doi.org/10.1128/MCB.19.6.4535

Burda, P., Aebi, M., 1999. The dolichol pathway of N-linked glycosylation. Biochim. Biophys. Acta BBA - Gen. Subj. 1426, 239-257. https://doi.org/10.1016/S0304-4165(98)00127-5

Castellano, B.M., Tang, D., Marsters, S., Lam, C., Liu, P., Rose, C.M., Sandoval, W., Ashkenazi, A., Snedecor, B., Misaghi, S., 2023. Activation of the PERK branch of the unfolded protein response during production reduces specific productivity in CHO cells via downregulation of PDGFRa and IRE1a signaling. Biotechnology Progress e3354.

Cherepanova, N., Shrimal, S., Gilmore, R., 2016. N-linked glycosylation and homeostasis of the endoplasmic reticulum. Curr. Opin. Cell Biol., Cell organelles 41, 57-65. https://doi.org/10.1016/j.ceb.2016.03.021

Christianson, J.C., Shaler, T.A., Tyler, R.E., Kopito, R.R., 2008. OS-9 and GRP94 deliver mutant α1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. Nat. Cell Biol. 10, 272-282. https://doi.org/10.1038/ncb1689

Cooper, G.M., 2000. The Endoplasmic Reticulum. Cell Mol. Approach 2nd Ed.

Costello, A., Coleman, O., Lao, N.T., Henry, M., Meleady, P., Barron, N., Clynes, M., 2018. Depletion of endogenous miRNA-378-3p increases peak cell density of CHO DP12 cells and is correlated with elevated levels of ubiquitin carboxyl-terminal hydrolase 14. Journal of biotechnology 288, 30-40.

Costello, A., Lao, N.T., Barron, N., Clynes, M., 2019. Continuous translation of circularized mRNA improves recombinant protein titer. Metab. Eng. 52, 284-292.

Cruciat, C.-M., Hassler, C., Niehrs, C., 2006. The MRH protein Erlectin is a member of the endoplasmic reticulum synexpression group and functions in N-glycan recognition. J. Biol. Chem. 281, 12986-12993. https://doi.org/10.1074/jbc.M511872200

Cunnea, P.M., Miranda-Vizuete, A., Bertoli, G., Simmen, T., Damdimopoulos, A.E., Hermann, S., Leinonen, S., Huikko, M.P., Gustafsson, J.-Å., Sitia, R., Spyrou, G., 2003. ERdj5, an Endoplasmic Reticulum (ER)-resident Protein Containing DnaJ and Thioredoxin Domains, Is Expressed in Secretory Cells or following ER Stress*. J. Biol. Chem. 278, 1059-1066. https://doi.org/10.1074/jbc.M206995200

Dahodwala, H., Sharfstein, S.T., 2017. The 'Omics revolution in CHO biology: roadmap to improved CHO productivity. Heterologous Protein Prod. CHO Cells Methods Protoc. 153-168.

Dhara, V.G., Naik, H.M., Majewska, N.I., Betenbaugh, M.J., 2018. Recombinant antibody production in CHO and NS0 cells: differences and similarities. BioDrugs 32, 571-584.

Du, Z., Treiber, D., McCoy, R.E., Miller, A.K., Han, M., He, F., Domnitz, S., Heath, C., Reddy, P., 2013. Non-invasive UPR monitoring system and its applications in CHO production cultures. Biotechnology and bioengineering 110, 2184-2194.

Ellgaard, L., Helenius, A., 2003. Quality control in the endoplasmic reticulum. Nat. Rev. Mol. Cell Biol. 4, 181-191.

Ellgaard, L., Molinari, M., Helenius, A., 1999. Setting the Standards: Quality Control in the Secretory Pathway. Science 286, 1882-1888. https://doi.org/10.1126/science.286.5446.1882

Ferris, S.P., Kodali, V.K., Kaufman, R.J., 2014. Glycoprotein folding and quality-control mechanisms in protein-folding diseases. Dis. Model. Mech. 7, 331-341. https://doi.org/10.1242/dmm.014589

Fischer, S., Handrick, R., Otte, K., 2015. The art of CHO cell engineering: A comprehensive retrospect and future perspectives. Biotechnol. Adv. 33, 1878-1896.

Forster, M.L., Sivick, K., Park, Y., Arvan, P., Lencer, W.I., Tsai, B., 2006. Protein disulfide isomerase-like proteins play opposing roles during retrotranslocation. J. Cell Biol. 173, 853-859.

Gething, M.-J., 1999. Role and regulation of the ER chaperone BiP, in: Seminars in Cell & Developmental Biology. Elsevier, pp. 465-472.

Hansen, H.G., Søltoft, C.L., Schmidt, J.D., Birk, J., Appenzeller-Herzog, C., Ellgaard, L., 2014. Biochemical evidence that regulation of Ero1β activity in human cells does not involve the isoform-specific cysteine 262. Biosci. Rep. 34, e00103. https://doi.org/10.1042/BSR20130124

Hartmann, E., Görlich, D., Kostka, S., Otto, A., Kraft, R., Knespel, S., Bürger, E., Rapoport, T.A., Prehn, S., 1993. A tetrameric complex of membrane proteins in the endoplasmic reticulum. Eur. J. Biochem. 214, 375-381.

Hebert, D.N., Lamriben, L., Powers, E.T., Kelly, J.W., 2014. The intrinsic and extrinsic effects of N-linked glycans on glycoproteostasis. Nat. Chem. Biol. 10, 902-910.

https://doi.org/10.1038/nchembio.1651

Helenius, A., Trombetta, E.S., Hebert, D.N., Simons, J.F., 1997. Calnexin, calreticulin and the folding of glycoproteins. Trends Cell Biol. 7, 193-200.

Hetz, C., 2012. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. Nat. Rev. Mol. Cell Biol. 13, 89-102.

Hetz, C., Zhang, K., Kaufman, R.J., 2020. Mechanisms, regulation and functions of the unfolded protein response. Nat. Rev. Mol. Cell Biol. 21, 421-438. https://doi.org/10.1038/s41580-020-0250-z

Ito, Y., Takeda, Y., Seko, A., Izumi, M., Kajihara, Y., 2015. Functional analysis of endoplasmic reticulum glucosyltransferase (UGGT): Synthetic chemistry's initiative in glycobiology. Semin. Cell Dev. Biol., Dendritic cell development and function & Protein quality control through glycosylation 41, 90-98. https://doi.org/10.1016/j.semcdb.2014.11.011

Johari, Y.B., Estes, S.D., Alves, C.S., Sinacore, M.S., James, D.C., 2015. Integrated cell and process engineering for improved transient production of a "difficult-to-express "fusion protein by CHO cells. Biotechnology and Bioengineering 112, 2527-2542.

Johnson, E.S., Ma, P.C.M., Ota, I.M., Varshavsky, A., 1995. A Proteolytic Pathway That Recognizes Ubiquitin as a Degradation Signal *. J. Biol. Chem. 270, 17442-17456. https://doi.org/10.1074/jbc.270.29.17442

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., Ishiguro-Watanabe, M., 2023. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 51, D587-D592.

Kästle, M., Grune, T., 2012. Interactions of the proteasomal system with chaperones: protein triage and protein quality control. Prog. Mol. Biol. Transl. Sci. 109, 113-160.

Kaufman, R.J., 2002. Orchestrating the unfolded protein response in health and disease. J. Clin. Invest. 110, 1389-1398. https://doi.org/10.1172/JCI0216886

Kim, T.-Y., Kim, E., Yoon, S.K., Yoon, J.-B., 2008. Herp enhances ER-associated protein degradation by recruiting ubiquilins. Biochem. Biophys. Res. Commun. 369, 741-746. https://doi.org/10.1016/j.bbrc.2008.02.086

Kitagawa, K., Skowyra, D., Elledge, S.J., Harper, J.W., Hieter, P., 1999. SGT1 Encodes an Essential Component of the Yeast Kinetochore Assembly Pathway and a Novel Subunit of the SCF Ubiquitin Ligase Complex. Mol. Cell 4, 21-33. https://doi.org/10.1016/S1097-2765(00)80184-7

Koegl, M., Hoppe, T., Schlenker, S., Ulrich, H.D., Mayer, T.U., Jentsch, S., 1999. A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly. Cell 96, 635-644. https://doi.org/10.1016/s0092-8674(00)80574-7

Komatsu, K., Kumon, K., Arita, M., Onitsuka, M., Omasa, T., Yohda, M., 2020a. Effect of the disulfide isomerase PDIa4 on the antibody production of Chinese hamster ovary cells. J. Biosci. Bioeng. 130, 637-643. https://doi.org/10.1016/j.jbiosc.2020.08.001

Kyeong, M., Lee, J.S., 2022. Endogenous BiP reporter system for simultaneous identification of ER

stress and antibody production in Chinese hamster ovary cells. Metab. Eng. 72, 35-45. https://doi.org/10.1016/j.ymben.2022.02.002

Lakshmanan, M., Kok, Y.J., Lee, A.P., Kyriakopoulos, S., Lim, H.L., Teo, G., Poh, S.L., Tang, W.Q., Hong, J., Tan, A.H., Bi, X., Ho, Y.S., Zhang, P., Ng, S.K., Lee, D., 2019. Multi-omics profiling of CHO parental hosts reveals cell line-specific variations in bioprocessing traits. Biotechnol. Bioeng. 116, 2117-2129. https://doi.org/10.1002/bit.27014

Latorre, Y., Torres, M., Vergara, M., Berrios, J., Sampayo, M.M., Gödecke, N., Wirth, D., Hauser, H., Dickson, A.J., Altamirano, C., 2023. Engineering of Chinese hamster ovary cells for co-overexpressing MYC and XBP1s increased cell proliferation and recombinant EPO production. Scientific Reports 13, 1482.

Li, Z.-M., Fan, Z.-L., Wang, X.-Y., Wang, T.-Y., 2022. Factors Affecting the Expression of Recombinant Protein and Improvement Strategies in Chinese Hamster Ovary Cells. Front. Bioeng. Biotechnol. 10, 880155. https://doi.org/10.3389/fbioe.2022.880155

Lim, P.J., Danner, R., Liang, J., Doong, H., Harman, C., Srinivasan, D., Rothenberg, C., Wang, H., Ye, Y., Fang, S., Monteiro, M.J., 2009. Ubiquilin and p97/VCP bind erasin, forming a complex involved in ERAD. J. Cell Biol. 187, 201-217. https://doi.org/10.1083/jcb.200903024

Lin, J.H., Walter, P., Yen, T.S.B., 2008. Endoplasmic Reticulum Stress in Disease Pathogenesis. Annu. Rev. Pathol. Mech. Dis. 3, 399-425. https://doi.org/10.1146/annurev.pathmechdis.3.121806.151434

Lindner, P., Christensen, S.B., Nissen, P., Møller, J.V., Engedal, N., 2020. Cell death induced by the ER stressor thapsigargin involves death receptor 5, a non-autophagic function of MAP1LC3B, and distinct contributions from unfolded protein response components. Cell Commun. Signal. 18, 12. https://doi.org/10.1186/s12964-019-0499-z

Lytton, J., Westlin, M., Hanley, M.R., 1991. Thapsigargin inhibits the sarcoplasmic or endoplasmic reticulum Ca-ATPase family of calcium pumps. J. Biol. Chem. 266, 17067-17071.

Malm, M., Kuo, C.-C., Barzadd, M.M., Mebrahtu, A., Wistbacka, N., Razavi, R., Volk, A.-L., Lundqvist, M., Kotol, D., Tegel, H., 2022. Harnessing secretory pathway differences between HEK293 and CHO to rescue production of difficult to express proteins. Metabolic engineering 72, 171-187.

Meacock, S.L., Lecomte, F.J.L., Crawshaw, S.G., High, S., 2002. Different transmembrane domains associate with distinct endoplasmic reticulum components during membrane integration of a polytopic protein. Mol. Biol. Cell 13, 4114-4129. https://doi.org/10.1091/mbc.e02-04-0198

Melnick, J., Aviel, S., Argon, Y., 1992. The endoplasmic reticulum stress protein GRP94, in addition to BiP, associates with unassembled immunoglobulin chains. J. Biol. Chem. 267, 21303-21306.

Nakao, H., Seko, A., Ito, Y., Sakono, M., 2017. PDI family protein ERp29 recognizes P-domain of

molecular chaperone calnexin. Biochem. Biophys. Res. Commun. 487, 763-767. https://doi.org/10.1016/j.bbrc.2017.04.139

Okumura, M., Kadokura, H., Inaba, K., 2015. Structures and functions of protein disulfide isomerase family members involved in proteostasis in the endoplasmic reticulum. Free Radic. Biol. Med. 83, 314-322.

Oslowski, C.M., Urano, F., 2011. Measuring ER Stress and the Unfolded Protein Response Using Mammalian Tissue Culture System, in: Methods in Enzymology. Elsevier, pp. 71-92. https://doi.org/10.1016/B978-0-12-385114-7.00004-0

Papadopoulos, C., Kirchner, P., Bug, M., Grum, D., Koerver, L., Schulze, N., Poehler, R., Dressler, A., Fengler, S., Arhzaouy, K., Lux, V., Ehrmann, M., Weihl, C.C., Meyer, H., 2017. VCP/p97 cooperates with YOD1, UBXD1 and PLAA to drive clearance of ruptured lysosomes by autophagy. EMBO J. 36, 135-150. https://doi.org/10.15252/embj.201695148

Park, J.H., Jin, J.H., Lim, M.S., An, H.J., Kim, J.W., Lee, G.M., 2017. Proteomic Analysis of Host Cell Protein Dynamics in the Culture Supernatants of Antibody-Producing CHO Cells. Sci. Rep. 7, 44246. https://doi.org/10.1038/srep44246

Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., Walzer, M., Wang, S., Brazma, A., Vizcaíno, J.A., 2022. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. Nucleic Acids Res. 50, D543-D552. https://doi.org/10.1093/nar/gkab1038

Pérez-Rodriguez, S., Wulff, T., Voldborg, B.G., Altamirano, C., Trujillo-Roldán, M.A., Valdez-Cruz, N.A., 2021. Compartmentalized Proteomic Profiling Outlines the Crucial Role of the Classical Secretory Pathway during Recombinant Protein Production in Chinese Hamster Ovary Cells. ACS Omega 6, 12439-12458. https://doi.org/10.1021/acsomega.0c06030

Pieper, L.A., Strotbek, M., Wenger, T., Olayioye, M.A., Hausser, A., 2017. ATF6β-based fine-tuning of the unfolded protein response enhances therapeutic antibody productivity of Chinese hamster ovary cells. Biotechnology and Bioengineering 114, 1310-1318.

Prashad, K., Mehra, S., 2015. Dynamics of unfolded protein response in recombinant CHO cells. Cytotechnology 67, 237-254. https://doi.org/10.1007/s10616-013-9678-8

Pybus, L.P., Dean, G., West, N.R., Smith, A., Daramola, O., Field, R., Wilkinson, S.J., James, D.C., 2014. Model-directed engineering of "difficult-to-express" monoclonal antibody production by Chinese hamster ovary cells. Biotechnology and bioengineering 111, 372-385.

Qiu, X.-B., Shao, Y.-M., Miao, S., Wang, L., 2006. The diversity of the DnaJ/Hsp40 family, the crucial partners for Hsp70 chaperones. Cell. Mol. Life Sci. CMLS 63, 2560-2570. https://doi.org/10.1007/s00018-006-6192-6

Radons, J., 2016. The human HSP70 family of chaperones: where do we stand? Cell Stress Chaperones 21, 379-404. https://doi.org/10.1007/s12192-016-0676-6

Rasmussen, U., Brøogger Christensen, S., Sandberg, F., 1978. Thapsigargine and thapsigargicine,

two new histamine liberators from Thapsia garganica L. Acta Pharm. Suec. 15, 133-140.

Read, A., Schröder, M., 2021. The unfolded protein response: an overview. Biology 10, 384

Reinhart, D., Sommeregger, W., Debreczeny, M., Gludovacz, E., Kunert, R., 2014. In search of expression bottlenecks in recombinant CHO cell lines—a case study. Applied microbiology and biotechnology 98, 5959-5965.

Roy, G., Zhang, S., Li, L., Higham, E., Wu, H., Marelli, M., Bowen, M.A., 2017. Development of a fluorescent reporter system for monitoring ER stress in Chinese hamster ovary cells and its application for therapeutic protein production. PLOS ONE 12, e0183694. https://doi.org/10.1371/journal.pone.0183694

Saghaleyni, R., Malm, M., Moruzzi, N., Zrimec, J., Razavi, R., Wistbacka, N., Thorell, H., Pintar, A., Hober, A., Edfors, F., Chotteau, V., Berggren, P.-O., Grassi, L., Zelezniak, A., Svensson, T., Hatton, D., Nielsen, J., Robinson, J.L., Rockberg, J., 2022. Enhanced metabolism and negative regulation of ER stress support higher erythropoietin production in HEK293 cells. Cell Rep. 39, 110936. https://doi.org/10.1016/j.celrep.2022.110936

Salama, N.R., Schekman, R.W., 1995. The role of coat proteins in the biosynthesis of secretory proteins. Curr. Opin. Cell Biol. 7, 536-543.

Saraogi, I., Shan, S., 2011. Molecular Mechanism of Co-translational Protein Targeting by the Signal Recognition Particle. Traffic 12, 535-542. https://doi.org/10.1111/j.1600-0854.2011.01171.x

Schwarz, D.S., Blower, M.D., 2016. The endoplasmic reticulum: structure, function and response to cellular signaling. Cell. Mol. Life Sci. 73, 79-94. https://doi.org/10.1007/s00018-015-2052-6

Senft, D., Ze'ev, A.R., 2015. UPR, autophagy, and mitochondria crosstalk underlies the ER stress response. Trends Biochem. Sci. 40, 141-148.

Shaheen, A., 2018. Effect of the unfolded protein response on ER protein export: a potential new mechanism to relieve ER stress. Cell Stress Chaperones 23, 797-806.

Shen, X., Ellis, R.E., Sakaki, K., Kaufman, R.J., 2005. Genetic interactions due to constitutive and inducible gene regulation mediated by the unfolded protein response in C. elegans. PLoS Genet. 1, e37.

Shergalis, A.G., Hu, S., Bankhead III, A., Neamati, N., 2020. Role of the ERO1-PDI interaction in oxidative protein folding and disease. Pharmacol. Ther. 210, 107525.

Shi, Y., Yuan, B., Zhu, W., Zhang, R., Li, L., Hao, X., Chen, S., Hou, F., 2017. Ube2D3 and Ube2N are essential for RIG-I-mediated MAVS aggregation in antiviral innate immunity. Nat. Commun. 8, 15138. https://doi.org/10.1038/ncomms15138

Shimada, O., Hara-Kuge, S., Yamashita, K., Tosaka-Shimada, H., Yanchao, L., Yongnan, L., Atsumi, S., Ishikawa, H., 2003. Clusters of VIP36-positive vesicles between endoplasmic reticulum and Golgi apparatus in GH3 cells. Cell Struct. Funct. 28, 155-163.

Silberstein, S., Gilmore, R., 1996. Biochemistry, molecular biology, and genetics of the

oligosaccharyltransferase. FASEB J. 10, 849-858. https://doi.org/10.1096/fasebj.10.8.8666161

Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., Jensen, L.J., von Mering, C., 2021. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 49, D605-D612. https://doi.org/10.1093/nar/gkaa1074

Tihanyi, B., Nyitray, L., 2020. Recent advances in CHO cell line development for recombinant protein production. Drug Discov. Today Technol. 38, 25-34.

Trombetta, S.E., Bosch, M., Parodi, A.J., 1989. Glucosylation of glycoproteins by mammalian, plant, fungal, and trypanosomatid protozoa microsomal membranes. Biochemistry 28, 8108-8116.

Wang, Y., Shen, J., Arenzana, N., Tirasophon, W., Kaufman, R.J., Prywes, R., 2000. Activation of ATF6 and an ATF6 DNA binding site by the endoplasmic reticulum stress response. J. Biol. Chem. 275, 27013-27020.

Wegele, H., Müller, L., Buchner, J., 2004. Hsp70 and Hsp90—a relay team for protein folding, in: Reviews of Physiology, Biochemistry and Pharmacology, Reviews of Physiology, Biochemistry and Pharmacology. Springer, Berlin, Heidelberg, pp. 1-44. https://doi.org/10.1007/s10254-003-0021-1

Xu, N., Ma, C., Ou, J., Sun, W.W., Zhou, L., Hu, H., Liu, X.M., 2017a. Comparative proteomic analysis of three Chinese hamster ovary (CHO) host cells. Biochem. Eng. J. 124, 122-129. https://doi.org/10.1016/j.bej.2017.05.007

Ye, J., Rawson, R.B., Komuro, R., Chen, X., Davé, U.P., Prywes, R., Brown, M.S., Goldstein, J.L., 2000. ER stress induces cleavage of membrane-bound ATF6 by the same proteases that process SREBPs. Mol. Cell 6, 1355-1364.

Yoshida, H., Okada, T., Haze, K., Yanagi, H., Yura, T., Negishi, M., Mori, K., 2000. ATF6 activated by proteolysis binds in the presence of NF-Y (CBF) directly to the cis-acting element responsible for the mammalian unfolded protein response. Mol. Cell. Biol. 20, 6755-6767.

Zito, E., 2015. ERO1: A protein disulfide oxidase and H2O2 producer. Free Radic. Biol. Med. 83, 299-304. https://doi.org/10.1016/j.freeradbiomed.2015.01.011

Chapter 3: Paper 2-Proteomic and Ubiquitinated Proteome Insights into ER Stress Responses in Chinese Hamster Ovary Cells Under Mild Hypothermic Conditions

Ryan, D., Sideri, C.-K., Henry, M., Karuppuchamy, S., Efeoglu, E. and Meleady, P. (2025), Proteomic and Ubiquitinated Proteome Insights into ER Stress Responses in Chinese Hamster Ovary Cells Under Mild Hypothermic Conditions. Biotechnology and Bioengineering. vol. 123,1 (2026): 5-25. doi:10.1002/bit.70081

This study explores changes in the proteome of CHO cells being cultured under mild hypothermic conditions, including the presence of ubiquitin in the proteome. Utilised throughout the biopharmaceutical industry in recombinant therapeutic production, mild hypothermia extends culture lifespan and increases product titre. While it has been extensively studied, how hypothermia causes this increase in productivity has yet to be comprehensively explained. Previous studies have suggested increased mRNA stability, enhanced recombinant gene transcription and larger cell size could be linked to the increase in productivity (Nguyen et al., 2020; C. m. Smales et al., 2004; Sunley et al., 2008). Aspects of ER function have also been linked to increase productivity at lower temperatures. Genes involved in the IRE1/XBP1 branch of the UPR have been found to be upregulated at lower temperature, with a simultaneous downregulation of proteins related to ERAD activity (Torres et al., 2021). Activation of the PERK-eIF2a pathway has also been found to help regulate the CHO cell response to the lowering of culture temperature (Wang et al., 2018).

In this study, label free LC-MS/MS was used to characterise the whole cell proteome of CHO cells cultured at 31°C and 37°C. The cultures were all initially grown at 37°C, and after a 48-h bulking-up period were lowered to 31°C. Proteomic analysis was carried out in cells on day 4 and day 7 of culture. This allowed for the comparison of change in the proteome over time in culture at different temperatures, yet still only represent snapshots of the dynamic nature of protein expression. This factor should be considered in future studies. In all three cell lines studied, a lower culture temperature resulted in reduced replication, but with extended culture lifespan. In both producer cell line, the 31°C cultures had increased final

product titre, as well as an increase in specific cell productivity. The downregulation of several proteins involved in the glycosylation of nascent polypeptides at 31°C, including DDOST, P4HB, PRKSCH and LMAN1, in all cell lines studied suggests that mild hypothermic shock disrupts the cell's normal ability to fold new proteins, leading to ER stress as the misfolded proteins accumulate in the ER, although the reduction in cell replication may mean that less proteins require folding, despite the increase in specific cell productivity.

A vital part of the ER stress response is the degradation of misfolded proteins thought the ERAD. As part of this, target proteins are tagged with ubiquitin, marking a protein for proteasomal degradation. Ubiquitin also plays a vital role in cell signalling regulation. Despite their prevalence and biological importance, PTMs are not regularly included in proteomic studies, due to technical limitations. In this study, we employed magnetic based immune-affinity enrichment to study the ubiquitinated proteome of CHO cells during mild hypothermia. This method uses a di-glycine remnant left on lysine residues following tryptic digestion and the distinct 114.1 Da mass shift on a peptide spectrum (van der Wal et al., 2018b). Enrichment in the ubiquitinated proteome indicates that pathways related to protein binding, enzyme binding and organic cyclic compound binding are under tight cellular regulation during mild hypothermic stress. Both ubiquitinated and non-ubiquitinated forms of proteins are present in cells at the same time, sometimes being differently regulated, suggesting both forms play a role in the cellular response to hypothermic stress.

This study demonstrates the importance of investigating the more complex aspects of the proteome and how this additional layer of detail can open new avenues for CHO cell engineering, allowing for consideration of ubiquitinated proteins in secretory pathway optimisation, clone screenings, and enhanced control of protein turnover.

In this work, I was responsible for the culturing of CHO-K1 and CHO-SK15-EPO cells for the whole cell proteomics, including temperature downshift, viability and productivity assays and cell pellet collection, with the aid of Christiana Kondylo Sideri for CHO-DP12 cell work. I prepared samples from all cell lines for proteomic analysis and performed differential expression analysis. For the ubiquitinated proteome work, I performed all large volume cell culture for all three cell lines. I carried out the prep scale protein digestion and immunoaffinity for ubiquitinated peptide enrichment, as well as the associated MS and DE analysis. In the manuscript preparation, I was the primary author for all sections, all data analysis and discussion of results.

Proteomic and Ubiquitinated Proteome Insights into ER Stress Responses in Chinese Hamster Ovary Cells Under Mild Hypothermic Conditions

David Ryan[1,2], david.ryan68@mail.dcu.ie

Christiana-Kondylo Sideri[1,2]; christianakondylo.sideri3@mail.dcu.ie

Michael Henry[1]; michael.henry@dcu.ie

Selvaprakash Karuppuchamy[1]; selvaprakash.karuppuchamy@dcu.ie

Esen Efeoglu[4]; esen.efeoglu@tudublin.ie

Paula Meleady[1,2,3*]; paula.meleady@dcu.ie

[1]Life Sciences Institute, Dublin City University, Glasnevin, Dublin 9, Ireland.

[2]School of Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland.

[3]SSPC Research Ireland Centre for Pharmaceuticals, Limerick, Ireland

[4]School of Food Science and Environmental Health, Technological University Dublin, Grangegorman, Dublin, Ireland,

*Corresponding author

Tel: +353-1-7005910; Email: paula.meleady@dcu.ie

Data availability

The mass spectrometry proteomics data has been submitted to and is available through the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2025) partner repository with the dataset identifier PXD063085 and 10.6019/PXD063085.

Acknowledgements/Funding

This research was funded by a Science Foundation Ireland Frontiers for the Future Award (grant no. 19/FFP/6759). The Orbitrap Fusion Tribrid mass spectrometer was funded by a Science Foundation Ireland Infrastructure Award to Dublin City University (SFI 16/RI/3701). Graphical abstract figure was created with BioRender.com.

Conflicts of interest

The authors declare that this research was conducted in the absence of any potential conflicts of interest.

Author contribution statement

David Ryan contributed to investigation, methodology, validation, and writing of the manuscript. Christiana-Kondylo Sideri contributed to investigation and validation of CHO-DP12 cell culture and productivity assay. Michael Henry contributed to the methodology and validation of the mass spectrometry data. Selvaprakash Karuppuchamy contributed to the methodology and validation of the ubiquitin enrichment and mass spectrometry analysis. Esen Efeoglu contributed to methodology and review of the manuscript. Paula Meleady contributed to conceptualisation, funding acquisition, project administration, supervision and review and editing of the manuscript. Authors have read and approved the submitted version.

3.1 Abstract

Chinese hamster ovary (CHO) cells, widely utilised in biopharmaceutical production, experience various stressors during cell culture that can affect protein expression and folding, particularly within the endoplasmic reticulum (ER). Mild hypothermia is widely employed in CHO cell bioproduction to improve recombinant protein yield and quality; however, its impact on ER-associated pathways, particularly those governing protein folding and stress responses, remains insufficiently characterized Mass spectrometry-based proteomics allows for the identification and relative quantification of proteins, enabling detailed insights into protein expression, modifications, and functional networks. This study investigates the impact of mild hypothermic conditions (31°C) on the whole cell proteome and ubiquitinated proteome of Chinese hamster ovary (CHO) cells, with a specific focus on ER proteins and ER stress. Using high-resolution mass spectrometry, we conducted a comprehensive proteomic and ubiquitinated proteomic analysis to quantify changes in protein abundance and ubiquitinated peptides under mild hypothermia. The downregulation of several proteins involved in the glycosylation of nascent polypeptides at 31°C, including DDOST, P4HB, PRKSCH and LMAN1, in all cell lines studied suggests that mild hypothermic shock disrupts the cell's normal ability to fold new proteins, leading to ER

stress as the misfolded proteins build up. When this is coupled with the maintained cell viability and increased productivity at 31°C, it indicates the ER stress response can mitigate the build-up of misfolded proteins. The differential regulation of the translation factor eIF2α, downregulated in non-producer cells but upregulated in producer cells at 31 °C, suggests that recombinant protein-producing CHO cells possess a more adaptive ER stress response, enabling more efficient function under hypothermic culture conditions. Enhanced ubiquitination of misfolded protein substrates highlights an increased reliance on ER-associated degradation (ERAD) pathways to alleviate proteotoxic stress, as well as the wide range of biological processes that are regulated by ubiquitination as part of the hypothermic stress response. These findings provide new insights into the cellular adaptation mechanisms of CHO cells to mild hypothermia, with implications for optimising bioproduction strategies to improve yield and quality of therapeutic proteins. Our study highlights the importance of understanding the more complex aspects of the proteome and how this additional layer of detail can open new avenues for CHO cell engineering.

Key Words

Chinese hamster ovary (CHO) cells, Endoplasmic Reticulum (ER) stress, unfolded protein response (UPR), Endoplasmic Reticulum associated degradation (ERAD), Hypothermia, Protein folding, Ubiquitination.

3.2 Introduction

The role of Chinese hamster ovary (CHO) cells in the biopharmaceutical industry is vital. Since the first CHO derived biotherapeautic, *Activase,* was given FDA approval in 1987, CHO cells have become the workhorse of the industry. Since then, CHO cells have become the preferred cell line for recombinant protein production, with 70% of recombinant protein therapies produced in CHO cells (W. Li et al., 2021). The popularity of CHO cells over other cell types is due to several key advantages, including: (i) reliable growth in chemically-defined and serum-free media, (ii) easily adapted to large-scale suspension culture, (iii) the

efficiency and accessibility of established gene editing and amplification techniques that allow for a specific productivity greater than other mammalian host cells, (iv) a well-documented history as a safe host cell line, with approval to market achieved by many CHO derived therapeutics, and (v) the ability to replicate human-like post-translational modifications on the recombinant proteins they produce (J. Y. Kim et al., 2012). Research on several CHO cell lines has shown that lowering culture temperature to between 31-33°C can have favourable effects on productivity. Lower culture temperature decreases cell growth rate, but this shift can result in an overall greater titre of protein product in the majority of CHO cell lines. It has been found that culture temperature of 35°C and the addition of 80 μM Zn2+ can increase monoclonal antibody (mAb) titre, and increase charge variants of the antibody, which is a crucial factor on product quality (Weng et al., 2020). CHO cells producing C-terminal α-amidating enzyme (Furukawa & Ohsuye, 1999), secreted alkaline phosphatase (Kaufmann et al., 1999) and EPO (Yoon et al., 2003) all demonstrated an increased recombinant protein yield when cultured at temperatures lower than 37°C. Yoon et al. suggested that differences in culture response to lower temperatures could be due to differences in the integration sites of the antibody gene, which occur during gene amplification caused by the sometimes random nature of chromosomal rearrangement (Yoon et al., 2004). It has been shown that a lower culture temperature (32°C) allowed for greater protein processing and folding capacity in rCHO cells by upregulating genes involved in the IRE1/XBP1 branch of the Unfolded Protein Response (UPR), while a simultaneous downregulation of proteins related to ERAD activity occurred (Torres et al., 2021). This study identified several physiological changes in rCHO cells cultured at reduced temperature, including the upregulation of genes encoding key UPR-specific transcriptional activators (e.g. *XBP1*s, *DDIT3*, *ATF5*) and ER-resident proteins (e.g. *GRP78, GRP94, TRIB3*), which are critical for maintaining protein folding homeostasis in the ER (Torres et al., 2021). In another study, a 25% increase in specific productivity of a mAb was observed

when the temperature was lowered to 32°C (Sou et al. 2015). However, a decreased proportion of more processed glycan structures, and a significant decrease in the expression of N-glycosyltransferases responsible for N-glycan branching and elongation of nascent peptides was observed in the lower temperature culture (Sou et al., 2015).

Protein folding within the ER is particularly important for glycoproteins, with 90% of glycoproteins undergoing the covalent addition of N-linked glycans in the ER (Helenius et al., 1997). The co-translational glycosylation of these polypeptides happens via the transfer of a preassembled oligosaccharide, $Glc_3Man_9GlcNAc_2$, to the nascent polypeptides (Weerapana & Imperiali, 2006). Glucosidases present in the ER continue to process the peptides to generate the monoglucosylated species ($Glc_1Man_9Glc\text{-}NAc_2$) that can be recognised by the membrane-bound lectin-chaperone calnexin (CNX) and its soluble orthologue calreticulin (CRT), both of which are needed to orchestrate the correct folding of the glycosylated polypeptides into glycoproteins (Rutkevich & Williams, 2011). If the newly formed proteins have been correctly folded, then the chaperones are released from the protein, and are recirculated in the ER (Ito et al., 2015). If, however, the proteins have been incorrectly folded, the glucosyltransferase UGGT1 re-glucosylates these proteins to create a peptide that will be recognised by CNX and CRT (Ito et al., 2015).

The disruption of normal ER function, termed ER stress, can be caused by internal or external factors such as high protein demand, viral infections, environmental toxins, inflammatory cytokines, and mutant protein expression (Oslowski & Urano, 2011). In response to the build-up of unfolded proteins, the cell activates the UPR (Bettigole & Glimcher, 2015). The complex machinery of the ER can become exhausted by overexpression of recombinant human wild-type proteins in CHO cells, proteins identical to those naturally produced in the human body, as shown in studies investigating blood coagulation factor VIII, a blood clotting factor that can be used in the treatment of haemophilia (Dorner et al., 1989) and antithrombin III, an anticoagulant that can be

administered to treat deep vein thrombosis or pulmonary embolism (Schröder & Friedl, 1997). This suggests that the extra workload CHO cells need to produce recombinant proteins may lead to an increase in misfolded proteins, inducing the conditions of ER stress. Three ER transmembrane sensors act as the main control of the UPR, inositol-requiring enzyme 1 (IRE1), activating transcription factor 6 (ATF6) and protein kinase R-like ER kinase (PERK) (Qu et al., 2021). When ER function is normal, these proteins are bound to the ER chaperone protein, BIP (GRP78), and are kept inactive; however, as misfolded proteins accumulate, BIP will dissociate from these proteins in order to bind to unfolded proteins, thus causing activation of the UPR (Gardner & Walter, 2011). A detailed understanding of ER function enables targeted genetic engineering of CHO cells to improve recombinant protein productivity. For example, enhanced expression of BiP, ATF6c, and XBP1s has been shown to increase IgG specific productivity (Qp) in the transiently transfected CHO-K1-derived cell line EB27 (Pybus et al., 2014b). Similarly, overexpression of ATF4 in CHO 13D-35D cells significantly boosted the production of recombinant human antithrombin III (Ohya et al., 2008). Additionally, knocking out the PERK branch of the UPR, along with the apoptosis-related factors Bax and Bak, in a DG44 mAb1-producing CHO cell line resulted in a 2-3-fold increase in specific productivity (Castellano et al., 2023).

The ERAD pathway is responsible for targeting misfolded proteins for degradation once they are recognised by the ER quality control machinery. Although the precise mechanisms of retro-translocation remain incompletely understood, current evidence suggests that this process is primarily mediated by the HRD1 complex, composed of HRD1, HRD3, DER1, USA1 and YOS9, alongside several other essential proteins and cofactors (Wu et al., 2020). ER Degradation Enhancing α-Mannosidase-Like proteins (EDEMs) recognise misfolded glycoproteins by binding to exposed mannose residues, thus targeting these glycoproteins for retro-translocation and subsequent degradation (T. Wang & Hebert, 2003). The energy needed for retro-translocation is provided by ATPases, such as p97/VCP (Ahlstedt et al.,

2022). Ubiquitin-regulatory X (UBX) proteins assist in directing substrates from the ERAD pathway towards autophagy (Kilgas & Ramadan, 2023), an alternative degradation route involving lysosomes.

In eukaryotic cells, the ubiquitin family of proteins vary in their peptide sequence, but as a result of folding, all show similar final structures, with most ubiquitin proteins terminating with a di-glycine sequence characteristic of this protein family, which is exposed following proteolytic processing (Pickart & Eddins, 2004). Ubiquitin is found free in the cell but is more commonly seen as part of a protein-ubiquitin conjugate (Bloom et al., 2003). This conjugation is formed via an isopeptide bond between the terminal glycine of ubiquitin and an amino group of the target protein, most commonly a lysine residue, but it is also possible for N-terminal ubiquitination to occur (Bloom et al., 2003).

Ubiquitination plays a critical role in the cellular response to ER stress. Ubiquitination influences UPR signalling and outcomes by regulating protein degradation and trafficking (Lopata et al., 2020). Ubiquitination targets misfolded proteins, marking them for proteasomal degradation or autophagic clearance. Additionally, ubiquitination modulates the stability and activity of key UPR and ERAD components, such as transcription factors and chaperones, fine-tuning the cellular response to ER stress (Carvalho et al., 2010; Denic et al., 2006). Although the ubiquitinated proteome of CHO cells is not yet well explored, this post-translational modification appears to be an important factor in protein signalling and ER stress pathways in CHO cells (Jung et al., 2015; Selvaprakash et al., 2024). Several studies have highlighted the relevance of ubiquitin-mediated mechanisms in the productivity of CHO cells. Xu et al. found that the addition of the ubiquitin-ligase inhibitor thalidomide prevented the intracellular degradation of the recombinant protein HSA-HGF at the later stages of CHO fed-batch culture, allowing for more efficient and stable secretion of the recombinant product (D. Xu et al., 2017). Tang et al. characterised a novel proteolysis/proteasome-dependent pathway involved in the degradation of antibody heavy

chains by the ligases UBR4 and UBR5, and found that reducing UBR4/UBR5 expression before the production phase, mAb productivity in CHO cells was increased (D. Tang et al., 2020). The inhibition of proteasomal activity in two producer CHO cell lines, expressing a model mAb and a fusion protein, during CHO host and recombinant cell pool selection, yields pools with improved product titres and specific productivity, including for a difficult-to-express molecule. This suggests that reducing ubiquitin-mediated clearance of recombinant proteins during selection favours cells with greater folding and secretion capacity (Knight et al., 2021). These findings indicate that ubiquitination is intrinsically linked with recombinant protein production in CHO cells. Investigating the ubiquitinated proteome further could identify additional degradation targets and pathways, possibly optimising production strategies across diverse biotherapeutic proteins. The importance of characterising other cellular post-translational modifications (PTMs) in proteomic studies has also been highlighted in recent publications. Characterisation of the global phosphoproteome of high/low producing industrially relevant CHO cell lines revealed phosphorylation of the kinase GCN2 is increased in high Qp cells, suggesting that phosphorylation plays an important role the function of this protein (Bryan et al., 2021). Histone modifications in CHO-K1 cells over a 9-day batch culture show correlations with gene expression levels, demonstrating regulation beyond simple presence/absence switches (Hernandez et al., 2019). Through metabolomic engineering, the build-up of toxic by-products can be reduced and protein production increased in CHO cells (Richelle & Lewis, 2017), and a detailed understanding of active PTMs in CHO cells could expand these results by increasing our knowledge of protein behaviour.

In-depth knowledge of CHO cell stress responses enables the development of more resilient cell lines. For example, Jiang et al. (2016) generated a stably transfected CHO cell line (CHO-GRP78) with a 1.5-2.5-fold increase in BiP compared to the parental cell line. It was found that these modified cells had increased viability and reduced apoptosis when exposed

to serum deprivation and oxidative stress (Jiang et al., 2016). It has also been shown that overexpression of BiP in CHO cells provides protection from ribose-induced cytotoxicity, while BiP knockdown led to a marked decrease in cell viability (Wu et al., 2018).

In this study, the role of ER stress mechanisms triggered by exposure to mild hypothermic conditions on the whole cell proteome and ubiquitinated proteome were investigated in three CHO cell lines, CHO-K1, CHO-DP12 which produces an anti-IL8-IgG, and CHO-SK15-EPO which produces human erythropoietin (EPO) using quantitative label-free mass spectrometry-based proteomics and subsequent gene ontology (GO) analysis. By lowering the culture temperature to the lowest temperature commonly used in industrial setting (31℃), we aimed to trigger a pronounced ER stress response. Cells were subjected to a temperature downshift from 37℃ to 31℃ forty-eight hours after seeding. Differential expression (DE) analysis of the whole cell proteome was carried out at two time points during culture, day four and day seven, followed by Gene Ontology (GO) analysis on the DE data to uncover which biological processes, molecular functions and KEGG pathways are enriched following a reduction in culture temperature. Ubiquitinated peptide enrichment was then carried out on the digested whole cell lysates, and DE and GO analysis performed on ubiquitinated peptides. With this first in-depth ubiquitinated proteomic study on hypothermia induced ER-stress mechanisms in CHO cells, we aim to improve our understanding of the role of ubiquitination in ER stress mechanisms, as well as protein folding processes, and highlight new possible targets for genetic engineering of CHO cells to enhance productivity and make cells more resilient to changes in culture conditions. While previous research has characterised global proteomic and transcriptomic changes in CHO cells subjected to temperature shifts, the specific impact of protein ubiquitination remains largely unexamined. By focusing on ubiquitin-mediated regulation, this work provides new insights into post-translational modifications that may underpin cellular adaptation and productivity changes during bioprocessing.

3.3 Methodology

3.3.1 Cell culture conditions

CHO-K1 (ATCC CCL-61), CHO-DP12 cell line (ATCC CRL-12445 clone #1934) which produces an anti-IL8-IgG, and CHO-SK15-EPO (an in-house CHO-K1 (ATCC CCL-61) derived cell line, expressing recombinant human erythropoietin (EPO) in a pcDNA3.1 vector (Invitrogen) modified with puromycin resistance as selection system) (Costello et al., 2019) were grown at 37°C under 5% CO_2 and 80% humidity at 170 RPM in BalanCD CHO Growth A (FujiFilm, 91128) cell culture medium supplemented with 4 mM L-glutamine (Sigma Aldrich, G7513) in a Climo-Shaker (Kühner). Cells were cultured in a working volume of 5 mL of cell culture media in 50 mL TPP® TubeSpin BioReactor Tubes (Merck, Z761028-180EA), with routine passaging at 2×10^5 cells/mL. The cells were routinely tested in-house for Mycoplasma contamination and were found to be negative. CHO-SK15-EPO cultures were pulsed with 10 mg/mL puromycin (Sigma-Aldrich/Merck, P8833) and CHO-DP12 with 400 nM MTX every 2-3 weeks to ensure only producing cells remained in culture.

3.3.2 Temperature Shift Culture Conditions

CHO-K1, CHO-DP12 and CHO-SK15-EPO cells were seeded at 2×10^5 cells/mL in triplicate tubes, cultured as described above and assayed using the COUNTESS II (Invitrogen, Thermo Scientific) and trypan blue staining (Invitrogen, Thermo Scientific) daily to collect growth and viability data. After two days, a subset of samples was transferred into an additional Kuhner incubator set at 31°C. Cell pellets for MS analysis were collected on day 4 and day 7 from the triplicate cultures from each of the three cell lines, and media samples for titre evaluation using ELISA were collected every second day, from separate triplicate cultures used for each measurement, so cell number or media volume was not impacted by sampling.

3.3.3 ELISA

hEPO titre from CHO-SK15-EPO cells and IgG titre from CHO-DP12 cells was measured using ELISA, as previously described (Ryan et al., 2023).

Specific productivity (pg/cell/day) was calculated as follows - productivity per viable cell on second time point minus productivity per viable cell on first time point, all divided by difference between time points.

Specific productivity (pg/cell/day) = (((P1-P0)/(((X1-X0)/2))))/(T1-T0)

Where,

P0 = productivity (pg/ml) by first point of analysis

P1 = productivity (pg/ml) by second point of analysis

X0 = Viable cell density ($x10^6$ cells/ml) by first point of analysis

X1 = Viable cell density ($x10^6$ cells/ml) by second point of analysis

T0 = Day of first point of analysis

T1 = Day of second point of analysis

3.3.4 Protein extraction from whole cell lysates

Cells were collected and pelleted at $3x10^6$ cells each, in triplicate, with three PBS wash steps between pelleting by centrifugation, then snap frozen and stored at -80 °C. Each sample was prepared for LC-MS/MS using the PreOmics iST kit, essentially as per manufacturer's instructions (PreOmics, P.O.00001). 100 µg of cell material was lysed with the lysis buffer (Preomics) at 95°C for 10 min at 1000 RPM. Following this, the Trypsin/LysC mix was added (as supplied in the iST kit), and the sample incubated at 37°C for 3 h at 500 RPM to allow proteolytic digestion to take place. Once the digestion was stopped, the peptide samples were washed and filtered through the cartridge. Once filtered, the peptide samples

were placed in the vacuum evaporator at 45°C until fully dry. Dried peptide samples were then stored at -80°C.

3.3.5 Ubiquitinated peptide enrichment

1.5mL of 8 M urea, 50 mM Tris, 100 mM NaCl (pH 8.5) buffer was used to lyse cells followed by sonication using an ultrasonicator with 30 s sonication pulses (×3 cycles) per sample, on ice. Cell debris was removed by centrifugation of the sample at 14,000×g for 10 min at 4 °C. The supernatant was collected, and the protein concentration was measured using the Bradford Assay (Thermo Scientific) at 660 nm. Five milligrams of protein from each sample is required for enrichment. Protein samples were adjusted to a volume of 1ml with 50 mM ammonium bicarbonate and 0.04% ProteaseMax (Promega). Samples are reduced at final concentration of 5 mM DTT (Sigma-Aldrich/Merck) at 56°C, at 500 RPM for 30 min. Samples are cooled and iodoacetamide (IAA) (Cell Signalling Technology) added to a final concentration of 25 mM and incubated at room temperature in the dark, at 500 RPM for 1 hour. Samples were then diluted to achieve a final concentration of urea <1.6 M. Lys-C (Cell Signalling Technology) was added at a 1:200 (w/w) ratio and incubated for 4 h at 37°C, with mixing every 30 min. Trypsin (Thermo Scientific) was added at a 1:100 (w/w) ratio and incubated for 18 h at 37°C.

Before beginning C-18 clean up, each sample was acidified using Trifluoroacetic Acid (TFA) at a 0.5% final concentration, giving a pH ~ 2. C-18 columns (Sep Pak, Waters) were used as per manufacturer's instructions. Columns were activated with 5 mL of 100% acetonitrile (ACN) (Thermo Scientific), then equilibrated with 0.1% TFA. Sample was then added, using 1 mL of 0.1% TFA to rinse the incubation tube. Elution was performed with 50% ACN:0.1% TFA, first with 4 mL, then with 2 mL. Eluates were flash frozen and then vacuum dried over 24 h.Magnetic bead immunoaffinity precipitation was completed using the PTMScan® HS Ubiquitin/SUMO Remnant Motif (K-ε-GG) Kit (Cell Signalling Technology) as per manufacturer's instructions. In a deviation from the manufacturer's

protocol, C-18 spin columns (Pierce, Thermo Fisher Scientific) were substituted for C-18 tips for greater certainty that the peptide amount is within the binding capacity of the column and used per manufacturer's instructions. After obtaining an eluate, the process was repeated with the remaining peptide solution to ensure the most peptides possible were obtained. The combined eluates were dried using a Speed-Vac for approximately 2 h, until entirely dried.

3.3.6 LC-MS/MS for Label free quantification (LFQ)

For LC-MS/MS analysis, an UltiMate 3000 nano RSLC (Thermo Scientific) system interfaced with an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific) was used. Using nano-flow HPLC separation, 1 µg of tryptic peptide from each sample was loaded onto the trapping column (PepMap100, C18, 300 µm × 5 mm) at a flow rate of 25 µL/min with 2% (v/v) ACN, 0.1% (v/v) TFA for a 3 min desalting and concentration step. Each sample was then resolved onto an analytical column (Acclaim PepMap 100, 75 µm × 50 cm, 3 µm bead diameter column) into the mass spectrometer. A binary gradient of: solvent A (0.1% (v/v) formic acid in LC-MS grade water) and solvent B (80% (v/v) ACN, 0.08% (v/v) formic acid in LC-MS grade water) using 2%–32% B for 75 min, 32%–90% B in 5 min and holding at 90% for 5 min at a flow rate of 300 nL/min was used to elute peptides.

For MS and MS/MS analysis, a temperature of 320°C and a voltage of 2.0 kV was used for peptide ionisation. The Orbitrap mass analyser with a resolution of 120,000 (at m/z 200), a maximum injection time of 50 ms and an automatic gain control (AGC) value of 4 × 105 was used to perform full MS scans for peptide mass information. Data-dependent acquisition was performed using a full scan range of 380–1500 m/z. A top-speed MS/MS acquisition algorithm was used to determine the number of selected precursor ions for fragmentation. A dynamic exclusion was applied to the analysed peptides after 60 s and peptides with a charge state between 2+ and 7+ were analysed. Peptides were fragmented using higher energy collision-induced dissociation (HCD) with a normalised collision energy of 29%, with a

maximum injection time of 35 ms and an AGC value of 2 × 104. The resulting MS/MS fragment ions were measured in the low-resolution linear ion trap.

3.3.7 Protein identification and differential expression analysis using quantitative label-free LC-MS/MS

Quantitative label-free analysis was performed using Progenesis QI for Proteomics to compare cells cultured at 31°C to 37°C on day 4 and day 7. Only peptide ions with charge states +2 and + 3 were allowed and normalised. The normalised peptide abundance data was transformed prior to statistical analysis, using an arcsinh transformation to meet the assumptions of the one-way ANOVA test. Peptides with a one-way ANOVA p-value ≤ 0.05 between experimental groups were exported and identified using Proteome Discoverer as described. Protein identifications were imported into Progenesis and considered differentially expressed proteins if they passed the following criteria: (i) a protein level one-way ANOVA p-value ≤ 0.05 between experimental groups, (ii) a minimum of 1 peptide contributing to a protein identification and (iii) a minimum fold change in abundance of ± 1.5-fold between culture temperatures.

Peptides were identified using Proteome Discoverer 2.4 (Thermo Scientific) against a Cricetulus griseus UniProt database (downloaded September 2022) using the SEQUEST HT algorithm and exported to Progenesis QI for Proteomics (NonLinear Dynamics, a Waters Company, UK) for relative quantitative analysis between experimental groups.

The following search parameters were set for protein identifications: (i) MS/MS mass tolerance set at 0.6 Da; (ii) peptide mass tolerance set to 20 ppm; (iii) trypsin as the enzyme; (iv) up to two missed cleavages were allowed; (v) cysteine carbamidomethylation as a static modification and (vi) methionine oxidation as a dynamic modification.

3.3.8 LC-MS/MS of ubiquitinated peptides and identification of differentially expressed ubiquitinated proteins

Ubiquitinated peptide LC-MS/MS and identification was carried out as described in (Selvaprakash et al., 2024). One microgram of Ubiquitinated peptides from each sample were loaded onto the trapping column at a flow rate of 25 µL/min with 2% (v/v) ACN and 0.1% (v/v) TFA for 3 min. Samples were resolved in the analytical column using a binary gradient of formic acid solution and ACN/formic acid solution using a flow rate of 280 nL/min to elute peptides. Data-dependent acquisition was used for all MS run, using a full scan range of 380-1500 *m/z*, over 140 min. MS1 spectra were collected at a resolution of 120,000 with an automated gain control (AGC) target 2×10^5 with a maximum injection time of 50 ms. Dynamic exclusion was applied to analyse peptides after 60 s, and peptides with a charge state between 2+ and 7+ were analysed. Peptides were fragmented via higher energy collision-induced dissociation with a collision energy of 29, an AGC value of 1×10^5, and a maximum injection time of 90 ms. Peptides were identified using Proteome Discoverer 2.4 (ThermoScientific) against a *Cricetulus griseus* UniProt database (downloaded September 2022) using the SEQUEST HT algorithm and exported to Progenesis QI for Proteomics (NonLinear Dynamics, a Waters Company, UK) for relative quantitative analysis between experimental groups. The following search parameters were set for protein identifications: (i) MS/MS mass tolerance set at 0.6 Da; (ii) peptide mass tolerance set to 20 ppm; (iii) trypsin as the enzyme; (iv) up to two missed cleavages were allowed; (v) cysteine carbamidomethylation as a static modification; (vi) methionine oxidation as a dynamic modification; and (vii) lysine with a K-ε-GG remnant (114.0429 Da) as a variable modification. Protein identifications were imported into Progenesis and considered differentially expressed proteins if they passed the following criteria: (i) a protein level one-way ANOVA p-value ≤ 0.05 between experimental groups, (ii) a minimum of 1

ubiquitinated peptide contributing to a protein identification and (iii) a minimum fold change in abundance of ± 1.5-fold between culture temperatures.

3.3.9 KEGG pathway and STRING association analysis

Whole cell proteins which were differentially expressed between control and test samples were inputted into STRING (https://www.string-db.org) (Szklarczyk et al., 2021), which was used to group the proteins based on protein-protein interactions and to provide groupings within KEGG biological pathways (https://www.genome.jp/kegg) (Kanehisa et al., 2023). The differentially expressed ubiquitinated proteins were also analysed using this same methodology.

3.3.10 Graph Plotting

All data plotting was performed using R, https://www.R-project.org/, package *ggplot2* (Wickham, 2016).

3.4 Results and Discussion

3.4.1 Culture viability and recombinant protein production

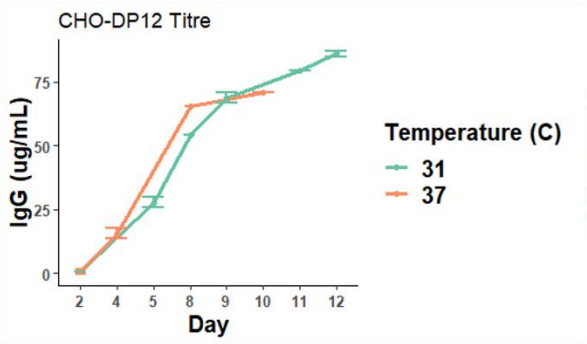
A reduction in CHO culture temperature is regularly used in the biopharma industry after the initial growth phase to increase productivity of cells or enhance product quality attributes of the recombinant protein or antibody (Torres et al., 2021). In this study, non-producer (CHO-K1) and producer (CHO-SK15 EPO and CHO DP-12) CHO cells were exposed to mild hypothermic conditions, and the ER stress response was analysed at the proteomic level. These cell lines allow for comparison between the native stress response in CHO-K1and the stress response in cells engineered for recombinant protein production (hEPO from CHO-SK15-EPO) or a more complex antibody (human IgG1 from CHO-DP12). Previous studies have reported on the lowering of culture temperatures from 37°C to between 35°C and 31°C. In our study we reduced the culture temperature to 31°C as this was the most extreme condition observed in these previous studies, in pursuit of the most comprehensive

view of ER stress induced by this method. Previous studies have also indicated that a moderate shift to 35°C decreased final titre relative to the unshifted control whereas a temperature shift to 32°C can significantly increase final titre (McHugh et al., 2020), suggesting a greater temperature change can be more beneficial to productivity.. Enhanced productivity in CHO cells has been linked to pathways involving mRNA stability (Oguchi et al., 2006), cell cycle arrest (Chen et al., 2004), and protein secretion (Yee et al., 2009), with additional studies implicating reduced apoptosis (Moore et al., 1997), increased secretory capacity (Smales et al., 2004), and elevated recombinant gene transcription (Kou et al., 2011). **Figure 1** shows the viable cell count and viability for the three cell lines studied, where cultures maintained at 37°C show an increase in cell density at a faster rate than cells cultured at 31°C, and reach a greater peak viable cell density than the cells cultured at 31°C, CHO-K1 between days 3-5, and between days 5-7 in CHO-DP12 and CHO-SK15-EPO. CHO-K1 starts to show a steady reduction in viability at 37°C by day 3 of culture, while at 31°C, there is a more gradual reduction over the culture period, with a noticeable drop only seen at day 9. Both CHO-DP12 and CHO-SK15-EPO cultures show a slower decline in percentage viability at 37°C at the earlier stages of culture, but a more rapid decline in later stages, in particular in CHO-SK15-EPO, where viability drops from 93% at day 7 to 43% at day 10. Viability remains steady in both producer cell lines at 31°C until later stages of culture, with a far more gradual decline than the rapid decline seen at 37°C. The two producer cell lines reach a peak viable cell count of approximately $5x10^6$ cells/mL at 31°C, less than half of the count reached by the 37°C cultures. These results align with previous studies, with all three cell lines maintaining culture viability for longer at 31°C than at 37°C, but with a much-reduced growth rate (Goey et al., 2017; C. M. Smales et al., 2004; Torres et al., 2021).

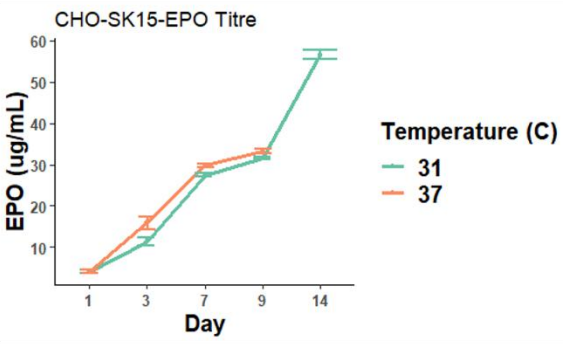
Figure 1. Viable cell density (cells/mL) and percentage cell viability for **(A)** CHO-K1, **(B)** CHO-DP12 and **(C)** CHO-SK15-EPO cells cultured at 37°C and 31°C, where culture temperature downshift occurs at 48 hrs post cell seeding. In all plots, orange represents 37°C, and blue represents 31°C. The data lines are the mean value of triplicate results, with error bars representing the standard deviation from the mean value at each time point.

The different culture temperatures can also affect productivity in the two producer cell lines, as shown in **Figure 2**. Although product titre is initially higher in cultures maintained at 37 °C, the extended viability of cells at 31 °C ultimately results in a greater final titre following the temperature shift. Comparisons of the specific cell productivity (Qp) show that cells subjected to the temperature shift are producing more recombinant protein than cells cultured at 37°C, with CHO-DP12 having a 4.4-fold increase in Qp at 31°C, and CHO-SK15-EPO having a smaller 1.7-fold increase at 31°C. The increase in productivity as a result of lowering the culture temperature is in line with previous hypothermia studies of CHO cells (Sou et al., 2015; Torres et al., 2021; Yoon et al., 2003), with previous publications suggesting that this increase may be due to increased transcriptional activity (C. M. Smales et al., 2004) or enhanced mRNA stability (Nguyen et al., 2020).

A



B



C

Qp	31°C	37°C
CHO-DP12 (Day 8)	12.78	2.61
CHO-SK15-EPO (Day 9)	1.48	0.85

Qp=pg/cell/day

Figure 2. (A) IgG titre from CHO-DP12 cells cultured at 31°C and 37°C, **(B)** hEPO titre from CHO-SK15-EPO cells cultured at 37°C and 31°C, and **(C)** specific cell productivity from both cell lines measured at a later time point in the 37°C-culture lifespan. Titre was measure by means of ELISA assay. In all plots, orange represents 37°C, and blue represents 31°C. The data lines are the mean value of triplicate results, with error bars representing the standard deviation from the mean value at each time point.

3.4.2 Gene ontology analysis of differentially expressed proteins in CHO cells under mild hypothermic conditions

In this study, whole-cell proteome analysis was performed on the three CHO cell lines subjected to a temperature shift from 37°C 31°C to investigate temperature-dependent variations in protein abundance. Samples were collected at two time points, day 4 and day 7, to study changes over the culture period. Quantitative label-free mass spectrometry was used to compare proteomic profiles, with the aim to gain insights into how culture temperature impacts culture viability and productivity. The total number of differentially expressed proteins for each cell line is shown in **Table 1**. All differentially expressed proteins are listed in **Supplementary Materials-1.**

Table 1: The number of proteins identified to be upregulated at each culture temperature on day 4 and day 7, in non-producing cell line CHO-K1, hEPO producer CHO-SK15-EPO and IgG producer CHO-DP12. Proteins identified show at least 1.5-fold change between conditions, an ANOVA p-value of less than or equal to 0.05 and are identified by at least one peptide matching the database sequence.

	CHO-K1		**CHO-SK15-EPO**		**CHO-DP12**	
	31°C	***37°C***	***31°C***	***37°C***	***31°C***	***37°C***
Day 4	526	380	347	514	561	547
Day 7	650	624	408	420	650	624

The differentially expressed proteins were subjected to GO analysis, focusing on KEGG pathways that are enriched in these proteins. All enriched KEGG pathways for all three cell lines can be found in **Supplementary Materials-2 (Table 1)**. At both time points, metabolism-related pathways including, 'Metabolic pathways' (cge0110), 'Carbon metabolism' (cge01200) and 'Purine metabolism' (cge0230), are shown to have changes in protein expression following a reduction in culture temperature, indicating that culture temperature affects how cells metabolise nutrients, in particular carbon, which may account for the reduction in growth. It was observed that a higher number of differentially expressed proteins mapped to the 'Proteasome' (cge03050) pathway at day 7 than day 4, with this pathway not showing up as enriched in CHO-SK15-EPO on day 4. The majority of these proteins showed an increase in levels of abundance on day 7 in the temperature shifted cultures, suggesting there are more misfolded proteins being degraded in these cultures. This is most likely due to the cells at the lower temperature producing more of the recombinant protein, whereas at 37°C cell viability is starting to decrease, so cells are likely producing less recombinant protein than earlier in culture. 'Apoptosis' (cge04210) is also enriched,

suggesting some cells are unable to restore ER homeostasis and apoptotic cell death is triggered.

'Protein processing in Endoplasmic reticulum' (cge04141) is a KEGG pathway which maps the proper functioning and homeostasis of the ER in CHO cells via multiple mechanisms such as the UPR, ERAD and proteasomal degradation. This pathway appears enriched at both time points in all three cell lines. This pathway is one of the main focuses of this study as the differentially expressed proteins allow for investigation of both the normal folding and quality control process, as well as those involved in the ER stress response. This allows for a greater focus on ER proteins that may otherwise be overlooked if a broader scope study was undertaken, looking at all differentially expressed proteins. The number of differentially expressed proteins for each cell line at day 4 and day 7 which contribute to the enrichment of this pathway are shown in **Table 2**, along with the strength and false discovery rate associated with this pathway according to STRING analysis. The KEGG pathways 'Apoptosis' (cge04210) is also enriched at day 4 in all three cell lines, with proteins related to caspase-mediated apoptosis differentially expressed (BAK1, EIF2α, CASP7), suggesting that cells where ER homeostasis cannot be restored will undergo apoptosis via these specific pathways.

Table 2: Comparison of the number of differentially expressed proteins which contribute to the enrichment of the KEGG Pathway: 'Processing in Endoplasmic Reticulum' (cge04141)

	Count	Strength	False Discovery Rate
Day 4			
CHO-K1	24/128	0.7	2.84e-08
CHO-SK15-EPO	27/128	0.77	4.45e-10
CHO-DP12	27/128	0.66	2.12e-08
Day 7			
CHO-K1	33/128	0.69	1.25e-08
CHO-SK15-EPO	22/128	0.7	7.06e-07
CHO-DP12	33/128	0.81	1.79e-13

3.4.3 Gene ontology analysis of differentially expressed ubiquitinated proteins in CHO cells under mild hypothermic conditions

The ubiquitinated proteome represents an under investigated aspect of the CHO proteome. The importance of ubiquitination in various cell processes, in particular relating to protein production, makes it a promising area for new insights into CHO productivity. Studies have shown that using targeted miRNAs to affect levels of some ubiquitin-proteasome related proteins has phenotypic effects on growth (Costello et al., 2018) and productivity (Fischer, Mathias, et al., 2015). Following tryptic digestion of proteins, a Lys-ε-diglycine (diGly) remnant is left on a peptide, indicating a site of ubiquitination. This tag has become the target for an immunoaffinity-based methods of enrichment of ubiquitinated peptides (using an anti-K-ε-GG antibody), allowing these peptides to be detected using mass spectrometry (van der Wal et al., 2018a). This is the method used here to investigate the effects of temperature shift

on the ubiquitinated proteome of non-producer CHO-K1, hEPO producer CHO-SK15-EPO and IgG producer CHO-DP12. Differentially expressed ubiquitinated proteins are defined here as protein identified by at least one peptide with a diGly remnant matching the database sequence, at least 1.5-fold change between conditions and an ANOVA p-value of less than or equal to 0.05. The number of proteins matching these parameters for each cell line at day 4 and day 7 is listed in **Table 3**. Complete lists of all differentially expressed ubiquitinated proteins can be found in **Supplementary Materials-3**.

Table 3: The number of ubiquitinated proteins identified to be up-regulated at each culture temperature on day 4 and day 7, in non-producing cell line CHO-K1, hEPO producer CHO-SK15-EPO and IgG producer CHO-DP12.

	CHO-K1		**CHO-SK15-EPO**		**CHO-DP12**	
	31°C	***37°C***	***31°C***	***37°C***	***31°C***	***37°C***
Day 4	76	209	124	405	33	85
Day 7	423	58	91	25	28	27

These proteins map to many cellular pathways, functions and processes, including some that are not identified as enriched in the whole cell proteome data. Complete lists of these enriched features are available in **Supplementary Materials-2 (Table 2).** The KEGG pathways, molecular functions and processes that are shared between the cell lines on day 4 and day 7 are listed in **Table 4.** On day 4, processes related to cellular organisation may hint at the cells attempt to reorganise its functions in response to the shock of the downshift in culture temperature. On day 4, the 'Proteasome' (cge03050) pathway is enriched in differentially expressed ubiquitinated proteins in both producer lines CHO-SK15-EPO and CHO-DP12. This suggests that both these cell lines use ubiquitination in the regulation of their response to the ER stress induced by temperature downshift, given the crucial role

played by the proteasome in degradation of misfolded proteins (Roos-Mattjus & Sistonen, 2004). On day 7, this pathway is enriched in differentially expressed ubiquitinated proteins in CHO-SK15-EPO and CHO-K1, but not in CHO-DP12, suggesting that CHO-DP12 no longer requires such tight regulation of this pathway, and that CHO-K1 now needs greater control of proteasomal activity than at day 4.

Table 4: The KEGG pathways, molecular functions and processes enriched in ubiquitinated differentially expressed proteins shared between CHO-K1, CHO-SK15-EPO and CHO-DP12 on day 4 **(A)** and day 7 **(B)**. No KEGG pathways were enriched in CHO-K1 at day 4.

A

		CHO-SK15-EPO			CHO-DP12		
KEGG	Description	Count	Strength	FDR	Count	Strength	FDR
cge05012	Parkinson disease	7	0.87	0.02	9	1.01	4.51E-05
cge04520	Adherens junction	4	1.2	0.0271	3	1.1	0.0432
cge04810	Regulation of actin cytoskeleton	6	0.88	0.0271	5	0.83	0.0246
cge05014	Amyotrophic lateral sclerosis	7	0.69	0.0271	9	0.82	0.00046
cge03050	Proteasome	3	1.19	0.0295	8	1.65	1.44E-08

		CHO-K1			CHO-SK15-EPO			CHO-DP12		
Term ID	Function	Count	Strength	FDR	Count	Strength	FDR	Count	Strength	FDR
GO:0005488	Binding	293	0.13	1.20E-11	111	0.19	2.28E-11	98	0.17	1.76E-06
GO:0005515	Protein binding	183	0.17	2.64E-07	79	0.3	1.32E-09	67	0.25	2.50E-05
GO:0097159	Organic cyclic compound binding	140	0.15	0.00057	55	0.24	0.0051	52	0.24	0.0041
GO:1901363	Heterocyclic compound binding	139	0.16	0.00057	54	0.23	0.0067	52	0.24	0.0035
GO:0019899	Enzyme binding	67	0.24	0.0018	34	0.43	6.93E-05	27	0.36	0.0167

		CHO-K1			CHO-SK15-EPO			CHO-DP12		
Term ID	Process	Count	Strength	FDR	Count	Strength	FDR	Count	Strength	FDR
GO:0009987	Cellular process	322	0.1	110	15994	0.12	0.00026	106	0.13	7.61E-06
GO:0016043	Cellular component organization	132	0.17	49	5621	0.22	0.0279	48	0.24	0.007
GO:0071840	Cellular component organization or biogenesis	137	0.16	52	5887	0.23	0.0146	51	0.25	0.0031
GO:0006996	Organelle organization	92	0.2	39	3632	0.32	0.0098	41	0.37	0.00017

B

		CHO-K1			CHO-SK15-EPO			CHO-DP12		
KEGG	Description	Count	Strength	FDR	Count	Strength	FDR	Count	Strength	FDR
cge04530	Tight junction	16	0.88	3.36E-07	5	0.99	0.0328	6	1.44	3.44E-05
cge05012	Parkinson disease	16	0.64	9.61E-05	8	0.95	0.0015	4	1.03	0.0193
cge04670	Leukocyte transendothelial migration	10	0.78	0.00042	4	0.99	0.0328	4	1.37	0.0034
cge04971	Gastric acid secretion	6	0.77	0.0071	3	1.08	0.0498	3	1.46	0.0124
cge04810	Regulation of actin cytoskeleton	9	0.47	0.0226	5	0.82	0.0328	4	1.11	0.0143
cge05034	Alcoholism	7	0.51	0.0299	4	0.88	0.0498	4	1.26	0.0066

		CHO-K1			CHO-SK15-EPO			CHO-DP12		
Term ID	Function	Count	Strength	FDR	Count	Strength	FDR	Count	Strength	FDR
GO:0005488	Binding	400	0.16	1.62E-29	96	0.15	2.95E-05	44	0.19	0.003
GO:0005515	Protein binding	254	0.22	1.56E-17	69	0.26	9.24E-06	30	0.28	0.0194
GO:0019899	Enzyme binding	104	0.33	4.66E-11	30	0.4	0.0011	16	0.51	0.014

		CHO-K1			CHO-SK15-EPO			CHO-DP12		
Term ID	Process	Count	Strength	FDR	Count	Strength	FDR	Count	Strength	FDR
GO:0032879	Regulation of localization	108	0.37	5.99E-14	29	0.41	0.0021	16	0.53	0.0067
GO:0051179	Localization	165	0.26	4.58E-13	45	0.31	0.0017	22	0.38	0.0144
GO:0051234	Establishment of localization	146	0.28	3.63E-12	42	0.35	0.00072	21	0.43	0.0071
GO:0008104	Protein localization	91	0.37	1.81E-11	24	0.41	0.0147	16	0.61	0.0025
GO:0006810	Transport	137	0.27	7.83E-11	39	0.34	0.0021	20	0.43	0.0106
GO:0045184	Establishment of protein localization	63	0.43	1.07E-09	17	0.47	0.0358	12	0.7	0.0044
GO:0051641	Cellular localization	106	0.3	2.06E-09	NA	NA	NA	17	0.49	0.0092
GO:0051049	Regulation of transport	79	0.31	2.12E-07	24	0.41	0.0134	13	0.52	0.0389

On day 7, all cell lines show several enriched processes related to protein localisation in the cell, as well as protein transport. This may be linked to the localisation of proteins to the ER as part of the UPR, or the movement of ubiquitinated proteins out of the ER and suggests that the organisation within the cell is more established and requires less regulation than at day 4. While none of the shared KEGG pathways at day 7 link to ER stress, the prevalence of protein binding as a shared molecular function could be attributed to stress, as many UPR and ERAD proteins are activated by either the binding or un-binding of other proteins. Binding functions are also shared at day 4, further suggesting the role of binding in the ER stress response.

In all three cell lines, KEGG pathways, biological processes and molecular functions related to protein localisation and cellular organisation are enriched in ubiquitinated protein DE data. This suggests the importance of ubiquitin in the regulation of fundamental cell processes in mild hypothermic conditions. The biological processes "Cellular component organisation", "Cellular component organisation or biogenesis", and "Organelle organisation" (GO:0016043, GO:0071840, GO:0006996) were enriched in all three cell lines on day 4. These show the extent of ubiquitination in key cellular functions, as the cells adapt following temperature downshift. These processes are not found across all cell lines at day 7, only in CHO-K1. This suggests that cells may undergo some rearrangement of organelles and other subcellular components in response to mild hypothermia. A reduction in culture temperature has been shown to arrest cells in the G1 phase of the cell cycle (Becerra et al., 2012), which is the phase of the cell cycle during which the cell grows, copies organelles and accumulates nutrients (Csikász-Nagy et al., 2006). The interruption of these processes may require ubiquitin as a crucial regulatory factor in this process or to target the proteins no longer needed in these processes for degradation. A study has found that 64% of cells remain in the G1 phase following a temperature downshift (Shahabi et al., 2023). This correlates with the reduction in cell replication, suggesting that cells arrested in the G1 phase may focus resources on protein production rather than cell replication, with producer cells focusing on recombinant protein production.

Protein localisation is another essential biological process that is seen to be enriched in the differentially expressed ubiquitinated proteins. The movement of proteins to the relevant subcellular location is crucial for correct cell function. Cellular signalling pathways are highly dependent on ubiquitin for responses to extracellular influences, and have been shown to modulate growth, cell cycle regulation, receptor function, development and the stress response (Wilkinson, 1999).. A 2011 review by Grabbe et al. highlighted the critical role of ubiquitin in regulating protein activity and localisation over time. Using the NF-κB pathway

as an example, the review demonstrated how different ubiquitin species coordinate the activation of kinase complexes across distinct branches of the pathway (Grabbe et al. 2011). The enrichment of localisation pathways would indicate that the effect of an external temperature downshift requires changes in the cellular organisation, and those changes require mediation by ubiquitination.

3.4.4 Differential expression of ubiquitinated proteins following a downshift in culture temperature

In all three cell lines, at both time points, there are multiple proteins found in the whole cell proteome data that do not show up as ubiquitinated in that data but are found to be ubiquitinated in the enriched data, as shown in **Table 5**. All shared proteins can be found in **Supplementary Materials-1**. These proteins come from a range of cellular pathways, indicating the expansive role ubiquitination plays in the proteome. It also highlights the transient nature of some proteins, with their function only required for a brief period, then being targeted for degradation. Pathways relating to proteasome function (KEGG: 'Proteasome', Cellular component: 'Proteasome complex, Proteasome regulatory particle', Process: 'Proteasomal protein catabolic process') are enriched in common between whole cell and ubiquitinated DE proteins of CHO-DP12 and CHO-SK15-EPO, highlighting the role of ubiquitin in the regulation of ubiquitin-mediated degradation. Across all cell lines, essential biological processes, including cellular metabolic process ('GO:0044237') and primary metabolic process ('GO:0044238'), show the crucial role of ubiquitin in basic cellular function, and also demonstrates how standard proteomic studies can miss the post-translation modifications present on detected proteins that are essential to normal cell function.

Table 5: The number of shared proteins between the whole cell proteome differentially expressed data and the ubiquitinated differentially expressed data, as well as the number of these proteins that are oppositely regulated between the data. The presence of the double-glycine ubiquitin protein remnant was detected on none of these proteins in the whole cell proteome data, only when the enriched peptides were analysed.

Day 4			Day 7		
CHO-K1	**CHO-SK15-EPO**	**CHO-DP12**	**CHO-K1**	**CHO-SK15-EPO**	**CHO-DP12**
48	105	34	163	7	19
24	55	3	45	5	9

Only a minority of these common proteins are found at both day 4 and day 7 in the respective cell lines, so it is likely that those found at day 4 only are proteins that are no longer needed in cell activity and are then degraded. CHO-K1 shows the greatest number of shared ubiquitinated differentially expressed proteins between timepoints, suggesting the need for a long-term response to hypothermic stress than that seen in either of the producer cell lines. CHO-K1 also shows the most differentially expressed proteins in common between the whole cell proteome and ubiquitinated proteome on day 7, more than on day 4, whereas in both producer cell lines the number of shared ubiquitinated proteins decreases between the time points. This may be suggestive of a lesser need for long term ubiquitin-mediated degradation and regulation in the producer cell lines, suggesting a greater ability of the ER stress response in these cells to restore ER homeostasis than the non-producing CHO-K1 line.

There is a subset of these proteins where they are oppositely regulated between the whole cell and ubiquitinated data sets, the number of which is also included in **Table 5**. These proteins have a range of biological functions, and from a range of cellular locations. This suggests that these proteins exist within the cell in both ubiquitinated and non-ubiquitinated states, again highlighting the role of the ubiquitin in the regulation of biological processes,

and that the differing forms of these proteins may serve different biological functions. GJA1 is one of such differentially regulated proteins, being differentially regulated in both producer cell lines on day 4 (CHO-DP12: WCP= +5.1, Ubi= -1.5; CHO-SK15-EPO: WCP= -1.5, Ubi= +24.5). This gap junction protein has been linked to protein localisation (Bazzoun et al., 2019), cell-cell signalling and signal transduction (Fishman et al., 1990). It has been found that the ubiquitination of GJA1 is crucial to its degradation via autophagy, with E3 ligase NEDD4 being responsible for this ubiquitination (Bejarano et al., 2012). NEDD4 is found to be upregulated at 31°C in both CHO-K1 and CHO-DP12.

The greatest difference in expression is in reticulocalbin in CHO-SK15-EPO, on day 4. Down-regulated in the whole cell proteome data (-1.52-fold) and upregulated in the ubiquitinated data (+30.73-fold), reticulocalbin is a calcium-binding protein located in the lumen of the ER (Weis et al., 1994). This suggests that while reticulocalbin is more abundant at 37 °C, it predominantly exists in a non-ubiquitinated form. In contrast, when focusing solely on the ubiquitinated form, its levels are higher at 31 °C, indicating temperature-dependent regulation of its post-translational modification. Proteins exhibiting similar fold-change patterns in both whole-cell and ubiquitinated datasets may indicate that ubiquitination occurs uniformly across all protein molecules. In such cases, the observed changes in the ubiquitinated form likely reflect shifts in total protein abundance, rather than an increase in the extent of ubiquitination per molecule. The proteins that are regulated in the same direction in both data sets can also show some large differences in abundance. Plectin shows the largest difference, with a 68-fold increase at 31°C of the ubiquitinated form in CHO-K1 on day 7. This protein is part of the plakin family, which involved in maintaining cell integrity but also serve as scaffolding platforms for the regulation of signalling complexes (Sonnenberg & Liem, 2007). Plectin has been studied in C2C12 myoblasts, where it was found to promote proliferation and inhibit apoptosis, and may act to inhibit ubiquitination of the Dvl-2 protein, inhibiting autophagy (Yin et al., 2021). There

is a lack of research on the ubiquitination of plectin itself, although studies of skin, muscle and nerve tissues suggest that phosphorylation of this protein is an important part of some of its functions (Castañón et al., 2013). This difference in expression can be explained by either differing functionality of the ubiquitinated and non-ubiquitinated forms of the protein, or that at the time point studied, we are observing the end of the functional period of a protein, and so most copies of that protein are tagged for degradation.

3.4.5 Differential expression of ER stress proteins following a downshift in culture temperature

In this study, the pathway is further broken down into three areas, relating to the different mechanisms present within the ER, as detailed in our previous study (Ryan et al., 2023). These are *Protein Folding*; the ER function relating to protein synthesis, glycosylation, glucosylation and folding, *ERAD*; ER function related to the degradation of misfolded proteins and *Unfolded Protein Response*; ER function relating to the ER mechanism of coping with an accumulation of misfolded proteins. **Figure 3** details the fold changes of the differentially expressed proteins in each of these 3 areas. The use of three different cell lines, an antibody producer, a protein producer and a non-producer, shows the differing ER responses that can occur within cells of the same species, engineered for different purposes. The differences in cell line specific UPR and ERAD pathways may be part of the reason it has so far proven difficult to find a comprehensive remedy to ER stress across CHO cells.

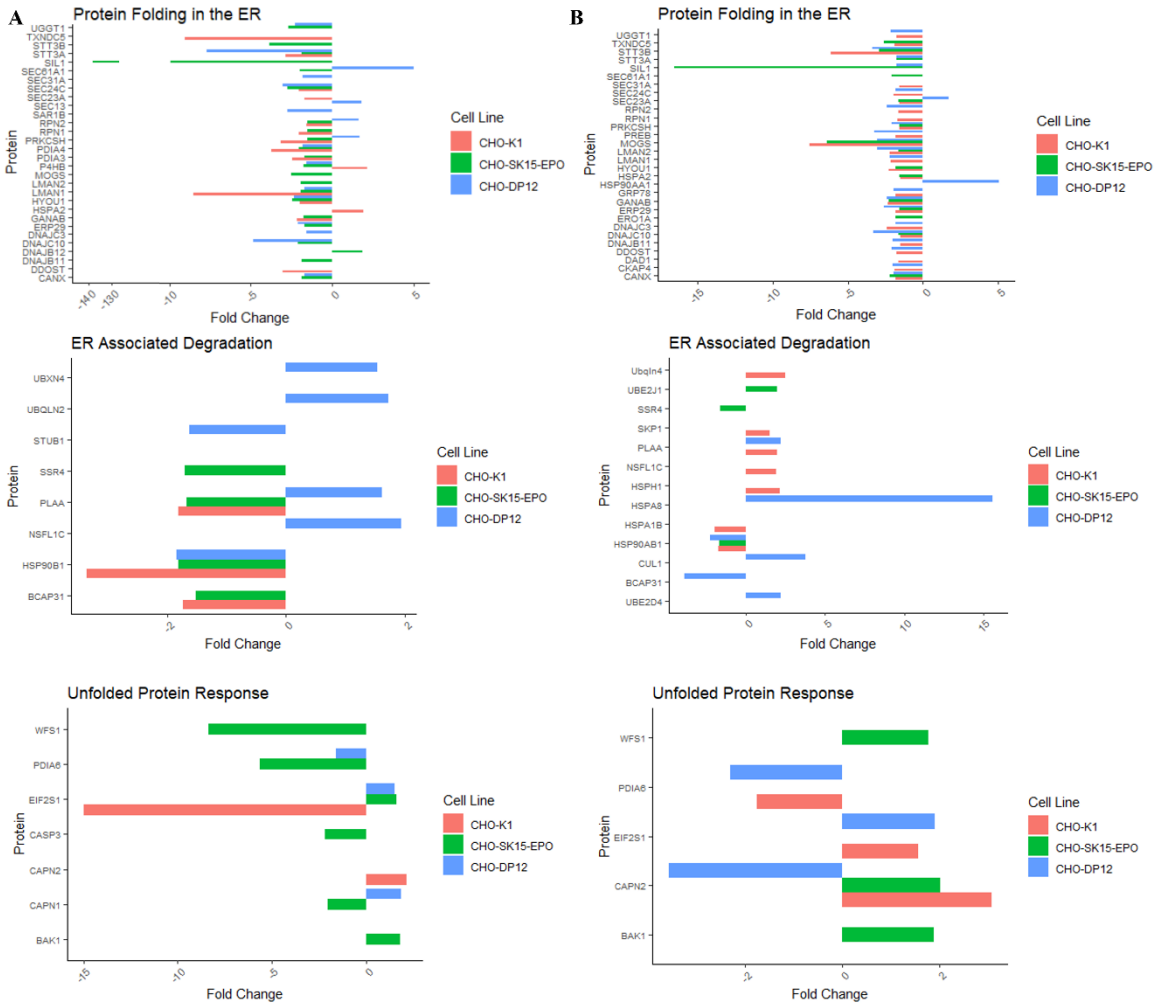


Figure 3. Fold changes of proteins in CHO-K1 (pink), CHO-SK15-EPO (green) and CHO-DP12 (blue) that are differentially expressed between cells culture at 31°C and cells cultured at 37°C, at **(A)** day 4 and **(B)** day 7. A minus value represents a protein that was less abundant in cells cultured at 31°C than in cells cultured at 37°C. In "Protein Folding in the ER" at day 4, where SEC61A1 is shown to have a +5-fold change in CHO-DP12, this value is for illustrative purposes, as this protein was only detected in the 31°C culture in this cell line, so has an "infinity" fold change between temperatures. In this same plot, an x-axis break is employed to represent the -136.5-fold change in the protein SIL1 in CHO-SK15-EPO, while keeping the other values within focus.

Protein Folding

Following protein translation, newly synthesised proteins enter the ER lumen, where these polypeptides undergo modifications such as glycosylation and disulfide bond formation, assisted by molecular chaperones and folding enzymes including BiP and several protein disulfide isomerases (PDI) (Okumura et al., 2015). Proper folding is critical for protein functionality, and misfolded proteins are targeted for degradation via the ERAD pathway to maintain cellular homeostasis. This processing ensures that only correctly folded proteins proceed to the Golgi apparatus for further maturation and secretion.

On day 4, there is down regulation of many proteins related to protein folding and synthesis at 31°C in CHO-DP12, CHO-K1 and CHO-SK15-EPO. This includes proteins involved in transport of nascent polypeptides into the ER. SEC61a1 is a component of the SEC61 channel-forming translocon complex that mediates transport of signal peptide-containing precursor polypeptides across the ER (Meacock et al., 2002a) and is downregulated in CHO-SK15-EPO cells at 31°C, but is up-regulated in CHO-DP12 at 31°C. Other transport proteins, SEC24C and SEC31α, are downregulated in CHO-DP12 at 31°C. SEC13 is needed for vesicle biogenesis for transport of proteins out of the ER (B. L. Tang et al., 1997) and is upregulated at 31°C in CHO-DP12, while SAR1B, which is a GTPase that functions in vesicle-mediated ER to Golgi transport (Cutrona et al., 2013), is downregulated at 31°C. SIL1 is required for protein translocation and folding in the ER, and functions as a nucleotide exchange factor for the ER lumenal chaperone HSPA5 (K. T. Chung et al., 2002). At day 4, there is a prominent fold change in SIL1 in CHO-SK15-EPO, being down-regulated 136.5-fold. SIL1 is a nucleotide exchange factor that has been linked with the release of BiP-clients in the ER lumen after sequestration with BiP, as well as suppression of PERK (Lemarié et al., 2021). Fellow HSP family member, HSPA2 is a molecular chaperone implicated in a wide variety of cellular processes including folding and transport of nascent polypeptides and the activation of proteolysis of misfolded proteins (Radons, 2016). It is upregulated in

CHO-K1 cells at 31°C. LMAN2 is downregulated in CHO-DP12 at 31°C and is important in the transportation of glycoproteins of unprocessed N-glycans from the Golgi to the ER, working with BiP in the ER (K. Yamamoto, 2021).

The glycosylation mechanism is also well represented in the differentially expressed proteins. Glucosidase GlcII (GANAB), required for cleaving glucose residues from immature glycoproteins, is a crucial part of the protein folding process and shows decreased expression at the lower temperature in CHO-K1 and CHO-SK-15-EPO but no differential expression between temperatures was found at day 4 in CHO-DP12. The lectin chaperone Calnexin and the luminal PDI it interacts with, ERP29, show a similar trend of decreased expression at 31°C compared to 37°C in CHO-SK15-EPO and CHO-DP12 but are not differentially expressed in CHO-K1. DDOST, which is part of the oligosaccharyl transferase (OST) complex involved in glycan transfer is down-regulated in CHO-K1 at 31°C (Ramírez et al., 2019). UGGT1 also works as a quality control checkpoint for protein folding by detecting minor folding defects in glycoproteins and re-glucosylating N-glycans allowing recognition by lectin-like molecular chaperones CNX and CRT (Ito et al., 2015), and similarly shows downregulation at 31°C in the producer cell lines CHO-DP12 and CHO-SK15-EPO. In addition to changes in the expression levels of GANAB and UGGT1, the other glycosyltransferases PRKCSH, RPN2, STT3a showed differential expression in all three cell lines and STT3a is downregulated at 31°C compared to 37°C in all three, whereas PRKCSH and RPN2 are downregulated at 31°C in CHO-K1 and CHO-SK15-EPO but upregulated at 31°C in CHO-DP12. Glycosyltransferases MOGS and STT3b, found only in CHO-SK15-EPO, and RPN1, found in CHO-SK15-EPO and CHO-K1 are all downregulated at 31°C. The downregulation of proteins related to folding and glucosylation in response to a temperature downshift is in line with the results reported by Bedoya-López et al (2016), who found a downregulation of UGGT1, MOGS, SSR1 and MAN1 (Bedoya-López et al., 2016). These results contrast with findings from other studies, who find that folding pathway

proteins are upregulated, and this is used to explain the increase in cell productivity (Torres et al., 2021; Yee et al., 2009). These differing results may be due to different cell lines, media, and the lower temperature of 31°C compared to 32/33°C. While the proteins involved in folding are downregulated, biological processes related to RNA activity are upregulated at 31°C (RNA splicing, mRNA processing, regulation of RNA splicing, mRNA metabolic processing).

Several PDIs were found to be downregulated in at least one of the cell lines at 31°C including P4HB, PDIA3 and PDIA4. The PDI family are crucial factors in both protein folding and quality control in the ER (Okumura et al., 2015). ERO1a is needed for proper function of PDIs and is similarly downregulated in CHO-SK15-EPO cells. DNAJB10, which is downregulated at 31°C, is an ER disulphide reductase required for efficient folding of proteins in the ER by catalysing the removal of non-native disulphide bonds formed during the folding of proteins (Cunnea et al., 2003). DNAJB11 is a co-chaperone for HSPA5 and it is required for proper folding, trafficking or degradation of proteins (Yu et al., 2000). No DNAJ-protein family members involved in protein folding were found to be differentially expressed on day 4 in CHO-K1. The DNAJ-protein family is a subset of heat shock proteins that are involved in the assembly and disassembly of macromolecular complexes (Stetler et al., 2010). The functional diversity of HSP70 proteins is reliant on the co-chaperone functions of DNAJ proteins (Hageman & Kampinga, 2009). DNAJC3 and DNAJC10, as well as stress response protein HYOU1 (Ozawa et al., 1999), are all downregulated in CHO-DP12 cells cultured at 31°C. It appears that in response to the arrest of cells in the G1 phase, the CHO cell lines reduce protein processing through the ER, possibly focusing on a reduction of the processing of cell cycle and growth-related proteins that are not needed in the G1 phase.

On day 7, the three cell lines studied showed similar trends to day 4, with protein transport, folding and glycosylation related proteins downregulated at 31°C. UGGT1, PRKCSH,

MOGS, STT3b, PDIA3, PDIA4, DNAJC10, LMAN1, ERP29 and CANX are all down regulated at 31°C in all cell lines. DNAJC3 and LMAN2 are downregulated in CHO-DP12. HSPA2 and HYOU1 are downregulated in CHO-SK15-EPO and CHO-K1. GANAB, STT3α and SIL1 are all downregulated in both CHO-SK15-EPO and CHO-DP12, while glycotransferases DDOST and RPN2 (Ramírez et al., 2019) all downregulated in both CHO-K1 and CHO-DP12, and RPN1 and DAD1 are downregulated in only CHO-K1. The chaperones BIP and DNAJB11 are both downregulated at 31°C in CHO-K1 and CHO-DP12. HSP90AA1 aids in the proper folding of specific target proteins through specific ATPase activity (Forsythe et al., 2001), and it is upregulated at 31°C in CHO-DP12. This suggests that these cell lines are processing less proteins overall at 31°C than at 37°C, possibly due to the cells needing less proteins for cell replication, as reflected by the reduced growth rate at 31°C.

SEC23α is the only transport protein differentially expressed in all three lines on day 7, downregulated in CHO-K1 and CHO-SK15-EPO but upregulated in CHO-DP12, with SEC24C downregulated in both CHO-K1 and CHO-DP12, SEC61α1 downregulated in CHO-SK15-EPO and SEC31α downregulated in CHO-K1. SEC12 (PREB) is part of the regulation of coat protein complex II/COPII in ER to Golgi vesicle-mediated transport (Melville et al., 2020) and is downregulated at 31°C in CHO-DP12. The transmembrane protein CKAP4, which facilitates microtubule anchoring to the ER (Vedrenne et al., 2005), is downregulated at 31°C in CHO-DP12. TXNDC5 introduces disulfide bonds into unfolded substrates rapidly, and then allows for more precise PDI-mediated disulfide bond formation (Okumura et al., 2015), and is downregulated in both CHO-K1 and CHO-SK15-EPO. This data suggests that the stress of temperature downshift affects the glycosylation ability of CHO cells, even at day 7, five days after the shift. The downregulation of so many proteins involved in the proper folding of nascent polypeptides would result in a mass build-up of misfolded proteins, triggering the ERAD and UPR. Due to the limitations of bottom-up

proteomics, it is not possible to confirm which proteins may be misfolded and if they relate to specific processes, functions or pathways. While ubiquitinated peptides may come from proteins targeted for degradation, it cannot be confirmed that these are from misfolded proteins, and not just proteins no longer useful to cell activity. It is also impossible to rule out the regulatory function of ubiquitination on these proteins, due to a lack of research on ubiquitination in CHO cells.

There are few proteins common to both producer cell lines only, and not in CHO-K1, with more being found in only one producer cell line. This indicates that CHO-SK15-EPO and CHO-DP12 may have developed different responses to the stress of recombinant protein production, impacting their response to mild hypothermic shock. On day 4, neither producer cell line show Sec23α as differentially expressed, it is only downregulated (1.6-fold) at 31°C in CHO-K1. On day 7 however, CHO-SK15-EPO shows a similar response as CHO-K1, downregulating this protein at 31°C (1.6-fold and 1.5-fold, respectively). On day 7, CHO-DP12 upregulates Sec23α (1.7-fold). This protein is part of the coat protein complex II (COPII) which promotes the formation of transport vesicles from the ER, for moving proteins out of the ER (Paccaud et al., 1996a). It is possible that CHO-DP12 may use an altered COPII to transport the increased IgG1 antibody being produced by cells at 31°C. Other proteins involved in transport (Sec24, Sec12, CKAP4) are downregulated at 31°C in CHO-DP12 but are not differentially expressed in CHO-SK15-EPO, suggesting that lower culture temperature does not illicit the same changes in transport pathways in these cells as in CHO-DP12. SIL1 (136-fold) is downregulated at 31°C in only CHO-SK15-EPO on day 4, and is still downregulated at day 7, but to a lesser extent (16-fold). On day 7 it also downregulated in CHO-DP12 at 31°C (1.7-fold) but is not differentially expressed in CHO-K1 at either time point, suggesting that the downregulation of this protein may play a role in the increased productivity of producer cells at 31°C. Given the role of SIL1 in the release of BiP-clients, as well as suppression of PERK (Lemarié et al., 2021), there is an indication

that producer cells, in particular CHO-SK15-EPO, may use this protein as a major part of their normal protein processing, and in response to a build-up of misfolded proteins, it must be downregulated to allow for an appropriate ER stress response.

On day 4, there are 6 DE proteins shared between the three cell lines, all downregulated at 31°C (ZC3HAV1, USP5, PI4KA, ARL2, PSMD13, CCT7). There is 9 DE proteins shared between all three cell lines on day 7. 4 of these proteins are upregulated at 31°C in all cell lines (GJA1, PSMD2, UBA1, HUWE1). Three are upregulated at 31°C in CHO-K1 and CHO-SK15-EPO but downregulated at 31°C in CHO-DP12 (GNAI3, MAP1B, PPP2R1A). RDX is downregulated at 31°C in CHO-K1 and CHO-DP12 but upregulated in CHO-SK15-EPO, and USP5 is upregulated at 31°C in CHO-K1 and CHO-DP12 but downregulated in CHO-SK15-EPO. USP5 is the only protein found in all cell lines differentially expressed at both day 4 and day 7. It is a ubiquitin-specific protease, involved in removing ubiquitin from the proximal end of the polyubiquitin chains not conjugated to a target protein, therefore maintaining the free ubiquitin pool (Clague et al., 2019).

ERAD

The ERAD pathway plays a crucial role in maintaining protein quality control, especially in CHO cells producing recombinant proteins. As these cells are engineered to express high levels of foreign proteins, the ER is put under extra pressure. ERAD recognises misfolded or improperly modified proteins in the ER and targets them for retro-translocation into the cytosol, where they are ubiquitinated and degraded by the proteasome. This process helps prevent the accumulation of defective proteins, mitigating the risk of apoptosis in cells, as well as contributing to quality of recombinant proteins produced.

On day 4, both Endoplasmin (HSP90AB1) and PLAA are differentially expressed at 31°C in all cell lines, with Endoplasmin being downregulated in all, whereas PLAA is downregulated in CHO-K1 and CHO-SK15-EPO but upregulated in CHO-DP12. BCAP31

is a chaperone protein that is one of the most abundant ER proteins and is involved in the export of secreted proteins in the ER, the recognition of abnormally folded proteins and their targeting to the ERAD (Annaert et al., 1997). It is downregulated in CHO-SK15-EPO but upregulated in CHO-K1 at 31°C. SSR4 is downregulated in CHO-SK15-EPO at 31°C.

Several proteins related to ubiquitin's role in the ERAD pathway are differentially expressed on day 4 in only CHO-DP12 cells. NSFL1c, also known as p47, has a C-terminal domain which includes a UBX domain (Yuan et al., 2001) and is known to interact with VCP (Uchiyama et al., 2003). UBXD2 is a membrane protein of the ER that binds VCP and is crucial in the promotion of ERAD (Lim et al., 2009). UBQLN2 is a vital regulator of different protein degradation pathways including the ubiquitin-proteasome system (UPS), autophagy and the ERAD pathway. It allows for proteasomal targeting of misfolded proteins for degradation by binding to their polyubiquitin chains and by interacting with subunits in the proteasome (Kleijnen et al., 2000). NSFLC1, UBXD2 and UBQLN2 are all upregulated at 31°C. STUB1, an E3 ubiquitin-protein ligase, is involved in targeting misfolded proteins towards proteasomal degradation (Ballinger et al., 1999), and is downregulated at 31°C in CHO-DP12. This indicates that there may be a temporal aspect to ERAD activity, given that most of the DE proteins from that pathway are upregulated in CHO-DP12, but STUB1 is downregulated, suggesting that not all proteins are required throughout the process. The presence of these proteins as differentially expressed in CHO-DP12 only highlights the differences between CHO cells line in their biological processes and response to environmental factors. CHO-SK15-EPO is a producer cell line derived directly from CHO-K1, which may a reason why CHO-DP12 has a different ERAD response to these two cell lines.

On day 7, only Endoplasmin shows differential expression in all cell lines, being downregulated. Endoplasmin engages with a wide variety of client proteins, and is hence, involved in several different cell functions (Retzlaff et al., 2009), and can be linked to

transport of cytosolic protein peptides through the ERGIC (endoplasmic reticulum-Golgi intermediate compartment) (M. Zhang et al., 2020). In CHO-SK15-EPO, only one other protein involved in ERAD is differentially expressed, UBE2J1, which is upregulated at 31°C. UBE2J1 catalyses the covalent attachment of ubiquitin to other proteins and allows for the selective degradation of misfolded proteins from the ER and is essential for cells to recover from ER stress (Elangovan et al., 2017). This is suggestive of a much-reduced ERAD response in CHO-SK15-EPO by day 7, again highlighting the between cell lie differences in response to temperature downshift. On day 7, BCAP31 is only differentially expressed in CHO-DP12, being downregulated at 31°C. HSPA8, which is upregulated at 31°C in CHO-DP12, is a versatile molecular chaperone involved in many cellular processes, including proteolysis of misfolded proteins, assembly of protein complexes, and is critical in the protein quality control system (Grove et al., 2011; Y. Yamamoto et al., 2010).

In CHO-K1 cells, PLAA, NSFL1c, and SKP1 were all upregulated at 31°C versus 37°C on day 7. PLAA is involved in protein ubiquitination, sorting and degradation through its association with VCP (Papadopoulos et al., 2017). SKP1 is one component of the SCF (SKP1-CUL1-F-box protein) ubiquitin ligase complex, which mediates ubiquitination of proteins (Man et al., 2015). Another component of this complex, CUL1, is upregulated in CHO-DP12 at 31°C. PLAA continues to be upregulated at 31°C in CHO-DP12. UBE2D4 is a ligase that enables ubiquitin conjugating enzyme activity, allowing protein polyubiquitination (David et al., 2010) and is also upregulated at 31°C in CHO-DP12. The upregulation of these proteins is suggestive of an increase in ubiquitination, likely due to misfolded protein build up. The shift from downregulation of PLAA on day 4 at 31°C in CHO-K1, to upregulation at 31°C on day 7 supports this conclusion given its function to help transfer ubiquitinated proteins to the proteasome for degradation (Qiu et al., 2010). The increase in number of proteins being differentially expressed at 31°C on day 7 compared to day 4 also indicates that the number of proteins requiring degradation is greater at day 7.

This is also seen in the number of differentially expressed ubiquitinated proteins seen on day 7, in particular in CHO-K1.

The upregulation of two ERAD proteins at 31°C, UBXD2 (1.5-fold) and UBQLN2 (1.7-fold), on day 4 in CHO-DP12, when neither protein is differentially expressed in CHO-K1 or CHO-SK15-EPO could suggest that the mild hypothermic conditions cause a build-up of misfolded proteins faster in CHO-DP12 than the other cell lines, requiring a more pronounced ERAD response than the other cell lines. Evidence for this is shown by 8 ubiquitinated DE proteins in the proteasome KEGG pathway in CHO-DP12 at day 4, and 17 DE proteins in that pathway form the whole cell data, more than in either CHO-K1 or CHO-SK15-EPO, again indicating that the difference in cell line lineage may cause differences in the ERAD pathway. There are also 15 DE proteins from the whole cell proteome identified in the KEGG pathway "Ubiquitin mediated proteolysis" pathway (cge04120).

UPR

The UPR is a key adaptive mechanism activated in response to the accumulation of misfolded or unfolded proteins in the ER. When recombinant proteins are overexpressed, the ER becomes stressed, triggering the UPR to restore protein homeostasis. The UPR works through three main signalling pathways-IRE1, PERK, and ATF6-that collectively reduce protein synthesis, enhance the expression of molecular chaperones, and promote protein degradation via ERAD (Stolz & Wolf, 2010). This balance helps cells cope with the increased protein load, ensuring optimal folding conditions and preventing cell death during large-scale production of recombinant proteins.

On day 4, PDIA6 is downregulated at 31°C in CHO-K1 and CHO-DP12. PDIA6 can limit UPR activity through inhibition of the crucial UPR activator, IRE1α, by converting oligomeric IRE1α to a monomeric form via a process of forming and breaking of disulfide bonds (Eletto et al., 2014). The mitochondrial apoptogenic protein, BAK1, has the ability to

localise to the ER during times of ER stress, allowing a progressive depletion of ER Ca^{2+} and induction of caspase 12 cleavage and has been proven to induce apoptosis through depletion of Ca^{2+} allowing for the caspase cascade to cause cell death (Zong et al., 2003). It is upregulated in CHO-SK15-EPO at 31°C. EIF2S1 is upregulated at 31°C in CHO-SK15-EPO and CHO-DP12 but downregulated in CHO-K1. EIF2α is associated with UPR-induced apoptosis, with studies linking EIF2α to caspase cleavage and possibly the caspase cascade (He et al., 2019; Nagata et al., 2007). Calpain 2 (CAPN2) is upregulated in CHO-DP12 at 31°C. EIF2α's function as a transcription factor can be limited during periods of ER stress, by being phosphorylated by the kinase PERK, limiting general protein translation in the cell, with the phosphorylation of EIF2α also inducing ATF4 mediated branch of the UPR (Cnop et al., 2017).

On day 7, PDIA6 continues to be downregulated in CHO-K1 and CHO-DP12 cells at 31°C. BAK1 shows a similar level of upregulation at 31°C in CHO-SK15-EPO as it did on day 4. EIF2S1 continues to be upregulated at 31°C in CHO-DP12 but is not differentially expressed CHO-SK15-EPO and CHO-K1 at day 7. The transmembrane protein Wolframin (WFS1) is downregulated on both day 4 and day 7 in CHO-SK15-EPO at 31°C, but with a greater fold-change at day 4 (8.38-fold) than day 7 (1.77-fold). It has previously been shown to be a UPR component that mitigates ER stress response in cells (Fonseca et al., 2005). CAPN2 is upregulated in CHO-SK15-EPO and CHO-K1 at 31°C on day 7, while CAPN1 is downregulated in CHO-DP12. Calpains are a family of Ca^{2+}-dependent intracellular cysteine proteases that cleave protein substrates, in a fashion that links them to a wide variety of biological functions including cell migration, cell cycle regulation, differentiation, and apoptosis (Goll et al., 2003). The upregulation of proteins such as BAK1 and CAPN2 and the downregulation of PDIA6 and WSF1 suggests that the UPR is engaging some apoptotic activity within both lines at 31°C on both day 4 and day 7 with no fold-change observed in proteins indicating any attempts to mitigate apoptosis, and the KEGG pathway 'Apoptosis'

enriched in whole cell proteome data in all three cell lines. Caspase-dependent apoptosis seems to be the CHO cell response when the UPR has been unsuccessful in mediating the build-up of misfolded proteins following temperature shift.

3.5 Conclusions

At 31 °C, both producer cell lines exhibit enhanced specific productivity compared to 37 °C. This increase in final titre is driven by prolonged culture viability under hypothermic conditions, where slower growth extends the production window and ultimately boosts final yield, with the increase in specific productivity likely driven by increased mRNA stability or enhanced recombinant gene transcription. The type of recombinant product being produced can affect the mechanisms of ER stress response, with the IgG producer and EPO producer having different expression of transport proteins at 31°C, including several secretory proteins (SEC12, SEC23, SEC24). The inclusion of the ubiquitinated proteome in this extensive profiling study provides new insights into the stress pathways of CHO cells, in response to an industrially relevant culture method. Enriching ubiquitinated peptides prior to mass spectrometry not only reveals proteins that are undetectable in whole-cell lysates but also uncovers ubiquitination events on proteins previously identified without any indication of post-translational modification. The prevalence of pathways related to protein binding, enzyme binding and organic cyclic compound binding in the ubiquitinated proteome indicates that these pathways are under tight cellular regulation during mild hypothermic stress. This is also seen in pathways related to cellular organisation on day 4, suggesting that proteins needed to redirect cell mechanisms during mild hypothermia are tightly regulated. Moreover, it may reflect the targeted degradation of proteins that have fulfilled their roles, indicating that by 48 hours post-temperature shift, cells may have already adapted and are actively clearing proteins no longer required. The differing expression of proteins between the whole cell proteome and ubiquitinated proteome shows that modified forms of proteins are differently regulated, and both forms of proteins play a role in the cellular response to

hypothermic stress. The ubiquitinated proteome shows a more in-depth view of the ER, and of the role of ubiquitin in the regulation of the cell's response to mild hypothermia. It has been found that proteins needed for correct ubiquitin regulation in the ER can themselves be subject to ubiquitination and given that this was seen in all three cell lines, this is likely a fundamental part of the CHO cell stress response. Furthermore, the data indicates that ubiquitination serves as a key regulatory strategy throughout different stages of cell culture, enabling precise control over protein function and turnover during adaptation.

This study underscores the importance of ubiquitin-dependent regulation in cellular adaptation to hypothermia and the data generated is a valuable resource for investigating the molecular underpinnings of CHO cell responses to environmental stress, highlighting the need for post-translational modification inclusive studies to truly understand proteome activity. These insights may guide the development of strategies to enhance CHO cell resilience and productivity in bioprocessing applications. The potential roles of proteins found in this study can be explored through cell line adaptation to under- or over-express protein targets of interest. A genome-wide CRISPR/Cas9 screening platform developed for CHO-K1 cells identified 180 genes linked to resistance to hyperosmotic stress, and two novel target genes from that screen, *Zfr* and *Pnp*, were then knocked out in rCHO cells and a corresponding increased resistance to hyperosmotic stress was observed, improving mAb production (S. H. Kim et al., 2023). In three IgG producing CHO cell lines, the over-expression of Sar1A, which is needed for COPII vesicle formation, antibody production was increased by the modification of the secretion process (Tsunoda et al., 2024). These studies highlight the need for in-depth profiling of CHO stress responses, to uncover novel targets for genetic engineering. ER stress marker proteins have also been used to develop reporter systems, for the detection of real-time changes in the UPR. Both BiP and IRE1 were used to create fluorescent monitoring systems in producer CHO cell lines, producing a GSR mAb and IgG, respectively (Kyeong & Lee, 2022; Roy et al., 2017). These systems could aid in

engineering strategies and cell line selection for improved production of biotherapeutics, and our results allow for a greater insight into potential targets for further development of such reporter systems.

To build on the findings of this study, future research should include profiling the ubiquitinated proteome of other CHO cell lines and exploring the functional consequences of specific ubiquitination events and their potential for targeted modulation to optimise biomanufacturing outcomes.

3.6 References

Aber, R., Chan, W., Mugisha, S., & Jerome-Majewska, L. A. (2019). Transmembrane emp24 domain proteins in development and disease. *Genetics Research*, *101*, e14. https://doi.org/10.1017/S0016672319000090

Aebi, M. (2013). N-linked protein glycosylation in the ER. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, *1833*(11), 2430-2437.

Ahlstedt, B. A., Ganji, R., & Raman, M. (2022). The functional importance of VCP to maintaining cellular protein homeostasis. *Biochemical Society Transactions*, *50*(5), 1457-1469. https://doi.org/10.1042/BST20220648

Annaert, W. G., Becker, B., Kistner, U., Reth, M., & Jahn, R. (1997). Export of Cellubrevin from the Endoplasmic Reticulum Is Controlled by BAP31. *The Journal of Cell Biology*, *139*(6), 1397-1410.

Arnold, P. K., & Finley, L. W. S. (2022). Regulation and function of the mammalian tricarboxylic acid cycle. *The Journal of Biological Chemistry*, *299*(2), 102838. https://doi.org/10.1016/j.jbc.2022.102838

Ballinger, C. A., Connell, P., Wu, Y., Hu, Z., Thompson, L. J., Yin, L. Y., & Patterson, C. (1999). Identification of CHIP, a novel tetratricopeptide repeat-containing protein that interacts with heat shock proteins and negatively regulates chaperone functions. *Molecular and Cellular Biology*, *19*(6), 4535-4545. https://doi.org/10.1128/MCB.19.6.4535

Bazzoun, D., Adissu, H. A., Wang, L., Urazaev, A., Tenvooren, I., Fostok, S. F., Chittiboyina, S., Sturgis, J., Hodges, K., Chandramouly, G., Vidi, P.-A., Talhouk, R. S., & Lelièvre, S. A. (2019). Connexin 43 maintains tissue polarity and regulates mitotic spindle orientation in the breast epithelium. *Journal of Cell Science*, *132*(10), jcs223313. https://doi.org/10.1242/jcs.223313

Becerra, S., Berrios, J., Osses, N., & Altamirano, C. (2012). Exploring the effect of mild hypothermia on CHO cell productivity. *Biochemical Engineering Journal*, *60*, 1-8. https://doi.org/10.1016/j.bej.2011.10.003

Bedard, K., Szabo, E., Michalak, M., & Opas, M. (2005). Cellular Functions of Endoplasmic Reticulum Chaperones Calreticulin, Calnexin, and ERp57. In *International Review of Cytology* (Vol. 245, pp. 91-121). Academic Press. https://doi.org/10.1016/S0074-7696(05)45004-4

Bedoya-López, A., Estrada, K., Sanchez-Flores, A., Ramírez, O. T., Altamirano, C., Segovia, L., Miranda-Ríos, J., Trujillo-Roldán, M. A., & Valdez-Cruz, N. A. (2016). Effect of Temperature Downshift on the Transcriptomic Responses of Chinese Hamster Ovary Cells Using Recombinant Human Tissue Plasminogen Activator Production Culture. *PLOS ONE*, *11*(3), e0151529. https://doi.org/10.1371/journal.pone.0151529

Bejarano, E., Girao, H., Yuste, A., Patel, B., Marques, C., Spray, D. C., Pereira, P., & Cuervo, A. M. (2012). Autophagy modulates dynamics of connexins at the plasma membrane in a ubiquitin-dependent manner. *Molecular Biology of the Cell*, *23*(11), 2156-2169. https://doi.org/10.1091/mbc.e11-10-0844

Bengtsson, L., & Otto, H. (2008). LUMA interacts with emerin and influences its distribution at the inner nuclear membrane. *Journal of Cell Science*, *121*(4), 536-548. https://doi.org/10.1242/jcs.019281

Bettigole, S. E., & Glimcher, L. H. (2015). Endoplasmic reticulum stress in immunity. *Annual Review of Immunology*, *33*, 107-138. https://doi.org/10.1146/annurev-immunol-032414-112116

Bloom, J., Amador, V., Bartolini, F., DeMartino, G., & Pagano, M. (2003). Proteasome-mediated degradation of p21 via N-terminal ubiquitinylation. *Cell*, *115*(1), 71-82.

Bonnon, C., Wendeler, M. W., Paccaud, J.-P., & Hauri, H.-P. (2010). Selective export of human GPI-anchored proteins from the endoplasmic reticulum. *Journal of Cell Science*, *123*(Pt 10), 1705-1715. https://doi.org/10.1242/jcs.062950

Bryan, L., Henry, M., Kelly, R. M., Lloyd, M., Frye, C. C., Osborne, M. D., Clynes, M., & Meleady, P. (2021). Global phosphoproteomic study of high/low specific productivity industrially relevant mAb producing recombinant CHO cell lines. *Current Research in Biotechnology*, *3*, 49-56. https://doi.org/10.1016/j.crbiot.2021.02.002

Capiaumont, J., Legrand, C., Carbonell, D., Dousset, B., Belleville, F., & Nabet, P. (1995). Methods for reducing the ammonia in hybridoma cell cultures. *Journal of Biotechnology*, *39*(1), 49-58.

Carvalho, P., Stanley, A. M., & Rapoport, T. A. (2010). Retro-translocation of a misfolded luminal ER protein by the ubiquitin-ligase Hrd1p. *Cell*, *143*(4), 579-591. https://doi.org/10.1016/j.cell.2010.10.028

Castañón, M. J., Walko, G., Winter, L., & Wiche, G. (2013). Plectin-intermediate filament partnership in skin, skeletal muscle, and peripheral nerve. *Histochemistry and Cell Biology*, *140*(1), 33-53. https://doi.org/10.1007/s00418-013-1102-0

Castellano, B. M., Tang, D., Marsters, S., Lam, C., Liu, P., Rose, C. M., Sandoval, W., Ashkenazi, A., Snedecor, B., & Misaghi, S. (2023). Activation of the PERK branch of the unfolded protein response during production reduces specific productivity in CHO cells via downregulation of PDGFRa and IRE1a signaling. *Biotechnology Progress*, e3354.

Chen, P., & Harcum, S. W. (2005). Effects of amino acid additions on ammonium stressed CHO cells. *Journal of Biotechnology*, *117*(3), 277-286. https://doi.org/10.1016/j.jbiotec.2005.02.003

Chen, P., & Harcum, S. W. (2006a). Effects of elevated ammonium on glycosylation gene expression in CHO cells. *Metabolic Engineering*, *8*(2), 123-132. https://doi.org/10.1016/j.ymben.2005.10.002

Chen, P., & Harcum, S. W. (2006b). Effects of elevated ammonium on glycosylation gene expression in CHO cells. *Metabolic Engineering*, *8*(2), 123-132. https://doi.org/10.1016/j.ymben.2005.10.002

Chen, P., & Harcum, S. W. (2007). Differential display identifies genes in chinese hamster ovary cells sensitive to elevated ammonium. *Applied Biochemistry and Biotechnology*, *141*(2), 349-359. https://doi.org/10.1007/BF02729072

Chitwood, D. G., Wang, Q., Elliott, K., Bullock, A., Jordana, D., Li, Z., Wu, C., Harcum, S. W., & Saski, C. A. (2021a). Characterization of metabolic responses, genetic variations, and microsatellite instability in ammonia-stressed CHO cells grown in fed-batch cultures. *BMC Biotechnology*, *21*(1), 4. https://doi.org/10.1186/s12896-020-00667-2

Chitwood, D. G., Wang, Q., Elliott, K., Bullock, A., Jordana, D., Li, Z., Wu, C., Harcum, S. W., & Saski, C. A. (2021b). Characterization of metabolic responses, genetic variations, and microsatellite instability in ammonia-stressed CHO cells grown in fed-batch cultures. *BMC Biotechnology*, *21*(1), 4. https://doi.org/10.1186/s12896-020-00667-2

Chung, K. T., Shen, Y., & Hendershot, L. M. (2002). BAP, a Mammalian BiP-associated Protein, Is a Nucleotide Exchange Factor That Regulates the ATPase Activity of BiP*. *Journal of Biological Chemistry*, *277*(49), 47557-47563. https://doi.org/10.1074/jbc.M208377200

Chung, M.-I., Lim, M.-H., Lee, Y.-J., Kim, I.-H., Kim, I.-Y., Kim, J.-H., Chang, K.-H., & Kim, H.-J. (2003). Reduction of Ammonia Accumulation and Improvement of Cell Viability by Expression of Urea Cycle Enzymes in Chinese Hamster Ovary Cells. *Journal of Microbiology and Biotechnology*, *13*(2), 217-224.

Clague, M. J., Urbé, S., & Komander, D. (2019). Breaking the chains: Deubiquitylating enzyme specificity begets function. *Nature Reviews Molecular Cell Biology*, *20*(6), 338-352.

Cnop, M., Toivonen, S., Igoillo-Esteve, M., & Salpea, P. (2017). Endoplasmic reticulum stress and eIF2α phosphorylation: The Achilles heel of pancreatic β cells. *Molecular Metabolism*, *6*(9), 1024-1039. https://doi.org/10.1016/j.molmet.2017.06.001

Connolly, D. J., O'Neill, L. A., & McGettrick, A. F. (2013). The GOLD domain-containing protein TMED1 is involved in interleukin-33 signaling. *Journal of Biological Chemistry*, *288*(8), 5616-5623.

Costello, A., Coleman, O., Lao, N. T., Henry, M., Meleady, P., Barron, N., & Clynes, M. (2018). Depletion of endogenous miRNA-378-3p increases peak cell density of CHO DP12 cells and is correlated with elevated levels of ubiquitin carboxyl-terminal hydrolase 14. *Journal of Biotechnology*, *288*, 30-40.

Costello, A., Lao, N. T., Barron, N., & Clynes, M. (2019). Continuous translation of circularized mRNA improves recombinant protein titer. *Metabolic Engineering*, *52*, 284-292.

Csikász-Nagy, A., Battogtokh, D., Chen, K. C., Novák, B., & Tyson, J. J. (2006). Analysis of a Generic Model of Eukaryotic Cell-Cycle Regulation. *Biophysical Journal*, *90*(12), 4361-4379. https://doi.org/10.1529/biophysj.106.081240

Cunnea, P. M., Miranda-Vizuete, A., Bertoli, G., Simmen, T., Damdimopoulos, A. E., Hermann, S., Leinonen, S., Huikko, M. P., Gustafsson, J.-Å., Sitia, R., & Spyrou, G. (2003). ERdj5, an Endoplasmic Reticulum (ER)-resident Protein Containing DnaJ and Thioredoxin Domains, Is Expressed in Secretory Cells or following ER Stress*. *Journal of Biological Chemistry*, *278*(2), 1059-1066. https://doi.org/10.1074/jbc.M206995200

Cutrona, M. B., Beznoussenko, G. V., Fusella, A., Martella, O., Moral, P., & Mironov, A. A. (2013). Silencing of mammalian Sar1 isoforms reveals COPII-independent protein sorting and transport. *Traffic (Copenhagen, Denmark)*, *14*(6), 691-708. https://doi.org/10.1111/tra.12060

Dahodwala, H., Hodzic, I., Slesarev, A., Cutak, B., Kuzin, A., Lal, R., Liu, J., Mahon, J., Narasimhan, R. L., Onuska, J., Ravellette, J., Reger, K., Sakurada, S., Stewart, F., Borgschulte, T., Cote, C., Lee, K. H., O'Dell, W. B., Kelman, Z., & Anderson, B. A. (2025). Development and Characterization of the NISTCHO Reference Cell Line. *Biotechnology Journal*, *20*(5), e70012. https://doi.org/10.1002/biot.70012

Dasarathy, S., & Hatzoglou, M. (2018). Hyperammonemia and proteostasis in cirrhosis. *Current Opinion in Clinical Nutrition and Metabolic Care*, *21*(1), 30-36. https://doi.org/10.1097/MCO.0000000000000426

David, Y., Ziv, T., Admon, A., & Navon, A. (2010). The E2 ubiquitin-conjugating enzymes direct polyubiquitination to preferred lysines. *The Journal of Biological Chemistry*, *285*(12), 8595-8604. https://doi.org/10.1074/jbc.M109.089003

Denic, V., Quan, E. M., & Weissman, J. S. (2006). A luminal surveillance complex that selects misfolded glycoproteins for ER-associated degradation. *Cell*, *126*(2), 349-359.

Dorner, A. J., Wasley, L. C., & Kaufman, R. J. (1989). Increased synthesis of secreted proteins induces expression of glucose-regulated proteins in butyrate-treated Chinese hamster ovary cells. *Journal of Biological Chemistry*, *264*(34), 20602-20607.

Dreger, M., Bengtsson, L., Schöneberg, T., Otto, H., & Hucho, F. (2001). Nuclear envelope proteomics: Novel integral membrane proteins of the inner nuclear membrane. *Proceedings of the National Academy of Sciences*, *98*(21), 11943-11948. https://doi.org/10.1073/pnas.211201898

Elahian, F., Sepehrizadeh, Z., Moghimi, B., & Mirzaei, S. A. (2014). Human cytochrome b5 reductase: Structure, function, and potential applications. *Critical Reviews in Biotechnology*, *34*(2), 134-143.

Elangovan, M., Chong, H. K., Park, J. H., Yeo, E. J., & Yoo, Y. J. (2017). The role of ubiquitin-conjugating enzyme Ube2j1 phosphorylation and its degradation by proteasome during endoplasmic stress recovery. *Journal of Cell Communication and Signaling*, *11*(3), 265-273. https://doi.org/10.1007/s12079-017-0386-6

Eletto, D., Eletto, D., Dersh, D., Gidalevitz, T., & Argon, Y. (2014). Protein disulfide isomerase A6 controls the decay of IRE1α signaling via disulfide-dependent association. *Molecular Cell*, *53*(4), 562-576.

Elliott, K., Harris, G., Harcum, S., Blakeman, K., Gavin, C., & Anderson, J. Y. (2020). Spent media analysis with an integrated CE-MS analyzer of Chinese hamster ovary cells grown in an ammonia-stressed parallel microbioreactor platform. *BioProcessing Journal*, *19*.

Fan, Y., Jimenez Del Val, I., Müller, C., Wagtberg Sen, J., Rasmussen, S. K., Kontoravdi, C., Weilguny, D., & Andersen, M. R. (2015). Amino acid and glucose metabolism in fed-batch CHO cell culture affects antibody production and glycosylation. *Biotechnology and Bioengineering*, *112*(3), 521-535. https://doi.org/10.1002/bit.25450

Fischer, S., Handrick, R., & Otte, K. (2015). The art of CHO cell engineering: A comprehensive retrospect and future perspectives. *Biotechnology Advances*, *33*(8), 1878-1896.

Fischer, S., Mathias, S., Schaz, S., Emmerling, V. V., Buck, T., Kleemann, M., Hackl, M., Grillari, J., Aschrafi, A., & Handrick, R. (2015). Enhanced protein production by microRNA-30 family in CHO cells is mediated by the modulation of the ubiquitin pathway. *Journal of Biotechnology*, *212*, 32-43.

Fishman, G. I., Spray, D. C., & Leinwand, L. A. (1990). Molecular characterization and functional expression of the human cardiac gap junction channel. *The Journal of Cell Biology*, *111*(2), 589-598. https://doi.org/10.1083/jcb.111.2.589

Fonseca, S. G., Fukuma, M., Lipson, K. L., Nguyen, L. X., Allen, J. R., Oka, Y., & Urano, F. (2005). WFS1 Is a Novel Component of the Unfolded Protein Response and Maintains Homeostasis of the Endoplasmic Reticulum in Pancreatic β-Cells *. *Journal of Biological Chemistry*, *280*(47), 39609-39615. https://doi.org/10.1074/jbc.M507426200

Fonseca, S. G., Ishigaki, S., Oslowski, C. M., Lu, S., Lipson, K. L., Ghosh, R., Hayashi, E., Ishihara, H., Oka, Y., Permutt, M. A., & Urano, F. (2010). Wolfram syndrome 1 gene negatively regulates ER stress signaling in rodent and human cells. *The Journal of Clinical Investigation*, *120*(3), 744-755. https://doi.org/10.1172/JCI39678

Forsythe, H. L., Jarvis, J. L., Turner, J. W., Elmore, L. W., & Holt, S. E. (2001). Stable association of hsp90 and p23, but Not hsp70, with active human telomerase. *The Journal of Biological Chemistry*, *276*(19), 15571-15574. https://doi.org/10.1074/jbc.C100055200

Furukawa, K., & Ohsuye, K. (1999). Enhancement of productivity of recombinant α-amidating enzyme by low temperature culture. *Cytotechnology*, *31*(1), 85-94.

Gardner, B. M., & Walter, P. (2011). Unfolded Proteins are Ire1-Activating Ligands that Directly Induce the Unfolded Protein Response. *Science (New York, N.Y.)*, *333*(6051), 1891-1894. https://doi.org/10.1126/science.1209126

Gawlitzek, M., Ryll, T., Lofgren, J., & Sliwkowski, M. B. (2000). Ammonium alters N-glycan structures of recombinant TNFR-IgG: Degradative versus biosynthetic mechanisms. *Biotechnology and Bioengineering*, *68*(6), 637-646.

Gerlach, J. Q., Sharma, S., Leister, K. J., & Joshi, L. (2012). A tight-knit group: Protein glycosylation, endoplasmic reticulum stress and the unfolded protein response. *Endoplasmic Reticulum Stress in Health and Disease*, 23-39.

Goey, C. H., Tsang, J., Bell, D., & Kontoravdi, C. (2017). Cascading effect in bioprocessing - the impact of mild hypothermia on CHO cell behavior and host cell protein composition. *2781*. https://doi.org/10.1002/bit.26437

Goll, D. E., Thompson, V. F., Li, H., Wei, W., & Cong, J. (2003). The Calpain System. *Physiological Reviews*, *83*(3), 731-801. https://doi.org/10.1152/physrev.00029.2002

Gozzelino, R., Jeney, V., & Soares, M. P. (2010). Mechanisms of cell protection by heme oxygenase-1. *Annual Review of Pharmacology and Toxicology*, *50*, 323-354. https://doi.org/10.1146/annurev.pharmtox.010909.105600

Grove, D. E., Fan, C.-Y., Ren, H. Y., & Cyr, D. M. (2011). The endoplasmic reticulum-associated Hsp40 DNAJB12 and Hsc70 cooperate to facilitate RMA1 E3-dependent degradation of nascent CFTRDeltaF508. *Molecular Biology of the Cell*, *22*(3), 301-314. https://doi.org/10.1091/mbc.E10-09-0760

Ha, T. K., Kim, Y.-G., & Lee, G. M. (2015). Understanding of altered N-glycosylation-related gene expression in recombinant Chinese hamster ovary cells subjected to elevated ammonium concentration by digital mRNA counting. *Biotechnology and Bioengineering*, *112*(8), 1583-1593. https://doi.org/10.1002/bit.25568

Hageman, J., & Kampinga, H. H. (2009). Computational analysis of the human HSPH/HSPA/DNAJ family and cloning of a human HSPH/HSPA/DNAJ expression library. *Cell Stress & Chaperones*, *14*(1), 1-21. https://doi.org/10.1007/s12192-008-0060-2

Haines, B. E., Wiest, O., & Stauffacher, C. V. (2013). The increasingly complex mechanism of HMG-CoA reductase. *Accounts of Chemical Research*, *46*(11), 2416-2426.

Halic, M., Gartmann, M., Schlenker, O., Mielke, T., Pool, M. R., Sinning, I., & Beckmann, R. (2006). Signal recognition particle receptor exposes the ribosomal translocon binding site. *Science (New York, N.Y.)*, *312*(5774), 745-747. https://doi.org/10.1126/science.1124864

Hall, J. C., Heel, K., & McCauley, R. (1996). Glutamine. *BJS (British Journal of Surgery)*, *83*(3), 305-312. https://doi.org/10.1002/bjs.1800830306

Han, Y. K., Kim, Y.-G., Kim, J. Y., & Lee, G. M. (2010). Hyperosmotic stress induces autophagy and apoptosis in recombinant Chinese hamster ovary cell culture. *Biotechnology and Bioengineering*, *105*(6), 1187-1192. https://doi.org/10.1002/bit.22643

Hansen, H. A., & Emborg, C. (1994). Influence of ammonium on growth, metabolism, and productivity of a continuous suspension Chinese hamster ovary cell culture. *Biotechnology Progress*, *10*(1), 121-124.

Hayter, P. M., Curling, E., Baines, A. J., Jenkins, N., Salmon, I., Strange, P. G., & Bull, A. T. (1991). Chinese hamster ovary cell growth and interferon production kinetics in stirred batch culture. *Applied Microbiology and Biotechnology*, *34*, 559-564.

He, Q., Wang, T., Ni, H., Liu, Q., An, K., Tao, J., Chen, Y., Xu, L., Zhu, C., & Yao, M. (2019). Endoplasmic reticulum stress promoting caspase signaling pathway-dependent apoptosis contributes to bone cancer pain in the spinal dorsal horn. *Molecular Pain*, *15*, 1744806919876150. https://doi.org/10.1177/1744806919876150

Helenius, A., Trombetta, E. S., Hebert, D. N., & Simons, J. F. (1997). Calnexin, calreticulin and the folding of glycoproteins. *Trends in Cell Biology*, *7*(5), 193-200.

Hernandez, I., Dhiman, H., Klanert, G., Jadhav, V., Auer, N., Hanscho, M., Baumann, M., Esteve-Codina, A., Dabad, M., & Gómez, J. (2019). Epigenetic regulation of gene expression in Chinese Hamster Ovary cells in response to the changing environment of a batch culture. *Biotechnology and Bioengineering*, *116*(3), 677-692.

Hong, Y., Maeda, Y., Watanabe, R., Inoue, N., Ohishi, K., & Kinoshita, T. (2000). Requirement of PIG-F and PIG-O for transferring phosphoethanolamine to the third mannose in glycosylphosphatidylinositol. *The Journal of Biological Chemistry*, *275*(27), 20911-20919. https://doi.org/10.1074/jbc.M001913200

Ito, Y., Takeda, Y., Seko, A., Izumi, M., & Kajihara, Y. (2015). Functional analysis of endoplasmic reticulum glucosyltransferase (UGGT): Synthetic chemistry's initiative in glycobiology. *Seminars in Cell & Developmental Biology*, *41*, 90-98. https://doi.org/10.1016/j.semcdb.2014.11.011

Jäkel, S., Albig, W., Kutay, U., Bischoff, F. R., Schwamborn, K., Doenecke, D., & Görlich, D. (1999). The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1. *The EMBO Journal*, *18*(9), 2411-2423. https://doi.org/10.1093/emboj/18.9.2411

Jiang, Q., Sun, Y., Guo, Z., Fu, M., Wang, Q., Zhu, H., Lei, P., & Shen, G. (2016). Overexpression of GRP78 enhances survival of CHO cells in response to serum deprivation and oxidative stress. *Engineering in Life Sciences*, *17*(2), 107-116. https://doi.org/10.1002/elsc.201500152

Johari, Y. B., Estes, S. D., Alves, C. S., Sinacore, M. S., & James, D. C. (2015). Integrated cell and process engineering for improved transient production of a “difficult-to-express “fusion protein by CHO cells. *Biotechnology and Bioengineering*, *112*(12), 2527-2542.

Jung, E. S., Hong, H., Kim, C., & Mook-Jung, I. (2015). Acute ER stress regulates amyloid precursor protein processing through ubiquitin-dependent degradation. *Scientific Reports*, *5*(1), 8805.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., & Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research*, *51*(D1), D587-D592.

Kaufmann, H., Mazur, X., Fussenegger, M., & Bailey, J. E. (1999). Influence of low temperature on productivity, proteome and protein phosphorylation of CHO cells. *Biotechnology and Bioengineering*, *63*(5), 573-582.

Kilgas, S., & Ramadan, K. (2023). Inhibitors of the ATPase p97/VCP: From basic research to clinical applications. *Cell Chemical Biology*, *30*(1), 3-21. https://doi.org/10.1016/j.chembiol.2022.12.007

Kim, J. Y., Kim, Y.-G., & Lee, G. M. (2012). CHO cells in biotechnology for production of recombinant proteins: Current state and further potential. *Applied Microbiology and Biotechnology*, *93*, 917-930.

Kim, S. H., Shin, S., Baek, M., Xiong, K., Karottki, K. J. la C., Hefzi, H., Grav, L. M., Pedersen, L. E., Kildegaard, H. F., Lewis, N. E., Lee, J. S., & Lee, G. M. (2023). Identification of hyperosmotic stress-responsive genes in Chinese hamster ovary cells via genome-wide virus-free CRISPR/Cas9 screening. *Metabolic Engineering*, *80*, 66-77. https://doi.org/10.1016/j.ymben.2023.09.006

Kinoshita, T. (2016). Glycosylphosphatidylinositol (GPI) Anchors: Biochemistry and Cell Biology: Introduction to a Thematic Review Series. *Journal of Lipid Research*, *57*(1), 4-5. https://doi.org/10.1194/jlr.E065417

Kleijnen, M. F., Shih, A. H., Zhou, P., Kumar, S., Soccio, R. E., Kedersha, N. L., Gill, G., & Howley, P. M. (2000). The hPLIC proteins may provide a link between the ubiquitination machinery and the proteasome. *Molecular Cell*, *6*(2), 409-419. https://doi.org/10.1016/s1097-2765(00)00040-x

Knight, T. J., Turner, S., Jaques, C. M., & Smales, C. M. (2021). Selection of CHO host and recombinant cell pools by inhibition of the proteasome results in enhanced product yields and cell specific productivity. *Journal of Biotechnology*, *337*, 35-45.

Kyeong, M., & Lee, J. S. (2022). Endogenous BiP reporter system for simultaneous identification of ER stress and antibody production in Chinese hamster ovary cells. *Metabolic Engineering*, *72*, 35-45. https://doi.org/10.1016/j.ymben.2022.02.002

Lang, S., Benedix, J., Fedeles, S. V., Schorr, S., Schirra, C., Schäuble, N., Jalal, C., Greiner, M., Hassdenteufel, S., Tatzelt, J., Kreutzer, B., Edelmann, L., Krause, E., Rettig, J., Somlo, S.,

Zimmermann, R., & Dudek, J. (2012). Different effects of Sec61α, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. *Journal of Cell Science*, *125*(Pt 8), 1958-1969. https://doi.org/10.1242/jcs.096727

Lao, M.-S., & Toth, D. (1997). Effects of Ammonium and Lactate on Growth and Metabolism of a Recombinant Chinese Hamster Ovary Cell Culture. *Biotechnology Progress*, *13*(5), 688-691. https://doi.org/10.1021/bp9602360

Latorre, Y., Torres, M., Vergara, M., Berrios, J., Sampayo, M. M., Gödecke, N., Wirth, D., Hauser, H., Dickson, A. J., & Altamirano, C. (2023). Engineering of Chinese hamster ovary cells for co-overexpressing MYC and XBP1s increased cell proliferation and recombinant EPO production. *Scientific Reports*, *13*(1), 1482.

Lemarié, M., Chatonnet, F., Caron, G., & Fest, T. (2021). Early Emergence of Adaptive Mechanisms Sustaining Ig Production: Application to Antibody Therapy. *Frontiers in Immunology*, *12*. https://doi.org/10.3389/fimmu.2021.671998

Li, D., Chen, S., Liu, C., Wei, B., & Li, X. (2022). Liver transcriptome analysis reveals biological pathways and transcription factors in response to high ammonia exposure. *Inhalation Toxicology*, *34*(7-8), 219-229. https://doi.org/10.1080/08958378.2022.2083275

Li, W., Fan, Z., Lin, Y., & Wang, T.-Y. (2021). Serum-Free Medium for Recombinant Protein Expression in Chinese Hamster Ovary Cells. *Frontiers in Bioengineering and Biotechnology*, *9*. https://doi.org/10.3389/fbioe.2021.646363

Liaci, A. M., Steigenberger, B., Telles de Souza, P. C., Tamara, S., Gröllers-Mulderij, M., Ogrissek, P., Marrink, S. J., Scheltema, R. A., & Förster, F. (2021). Structure of the human signal peptidase complex reveals the determinants for signal peptide cleavage. *Molecular Cell*, *81*(19), 3934-3948.e11. https://doi.org/10.1016/j.molcel.2021.07.031

Lim, P. J., Danner, R., Liang, J., Doong, H., Harman, C., Srinivasan, D., Rothenberg, C., Wang, H., Ye, Y., Fang, S., & Monteiro, M. J. (2009). Ubiquilin and p97/VCP bind erasin, forming a complex involved in ERAD. *The Journal of Cell Biology*, *187*(2), 201-217. https://doi.org/10.1083/jcb.200903024

Liu, K.-T., Yeh, I.-J., Chou, S.-K., Yen, M.-C., & Kuo, P.-L. (2018). Regulatory mechanism of fatty acid-CoA metabolic enzymes under endoplasmic reticulum stress in lung cancer. *Oncology Reports*, *40*(5), 2674-2682.

Lopata, A., Kniss, A., Löhr, F., Rogov, V. V., & Dötsch, V. (2020). Ubiquitination in the ERAD Process. *International Journal of Molecular Sciences*, *21*(15), 5369. https://doi.org/10.3390/ijms21155369

Malm, M., Kuo, C.-C., Barzadd, M. M., Mebrahtu, A., Wistbacka, N., Razavi, R., Volk, A.-L., Lundqvist, M., Kotol, D., & Tegel, H. (2022). Harnessing secretory pathway differences between HEK293 and CHO to rescue production of difficult to express proteins. *Metabolic Engineering*, *72*, 171-187.

Man, X., Megraw, T. L., & Lim, Y. P. (2015). Cep68 can be regulated by Nek2 and SCF complex. *European Journal of Cell Biology*, *94*(3-4), 162-172. https://doi.org/10.1016/j.ejcb.2015.01.004

McHugh, K. P., Xu, J., Aron, K. L., Borys, M. C., & Li, Z. J. (2020). Effective temperature shift strategy development and scale confirmation for simultaneous optimization of protein productivity and quality in Chinese hamster ovary cells. *Biotechnology Progress*, *36*(3), e2959. https://doi.org/10.1002/btpr.2959

Meacock, S. L., Lecomte, F. J. L., Crawshaw, S. G., & High, S. (2002a). Different transmembrane domains associate with distinct endoplasmic reticulum components during membrane integration of a polytopic protein. *Molecular Biology of the Cell*, *13*(12), 4114-4129. https://doi.org/10.1091/mbc.e02-04-0198

Meacock, S. L., Lecomte, F. J. L., Crawshaw, S. G., & High, S. (2002b). Different transmembrane domains associate with distinct endoplasmic reticulum components during membrane integration of a polytopic protein. *Molecular Biology of the Cell*, *13*(12), 4114-4129. https://doi.org/10.1091/mbc.e02-04-0198

Melville, D. B., Studer, S., & Schekman, R. (2020). Small sequence variations between two mammalian paralogs of the small GTPase SAR1 underlie functional differences in coat protein complex II assembly. *The Journal of Biological Chemistry*, *295*(25), 8401-8412. https://doi.org/10.1074/jbc.RA120.012964

Mitchell, S., Ellingson, C., Coyne, T., Hall, L., Neill, M., Christian, N., Higham, C., Dobrowolski, S. F., Tuchman, M., Summar, M., & Urea Cycle Disorder Consortium. (2009). Genetic variation in the urea cycle: A model resource for investigating key candidate genes for common diseases. *Human Mutation*, *30*(1), 56-60. https://doi.org/10.1002/humu.20813

Nagata, T., Ilieva, H., Murakami, T., Shiote, M., Narai, H., Ohta, Y., Hayashi, T., Shoji, M., & Abe, K. (2007). Increased ER stress during motor neuron degeneration in a transgenic mouse model of amyotrophic lateral sclerosis. *Neurological Research*, *29*(8), 767-771. https://doi.org/10.1179/016164107X229803

Nery, F. C., Armata, I. A., Farley, J. E., Cho, J. A., Yaqub, U., Chen, P., da Hora, C. C., Wang, Q., Tagaya, M., Klein, C., Tannous, B., Caldwell, K. A., Caldwell, G. A., Lencer, W. I., Ye, Y., & Breakefield, X. O. (2011). TorsinA participates in endoplasmic reticulum-associated degradation. *Nature Communications*, *2*(1), 393. https://doi.org/10.1038/ncomms1383

Nguyen, L. N., Novak, N., Baumann, M., Koehn, J., & Borth, N. (2020). Bioinformatic Identification of Chinese Hamster Ovary (CHO) Cold-Shock Genes and Biological Evidence of their Cold-Inducible Promoters. *Biotechnology Journal*, *15*(3), 1900359. https://doi.org/10.1002/biot.201900359

NISTCHO. (2016). *NIST*. https://www.nist.gov/programs-projects/nistcho

Ohishi, K., Inoue, N., & Kinoshita, T. (2001). PIG-S and PIG-T, essential for GPI anchor attachment to proteins, form a complex with GAA1 and GPI8. *The EMBO Journal*, *20*(15), 4088-4098. https://doi.org/10.1093/emboj/20.15.4088

Ohkuni, A., Ohno, Y., & Kihara, A. (2013). Identification of acyl-CoA synthetases involved in the mammalian sphingosine 1-phosphate metabolic pathway. *Biochemical and Biophysical Research Communications*, *442*(3-4), 195-201. https://doi.org/10.1016/j.bbrc.2013.11.036

Ohya, T., Hayashi, T., Kiyama, E., Nishii, H., Miki, H., Kobayashi, K., Honda, K., Omasa, T., & Ohtake, H. (2008). Improved production of recombinant human antithrombin III in Chinese hamster ovary cells by ATF4 overexpression. *Biotechnology and Bioengineering*, *100*(2), 317-324. https://doi.org/10.1002/bit.21758

Okumura, M., Kadokura, H., & Inaba, K. (2015). Structures and functions of protein disulfide isomerase family members involved in proteostasis in the endoplasmic reticulum. *Free Radical Biology and Medicine*, *83*, 314-322.

Oliveros, J. C. (2007). VENNY. An interactive tool for comparing lists with Venn Diagrams. *Http://Bioinfogp. Cnb. Csic. Es/Tools/Venny/Index. Html*.

Oslowski, C. M., & Urano, F. (2011). Measuring ER stress and the unfolded protein response using mammalian tissue culture system. In *Methods in enzymology* (Vol. 490, pp. 71-92). Elsevier.

Ozawa, K., Kuwabara, K., Tamatani, M., Takatsuji, K., Tsukamoto, Y., Kaneda, S., Yanagi, H., Stern, D. M., Eguchi, Y., Tsujimoto, Y., Ogawa, S., & Tohyama, M. (1999). 150-kDa oxygen-regulated protein (ORP150) suppresses hypoxia-induced apoptotic cell death. *The Journal of Biological Chemistry*, *274*(10), 6397-6404. https://doi.org/10.1074/jbc.274.10.6397

Paccaud, J. P., Reith, W., Carpentier, J. L., Ravazzola, M., Amherdt, M., Schekman, R., & Orci, L. (1996a). Cloning and functional characterization of mammalian homologues of the COPII component Sec23. *Molecular Biology of the Cell*, *7*(10), 1535-1546. https://doi.org/10.1091/mbc.7.10.1535

Paccaud, J. P., Reith, W., Carpentier, J. L., Ravazzola, M., Amherdt, M., Schekman, R., & Orci, L. (1996b). Cloning and functional characterization of mammalian homologues of the COPII component Sec23. *Molecular Biology of the Cell*, *7*(10), 1535-1546.

Papadopoulos, C., Kirchner, P., Bug, M., Grum, D., Koerver, L., Schulze, N., Poehler, R., Dressler, A., Fengler, S., Arhzaouy, K., Lux, V., Ehrmann, M., Weihl, C. C., & Meyer, H. (2017). VCP/p97 cooperates with YOD1, UBXD1 and PLAA to drive clearance of ruptured lysosomes by autophagy. *The EMBO Journal*, *36*(2), 135-150. https://doi.org/10.15252/embj.201695148

Park, S.-Y., Egan, S., Cura, A. J., Aron, K. L., Xu, X., Zheng, M., Borys, M., Ghose, S., Li, Z., & Lee, K. (2021). Untargeted proteomics reveals upregulation of stress response pathways

during CHO-based monoclonal antibody manufacturing process leading to disulfide bond reduction. *mAbs*, *13*(1), 1963094. https://doi.org/10.1080/19420862.2021.1963094

Pereira, S., Kildegaard, H. F., & Andersen, M. R. (2018). Impact of CHO Metabolism on Cell Growth and Protein Production: An Overview of Toxic and Inhibiting Metabolites and Nutrients. *Biotechnology Journal*, *13*(3), 1700499. https://doi.org/10.1002/biot.201700499

Perez-Riverol, Y., Bandla, C., Kundu, D. J., Kamatchinathan, S., Bai, J., Hewapathirana, S., John, N. S., Prakash, A., Walzer, M., & Wang, S. (2025). The PRIDE database at 20 years: 2025 update. *Nucleic Acids Research*, *53*(D1), D543-D553.

Pickart, C. M., & Eddins, M. J. (2004). Ubiquitin: Structures, functions, mechanisms. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1695*(1), 55-72. https://doi.org/10.1016/j.bbamcr.2004.09.019

Pieper, L. A., Strotbek, M., Wenger, T., Olayioye, M. A., & Hausser, A. (2017). ATF6β-based fine-tuning of the unfolded protein response enhances therapeutic antibody productivity of Chinese hamster ovary cells. *Biotechnology and Bioengineering*, *114*(6), 1310-1318.

Prashad, K., & Mehra, S. (2015). Dynamics of unfolded protein response in recombinant CHO cells. *Cytotechnology*, *67*, 237-254.

Pybus, L. P., Dean, G., West, N. R., Smith, A., Daramola, O., Field, R., Wilkinson, S. J., & James, D. C. (2014a). Model-d irected engineering of "difficult-to-e xpress" monoclonal antibody production by Chinese hamster ovary cells. *Biotechnology and Bioengineering*, *111*(2), 372-385.

Pybus, L. P., Dean, G., West, N. R., Smith, A., Daramola, O., Field, R., Wilkinson, S. J., & James, D. C. (2014b). Model-directed engineering of "difficult-to-express" monoclonal antibody production by Chinese hamster ovary cells. *Biotechnology and Bioengineering*, *111*(2), 372-385. https://doi.org/10.1002/bit.25116

Qin, S.-Y., Hu, D., Matsumoto, K., Takeda, K., Matsumoto, N., Yamaguchi, Y., & Yamamoto, K. (2012). Malectin forms a complex with ribophorin I for enhanced association with misfolded glycoproteins. *The Journal of Biological Chemistry*, *287*(45), 38080-38089. https://doi.org/10.1074/jbc.M112.394288

Qiu, L., Pashkova, N., Walker, J. R., Winistorfer, S., Allali-Hassani, A., Akutsu, M., Piper, R., & Dhe-Paganon, S. (2010). Structure and Function of the PLAA/Ufd3-p97/Cdc48 Complex *. *Journal of Biological Chemistry*, *285*(1), 365-372. https://doi.org/10.1074/jbc.M109.044685

Qu, J., Zou, T., & Lin, Z. (2021). The Roles of the Ubiquitin-Proteasome System in the Endoplasmic Reticulum Stress Pathway. *International Journal of Molecular Sciences*, *22*(4), 1526. https://doi.org/10.3390/ijms22041526

Radons, J. (2016). The human HSP70 family of chaperones: Where do we stand? *Cell Stress & Chaperones*, *21*(3), 379-404. https://doi.org/10.1007/s12192-016-0676-6

Ramírez, A. S., Kowal, J., & Locher, K. P. (2019). Cryo-electron microscopy structures of human oligosaccharyltransferase complexes OST-A and OST-B. *Science (New York, N.Y.)*, *366*(6471), 1372-1375. https://doi.org/10.1126/science.aaz3505

Read, A., & Schröder, M. (2021). The unfolded protein response: An overview. *Biology*, *10*(5), 384.

Reinhart, D., Sommeregger, W., Debreczeny, M., Gludovacz, E., & Kunert, R. (2014). In search of expression bottlenecks in recombinant CHO cell lines—A case study. *Applied Microbiology and Biotechnology*, *98*(13), 5959-5965. https://doi.org/10.1007/s00253-014-5584-z

Retzlaff, M., Stahl, M., Eberl, H. C., Lagleder, S., Beck, J., Kessler, H., & Buchner, J. (2009). Hsp90 is regulated by a switch point in the C-terminal domain. *EMBO Reports*, *10*(10), 1147-1153. https://doi.org/10.1038/embor.2009.153

Rezvani, K. (2016). UBXD Proteins: A Family of Proteins with Diverse Functions in Cancer. *International Journal of Molecular Sciences*, *17*(10), Article 10. https://doi.org/10.3390/ijms17101724

Richelle, A., & Lewis, N. E. (2017). Improvements in protein production in mammalian cells from targeted metabolic engineering. *Current Opinion in Systems Biology*, *6*, 1-6.

Roos-Mattjus, P., & Sistonen, L. (2004). The ubiquitin-proteasome pathway. *Annals of Medicine*, *36*(4), 285-295. https://doi.org/10.1080/07853890310016324

Roy, G., Zhang, S., Li, L., Higham, E., Wu, H., Marelli, M., & Bowen, M. A. (2017). Development of a fluorescent reporter system for monitoring ER stress in Chinese hamster ovary cells and its application for therapeutic protein production. *PLoS ONE*, *12*(8), e0183694. https://doi.org/10.1371/journal.pone.0183694

Rutkevich, L. A., & Williams, D. B. (2011). Participation of lectin chaperones and thiol oxidoreductases in protein folding within the endoplasmic reticulum. *Current Opinion in Cell Biology*, *23*(2), 157-166.

Rutkowski, D. T., & Kaufman, R. J. (2004). A trip to the ER: coping with stress. *Trends in Cell Biology*, *14*(1), 20-28. https://doi.org/10.1016/j.tcb.2003.11.001

Ryan, D., Sideri, C.-K., Henry, M., Efeoglu, E., & Meleady, P. (2023). Label-Free Quantitative Proteomics Analysis of Producer and Non-producer Chinese Hamsters Ovary (CHO) Cells under ER Stress Conditions. *Current Research in Biotechnology*, 100141.

Schröder, M., & Friedl, P. (1997). Overexpression of recombinant human antithrombin III in Chinese hamster ovary cells results in malformation and decreased secretion of recombinant protein. *Biotechnology and Bioengineering*, *53*(6), 547-559.

Selvaprakash, K., Sideri, C.-K., Henry, M., Efeoglu, E., Ryan, D., & Meleady, P. (2024). Characterization of the Ubiquitin-Modified Proteome of Recombinant Chinese Hamster Ovary Cells in Response to Endoplasmic Reticulum Stress. *Biotechnology Journal*, *19*(11), e202400413. https://doi.org/10.1002/biot.202400413

Shahabi, F., Abdoli, S., Bazi, Z., Shamsabadi, F., Yamchi, A., & Shahbazi, M. (2023). Enhancing productivity of Chinese hamster ovary (CHO) cells: Synergistic strategies combining low-temperature culture and mTORC1 signaling engineering. *Frontiers in Bioengineering and Biotechnology*, *11*. https://doi.org/10.3389/fbioe.2023.1268048

Smales, C. m., Dinnis, D. m., Stansfield, S. h., Alete, D., Sage, E. a., Birch, J. r., Racher, A. j., Marshall, C. t., & James, D. c. (2004). Comparative proteomic analysis of GS-NS0 murine myeloma cell lines with varying recombinant monoclonal antibody production rate. *Biotechnology and Bioengineering*, *88*(4), 474-488. https://doi.org/10.1002/bit.20272

Smales, C. M., Dinnis, D. M., Stansfield, S. H., Alete, D., Sage, E. A., Birch, J. R., Racher, A. J., Marshall, C. T., & James, D. C. (2004). Comparative proteomic analysis of GS-NS0 murine myeloma cell lines with varying recombinant monoclonal antibody production rate. *Biotechnology and Bioengineering*, *88*(4), 474-488.

Sonnenberg, A., & Liem, R. K. H. (2007). Plakins in development and disease. *Experimental Cell Research*, *313*(10), 2189-2203. https://doi.org/10.1016/j.yexcr.2007.03.039

Sou, S. N., Sellick, C., Lee, K., Mason, A., Kyriakopoulos, S., Polizzi, K. M., & Kontoravdi, C. (2015). How does mild hypothermia affect monoclonal antibody glycosylation? *1176*. https://doi.org/10.1002/bit.25524

Soupene, E., & Kuypers, F. A. (2008). Mammalian Long-Chain Acyl-CoA Synthetases. *Experimental Biology and Medicine*, *233*(5), 507-521. https://doi.org/10.3181/0710-MR-287

Stec, D. E., & Hinds, T. D. (2020). Natural Product Heme Oxygenase Inducers as Treatment for Nonalcoholic Fatty Liver Disease. *International Journal of Molecular Sciences*, *21*(24), 9493. https://doi.org/10.3390/ijms21249493

Stetler, R. A., Gan, Y., Zhang, W., Liou, A. K., Gao, Y., Cao, G., & Chen, J. (2010). Heat Shock Proteins: Cellular and molecular mechanisms in the CNS. *Progress in Neurobiology*, *92*(2), 184-211. https://doi.org/10.1016/j.pneurobio.2010.05.002

Stolz, A., & Wolf, D. H. (2010). Endoplasmic reticulum associated protein degradation: A chaperone assisted journey to hell. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1803*(6), 694-705. https://doi.org/10.1016/j.bbamcr.2010.02.005

Strating, J. R. P. M., & Martens, G. J. M. (2009). The p24 family and selective transport processes at the ER-Golgi interface. *Biology of the Cell*, *101*(9), 495-509. https://doi.org/10.1042/BC20080233

Sulaj, E., Schwaigerlehner, L., Sandell, F. L., Dohm, J. C., Marzban, G., & Kunert, R. (2024). Quantitative proteomics reveals cellular responses to individual mAb expression and tunicamycin in CHO cells. *Applied Microbiology and Biotechnology*, *108*(1), 1-19. https://doi.org/10.1007/s00253-024-13223-1

Sunley, K., Tharmalingam, T., & Butler, M. (2008). CHO cells adapted to hypothermic growth produce high yields of recombinant β-interferon. *Biotechnology Progress*, *24*(4), 898-906.

Synoground, B. F., McGraw, C. E., Elliott, K. S., Leuze, C., Roth, J. R., Harcum, S. W., & Sandoval, N. R. (2021). Transient ammonia stress on Chinese hamster ovary (CHO) cells yield alterations to alanine metabolism and IgG glycosylation profiles. *Biotechnology Journal*, *16*(7), 2100098. https://doi.org/10.1002/biot.202100098

Szklarczyk, D., Gable, A. L., Nastou, K. C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N. T., Legeay, M., Fang, T., Bork, P., Jensen, L. J., & von Mering, C. (2021). The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Research*, *49*(D1), D605-D612. https://doi.org/10.1093/nar/gkaa1074

Tang, B. L., Peter, F., Krijnse-Locker, J., Low, S. H., Griffiths, G., & Hong, W. (1997). The mammalian homolog of yeast Sec13p is enriched in the intermediate compartment and is essential for protein transport from the endoplasmic reticulum to the Golgi apparatus. *Molecular and Cellular Biology*, *17*(1), 256-266. https://doi.org/10.1128/MCB.17.1.256

Tang, B. L., Zhang, T., Low, D. Y., Wong, E. T., Horstmann, H., & Hong, W. (2000). Mammalian homologues of yeast sec31p. An ubiquitously expressed form is localized to endoplasmic reticulum (ER) exit sites and is essential for ER-Golgi transport. *The Journal of Biological Chemistry*, *275*(18), 13597-13604. https://doi.org/10.1074/jbc.275.18.13597

Tang, D., Sandoval, W., Lam, C., Haley, B., Liu, P., Xue, D., Roy, D., Patapoff, T., Louie, S., & Snedecor, B. (2020). UBR E3 ligases and the PDIA3 protease control degradation of unfolded antibody heavy chain by ERAD. *Journal of Cell Biology*, *219*(7), e201908087.

Tani, K., Oyama, Y., Hatsuzawa, K., & Tagaya, M. (1999). Hypothetical protein KIAA0079 is a mammalian homologue of yeast Sec24p. *FEBS Letters*, *447*(2-3), 247-250. https://doi.org/10.1016/s0014-5793(99)00303-8

Torres, M., Akhtar, S., McKenzie, E. A., & Dickson, A. J. (2021). Temperature Down-Shift Modifies Expression of UPR-/ERAD-Related Genes and Enhances Production of a Chimeric Fusion Protein in CHO Cells. *Biotechnology Journal*, *16*(2), 2000081. https://doi.org/10.1002/biot.202000081

Tsunoda, Y., Yamano-Adachi, N., Koga, Y., & Omasa, T. (2024). Sar1A overexpression in Chinese hamster ovary cells and its effects on antibody productivity and secretion. *Journal of Bioscience and Bioengineering*, *138*(2), 171-180. https://doi.org/10.1016/j.jbiosc.2024.05.003

Uchiyama, K., Jokitalo, E., Lindman, M., Jackman, M., Kano, F., Murata, M., Zhang, X., & Kondo, H. (2003). The localization and phosphorylation of p47 are important for Golgi disassembly-assembly during the cell cycle. *The Journal of Cell Biology*, *161*(6), 1067-1079. https://doi.org/10.1083/jcb.200303048

van der Wal, L., Bezstarosti, K., Sap, K. A., Dekkers, D. H., Rijkers, E., Mientjes, E., Elgersma, Y., & Demmers, J. A. (2018a). Improvement of ubiquitylation site detection by Orbitrap mass spectrometry. *Journal of Proteomics*, *172*, 49-56.

van der Wal, L., Bezstarosti, K., Sap, K. A., Dekkers, D. H. W., Rijkers, E., Mientjes, E., Elgersma, Y., & Demmers, J. A. A. (2018b). Improvement of ubiquitylation site detection by Orbitrap mass spectrometry. *Journal of Proteomics*, *172*, 49-56. https://doi.org/10.1016/j.jprot.2017.10.014

Vedrenne, C., Klopfenstein, D. R., & Hauri, H.-P. (2005). Phosphorylation controls CLIMP-63-mediated anchoring of the endoplasmic reticulum to microtubules. *Molecular Biology of the Cell*, *16*(4), 1928-1937. https://doi.org/10.1091/mbc.e04-07-0554

Vorum, H., Hager, H., Christensen, B. M., Nielsen, S., & Honoré, B. (1999). Human calumenin localizes to the secretory pathway and is secreted to the medium. *Experimental Cell Research*, *248*(2), 473-481. https://doi.org/10.1006/excr.1999.4431

Wang, K., Zhang, T., Chen, J., Liu, C., Tang, J., & Xie, Q. (2018). The effect of culture temperature on the aggregation of recombinant TNFR-Fc is regulated by the PERK-eIF2a pathway in CHO cells. *Protein and Peptide Letters*, *25*(6), 570-579.

Wang, T., & Hebert, D. N. (2003). EDEM an ER quality control receptor. *Nature Structural & Molecular Biology*, *10*(5), 319-321. https://doi.org/10.1038/nsb0503-319

Weerapana, E., & Imperiali, B. (2006). Asparagine-linked protein glycosylation: From eukaryotic to prokaryotic systems. *Glycobiology*, *16*(6), 91R-101R.

Weis, K., Griffiths, G., & Lamond, A. I. (1994). The endoplasmic reticulum calcium-binding protein of 55 kDa is a novel EF-hand protein retained in the endoplasmic reticulum by a carboxyl-terminal His-Asp-Glu-Leu motif. *The Journal of Biological Chemistry*, *269*(29), 19142-19150. https://doi.org/10.1016/s0021-9258(17)32286-x

Weng, Z., Jin, J., Shao, C., & Li, H. (2020). Reduction of charge variants by CHO cell culture process optimization. *Cytotechnology*, *72*(2), 259-269. https://doi.org/10.1007/s10616-020-00375-x

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer International Publishing.

Wilkinson, K. D. (1999). Ubiquitin-Dependent Signaling: The Role of Ubiquitination in the Response of Cells to Their Environment.12. *The Journal of Nutrition*, *129*(11), 1933-1936. https://doi.org/10.1093/jn/129.11.1933

Wu, B., Yu, L., Hu, P., Lu, Y., Li, J., Wei, Y., & He, R. (2018). GRP78 protects CHO cells from ribosylation. *Biochimica Et Biophysica Acta. Molecular Cell Research*, *1865*(4), 629-637. https://doi.org/10.1016/j.bbamcr.2018.02.001

Xu, D., Wan, A., Zhang, J., Peng, L., Chen, Y., He, Y., Yang, J., & Jin, J. (2017). Inhibition of the ubiquitin ligase activity improves the production of biologically active fusion protein HSA-HGF in Chinese hamster ovary cells. *Bioengineered*, *8*(3), 256-264.

Xu, S., Gavin, J., Jiang, R., & Chen, H. (2017). Bioreactor productivity and media cost comparison for different intensified cell culture processes. *Biotechnology Progress*, *33*(4), 867-878.

Yamaguchi, T., Dulubova, I., Min, S.-W., Chen, X., Rizo, J., & Südhof, T. C. (2002). Sly1 binds to Golgi and ER syntaxins via a conserved N-terminal peptide motif. *Developmental Cell*, *2*(3), 295-305.

Yamamoto, K. (2021). *Expression and binding assay of vesicular integral membrane protein 36 (VIP36, LMAN2*. https://www.ncbi.nlm.nih.gov/books/NBK593891

Yamamoto, Y., Kimura, T., Momohara, S., Takeuchi, M., Tani, T., Kimata, Y., Kadokura, H., & Kohno, K. (2010). A novel ER J-protein DNAJB12 accelerates ER-associated degradation of membrane proteins including CFTR. *Cell Structure and Function*, *35*(2), 107-116. https://doi.org/10.1247/csf.10023

Yang, M., & Butler, M. (2000a). Effect of ammonia on the glycosylation of human recombinant erythropoietin in culture. *Biotechnology Progress*, *16*(5), 751-759. Scopus. https://doi.org/10.1021/bp000090b

Yang, M., & Butler, M. (2000b). Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation. *Biotechnology and Bioengineering*, *68*(4), 370-380.

Yee, J. C., Gerdtzen, Z. P., & Hu, W.-S. (2009). Comparative transcriptome analysis to unveil genes affecting recombinant protein productivity in mammalian cells. *Biotechnology and Bioengineering*, *102*(1), 246-263.

Yin, H., Han, S., Cui, C., Wang, Y., Li, D., & Zhu, Q. (2021). Plectin regulates Wnt signaling mediated-skeletal muscle development by interacting with Dishevelled-2 and antagonizing autophagy. *Gene*, *783*, 145562. https://doi.org/10.1016/j.gene.2021.145562

Yoon, S. K., Hwang, S. O., & Lee, G. M. (2004). Enhancing effect of low culture temperature on specific antibody productivity of recombinant Chinese hamster ovary cells: Clonal variation. *Biotechnology Progress*, *20*(6), 1683-1688.

Yoon, S. K., Song, J. Y., & Lee, G. M. (2003). Effect of low culture temperature on specific productivity, transcription level, and heterogeneity of erythropoietin in Chinese hamster ovary cells. *Biotechnology and Bioengineering*, *82*(3), 289-298.

Yu, M., Haslam, R. H., & Haslam, D. B. (2000). HEDJ, an Hsp40 co-chaperone localized to the endoplasmic reticulum of human cells. *The Journal of Biological Chemistry*, *275*(32), 24984-24992. https://doi.org/10.1074/jbc.M000739200

Yuan, X., Shaw, A., Zhang, X., Kondo, H., Lally, J., Freemont, P. S., & Matthews, S. (2001). Solution structure and interaction surface of the C-terminal domain from p47: A major p97-cofactor involved in SNARE disassembly11Edited by M. F. Summers. *Journal of Molecular Biology*, *311*(2), 255-263. https://doi.org/10.1006/jmbi.2001.4864

Zhang, J., Li, C., Tang, X., Lu, Q., Sa, R., & Zhang, H. (2015). High Concentrations of Atmospheric Ammonia Induce Alterations in the Hepatic Proteome of Broilers (Gallus gallus): An iTRAQ-Based Quantitative Proteomic Analysis. *PloS One*, *10*(4), e0123596. https://doi.org/10.1371/journal.pone.0123596

Zhang, M., Liu, L., Lin, X., Wang, Y., Li, Y., Guo, Q., Li, S., Sun, Y., Tao, X., Zhang, D., Lv, X., Zheng, L., & Ge, L. (2020). A Translocation Pathway for Vesicle-Mediated Unconventional Protein Secretion. *Cell*, *181*(3), 637-652.e15. https://doi.org/10.1016/j.cell.2020.03.031

Zhang, S., Liu, C., Li, Y., Imam, M. U., Huang, H., Liu, H., Xin, Y., & Zhang, H. (2016). Novel Role of ER Stress and Autophagy in Microcystin-LR Induced Apoptosis in Chinese Hamster Ovary Cells. *Frontiers in Physiology*, *7*. https://doi.org/10.3389/fphys.2016.00527

Zong, W.-X., Li, C., Hatzivassiliou, G., Lindsten, T., Yu, Q.-C., Yuan, J., & Thompson, C. B. (2003). Bax and Bak can localize to the endoplasmic reticulum to initiate apoptosis. *The Journal of Cell Biology*, *162*(1), 59-69. https://doi.org/10.1083/jcb.200302084

Chapter 4: Paper 3-Proteomic Analysis of Endoplasmic Reticulum (ER) Protein Expression under Ammonia-Induced Stress in Chinese Hamster Ovary (CHO) Cell Cultures

Ryan, D., Sideri, C.-K., Henry, M., Efeoglu, E. and Meleady, P. (2026)., Proteomic Analysis of Endoplasmic Reticulum (ER) Protein Expression under Ammonia-Induced Stress in Chinese Hamster Ovary (CHO) Cell Cultures. Journal of Proteomics, Volume 328, 2026, 105655, ISSN 1874-3919, https://doi.org/10.1016/j.jprot.2026.105655.

This study explores the relationship between the cellular waste product, ammonia, and culture viability, productivity and ER function. A by-production of amino acid catabolism, ammonia can cause toxicity but also acts a vital source of nitrogen for amino acid creation. In batch cultures, ammonia levels can reach levels of 20- to 30 mM (Capiaumont et al., 1995; Chitwood et al., 2021b; Yang & Butler, 2000b). It has been found levels over 20 mM can be inhibitory to culture growth and productivity (Chen & Harcum, 2006a; Lao & Toth, 1997). This negative affect on growth has been linked to changes in carbon/nitrogen metabolism, cell cycle and cell longevity pathways (Chitwood et al., 2021a). Compared to other environmental stressors, ammonia stress has been relatively under investigated, in particular using in-depth proteomics. Existing studies link ammonia stress to alterations in glycosylation pathways (Synoground et al., 2021; Yang & Butler, 2000a). Yang & Butler have found that a moderate increase in ammonia level (10 mM) can lead to an increase in recombinant product production (Yang & Butler, 2000b). While the use of Glutamine Synthetase (GS) expression system has reduced the likelihood of extreme level of ammonia in culture, it has not fully eliminated the issue from CHO bioprocesses.

In this study, we use label free tandem mass spectrometry to identify proteomic changes to three CHO cell lines exposed to 10 mM and 30 mM ammonia. Two IgG producers, NISTCHO and CHO-DP12 were used, with proteomic analysis performed 48- and 120 hours after ammonia addition. Both concentrations reduced cell growth and product titre, with 10

mM causing a mild increase in productivity in both producer cell lines. Gene Ontology analysis showed the tricarboxylic acid cycle, metabolic and catabolic processes, and cellular biogenesis processes are all affected by elevated ammonia. ER membrane bound proteins are more affected by elevated ammonia than those found in the ER lumen. In longer-term cultures, transmembrane proteins linked to Golgi-mediated pathways are upregulated, while vesicle transport associated proteins are downregulated, indicating altered intracellular trafficking under ammonia stress.

In this work, I was responsible for the culturing of NISTCHO and CHO-DP12 cells, including stressing the cells, viability and productivity assays and cell pellet collection. I prepared samples from these cell lines for proteomic analysis and performed differential expression study, through MS analysis. In the manuscript preparation, I was the primary author for all sections, all data analysis and discussion of results.

Proteomic Analysis of Ammonia-Induced Stress in Chinese Hamster Ovary (CHO) Cell Cultures

David Ryan [1,2], david.ryan68@mail.dcu.ie

Michael Henry [1]; michael.henry@dcu.ie

Christiana-Kondylo Sideri [1,2]; christianakondylo.sideri3@mail.dcu.ie

Esen Efeoglu [4]; esen.efeoglu@tudublin.ie

Paula Meleady [1,2,3]*; paula.meleady@dcu.ie

[1]Life Sciences Institute, Dublin City University, Glasnevin, Dublin 9, Ireland.

[2]School of Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland.

[3] SSPC Research Ireland Centre for Pharmaceuticals, Limerick, Ireland

[4]School of Food Science and Environmental Health, Technological University Dublin, Grangegorman, Dublin, Ireland,

*Corresponding author

Tel: +353-1-7005910; Email: paula.meleady@dcu.ie

Author contribution statement

David Ryan: conceptualisation, investigation, methodology, data handling and writing of the manuscript. Michael Henry: MS methodology (supporting), MS data handling and validation (supporting). Christiana-Kondylo Sideri: investigation (supporting). Esen Efeoglu: methodology (supporting) and review of the manuscript (supporting). Paula Meleady: conceptualisation, funding acquisition, project administration, supervision and review and editing of the manuscript (lead). Authors have read and approved the submitted version.

Abstract

Ammonia (NH_3) and its ionic form ammonium (NH_4^+) are both metabolic waste products and essential nitrogen sources within Chinese hamster ovary (CHO) cell cultures. Although necessary for amino acid synthesis, excessive accumulation in the extracellular environment can exert stress, reducing cell proliferation and impairing the efficiency of recombinant protein production. Proper endoplasmic reticulum (ER) function is critical for CHO cells as biotherapeutic producers. Previous work has linked elevated ammonia concentrations to reduced productivity via altered N-glycosylation pathways, but its broader effects on ER biology remain unclear. In this study, we applied high-resolution mass spectrometry to perform a comprehensive analysis of changes in the ER proteome in CHO cells exposed to two ammonia concentrations, 10 mM and 30 mM, 48 and 120 hours after supplementation. Both conditions suppressed cell growth and reduced product titre; however, the 10 mM supplementation resulted in a minor increase in specific cell productivity. Gene Ontology analysis revealed that ammonia strongly affected the tricarboxylic acid cycle, as well as key metabolic, catabolic and biogenetic processes. Several ER membrane proteins, including HMGCR and PREB, were consistently downregulated. In extended cultures, transmembrane proteins linked to Golgi-transport were upregulated, while vesicle transport associated proteins were downregulated, indicating altered intracellular trafficking.

Key Words

Chinese hamster ovary (CHO) cells, Cell culture, Endoplasmic reticulum (ER), Glycosylation, Endoplasmic reticulum associated degradation (ERAD), Ammonia, Ammonium toxicity.

Introduction

Chinese hamster ovary (CHO) cells have become established as a cornerstone in biopharmaceutical production, representing the gold standard for the manufacturing of

recombinant therapeutic proteins, including monoclonal antibodies, clotting factors, and other biologics (Fischer et al., 2015). Their widespread use stems from their ability to perform post-translational modifications, such as glycosylation, that are critical for the efficacy and safety of therapeutic proteins intended for human use (Kim et al., 2012). Despite these advantages, the metabolic demands of high-density CHO cultures and the associated accumulation of metabolic byproducts pose significant challenges to process efficiency and product integrity.

Ammonia is one of the metabolic byproducts produced during CHO cell culture. Ammonia is predominantly generated through glutamine metabolism, a key energy and biosynthesis pathway in rapidly proliferating cells (Pereira et al., 2018). While glutamine is an essential nutrient for CHO cultures, its catabolism yields equimolar quantities of ammonia (Hall et al., 1996), which accumulates over time, particularly in fed-batch and perfusion culture systems designed for prolonged production runs. During high-cell-density CHO cultures with viable cell densities (VCD) up to 50 million cells per mL, ammonia levels have been measured as varying between 10 and 20 mM (Fan et al., 2015; Xu et al., 2017). Previously, 33 mM has been reported as the ammonia concentration that causes a 50% reduction in growth (IC-50) in EPO-producing CHO cells (Yang and Butler, 2000a). Consequently, ammonia accumulation remains a significant bottleneck in the optimisation of CHO-based bioprocesses. Various methods have been studied in the pursuit of overcoming this problem. Given that glutamine metabolism generates ammonia, substitution of glutamate has been trialled, but cell growth was found to decrease (Capiaumont et al., 1995). The addition of threonine, proline, and glycine supplements was found to improve CHO cell growth and recombinant protein production in ammonia-stressed cells (Chen and Harcum, 2005). Five enzymes (Carbamoyl phosphate synthetase I, Ornithine transcarbamylase, Argininosuccinate synthetase, Argininosuccinate lyase, and Arginase) are involved in the urea cycle, converting ammonia into urea, which can be excreted from the body (Mitchell et

al., 2009). The incorporation of between three, four or five of these enzymes into the CHO cell genome using transfection was shown to increase cell viability by up to 44% and lower ammonia levels by up to 40% compared to non-transfected cells, with the cells expressing the first four enzymes showing the greatest ability to remove ammonia and greatest cell viability (Chung et al., 2003).

The impact of ammonia stress on CHO cells has been the focus of significant research efforts. Early investigations concentrated on characterising the kinetics of ammonia production and its direct cytotoxic effects on CHO cultures (Henrik A Hansen and Emborg, 1994; Hayter et al., 1991). More recent studies have focused on the cellular biology of CHO cells exposed to elevated ammonia levels, showing its impact on culture growth (Chitwood et al., 2021), energy metabolism (Chen and Harcum, 2007) and cell cycle (Chen and Harcum, 2007).

The impact of ammonia concentrations on the productivity of CHO cells, along with the underlying mechanisms that govern accurate protein production, constitutes a significant area of inquiry, particularly given the inevitable accumulation of ammonia in cell cultures. However, there has been relatively little study into this, with few in-depth studies of biological process and pathways affected by elevated ammonia. Elliott et al (2020) showed that CHO cells exposed to an additional 10 or 30 mM ammonia had reduced viable cell density as well as IgG productivity (Elliott et al., 2020). Moreover, ammonia can impair the fidelity of post-translational modifications, especially glycosylation, by altering intracellular pH and the balance of critical cofactors in the Golgi apparatus (Gawlitzek et al., 2000). Ha et al (2015) found 13 N- glycosylation related genes up-regulated in CHO cells exposed to 10 mM ammonia. It has also been found that ammonia can affect the glycosylation profile of the recombinant protein produced. Glycosylation is crucial for the correct structure of proteins and is a key function of the ER. Increased ammonia concentrations can interfere with glycosylation events within the Golgi apparatus, whereas glycosylation processes in the

ER, such as N-glycosylation, are less disturbed by elevated ammonia than later Golgi-based modifications like galactosylation and sialylation (Chen and Harcum, 2006). Often, errors in glycosylation can be linked to the ER and are related to the ER stress response through the unfolded protein response (UPR) and ER-associated degradation (ERAD) (Aebi, 2013; Chen and Harcum, 2006; Gerlach et al., 2012; Ha et al., 2015). This indicates that elevated ammonia affects CHO cell glycosylation through a differing mechanism than that seen in well studied ER stress inducers. This study aims to increase the understanding of the changes in the ER proteome in elevated ammonia cultures. ER function is crucial to the production of useable biologics.

The use of high-throughput proteomics technologies, such as MS-based proteomics, has been widely used to allow researchers to study the cell's protein response to stress. Label-free tandem mass spectrometry has been previously used to study the ER stress response of CHO cells to thapsigargin (Ryan et al., 2023), tunicamycin (Selvaprakash et al., 2024) and mild hypothermia (Ryan et al., n.d.). This study represents the first use of high-throughput label-free MS/MS to explore the response of CHO cells to elevated ammonia levels in culture.

In this study, we investigated the effects of elevated ammonia levels on two IgG producer cell lines, CHO-DP12 and NISTCHO, cultured in different media to identify both cell line specific changes and robust, common proteomic changes. Using label-free quantification and tandem mass spectrometry, we were able to study the differential expression of proteins from across the whole cell proteome, with a focus on ER-resident proteins. This was achieved by the supplementation of two concentrations of ammonia, 10 mM and 30 mM, added after 48 hours of culture, with viability and productivity studied across the period leading up to 120 hours post exposure. While 30 mM ammonia would not be reached in most standard cultures, this concentration was chosen to represent the most extreme condition. This was done to allow for greater insight into the CHO response to elevated ammonia. The

method of spiking-in of ammonia into cultures has been used as the most common method of modelling ammonia stress, despite its difference to gradual ammonia build-up in conventional culture (Chen and Harcum, 2007; Chitwood et al., 2021; Ha et al., 2015; Synoground et al., 2021; Yang and Butler, 2000a). The early addition of ammonia ensures that all replicates experience the same stress level at the same time, improving reproducibility and making it easier to attribute observed effects to ammonia exposure rather than culture variability. The clear, defined induction of stress also allows for greater confidence in linking mechanistic changes to ammonia level rather than other changes in the culture environment or cell cycle. While it may not directly mirror the bioprocess conditions, it can still provide insight into the stress response. Proteomic analysis was carried out in both cell lines at the 48-hour and 120-hour timepoint post supplementation. In this in-depth proteomic study of ER proteins in CHO cells exposed to elevated ammonia levels, we aim to improve our understanding of the impact of excessive ammonia on cell culture growth and productivity, as well as identify the mechanisms CHO cells use to cope with this stress, and to uncover potential cell line engineering targets that may enhance tolerance to excess ammonia, allowing for extended culture lifespans.

Methodology

Cell culture conditions

CHO-DP12 cells (ATCC CRL-12445 clone #1934), which produces an anti-IL8-IgG, were cultured in suspension at 37°C under 5% CO_2 and 80% humidity at 170 RPM in BalanCD CHO Growth A (FujiFilm, NL) with 4 mM L-glutamine (Sigma Aldrich, IE) supplementation in a Climo-Shaker (Kühner, SW). NISTCHO (National Institute of Standards and Technology), a CHO-K1 derived cell line, (Dahodwala et al., 2025; "NISTCHO," 2016)), a fully humanized IgG1κ antibody producer, was cultured in the same environmental conditions in EX-CELL® ADVANVCED CHO Medium (Merck, IE)

supplemented with 4 mM L-glutamine. Cells were grown in a 5 mL working volume of media in 50 mL TubeSpin BioReactor Tubes (Merck, IE), with routine passaging at 2×10^5 cells/mL. The cells were tested in-house for Mycoplasma contamination and found to be negative. CHO-DP12 cultures were pulsed with 400 nM Methotrexate every 2-3 weeks to ensure only producing cells remained in culture.

Ammonia Supplementation

CHO-DP12 and NISTCHO cells were seeded at 2×10^5 cells/mL, cultured in conditions as described above and assayed using the COUNTESS II (Invitrogen, Thermo Scientific, IE) and trypan blue (Invitrogen, Thermo Scientific, IE) daily, as per manufacturer's instructions, to generate viable cell number and viability data. After 48 hours, cells were stressed through the addition of 10 mM or 30 mM ammonium chloride to the culture media, with the relevant growth medium used to normalise the volume of the 0 mM (non-supplemented) cultures to the 10 mM and 30 mM ammonia-supplemented cultures. Separate biological triplicate samples were set for the collection of media and cell pellets alongside the triplicate samples used for viability assays.

IgG Titre Assay

IgG titre from CHO-DP12 and NISTCHO was measured using Valita Titer Rapid high-throughput IgG quantification assay 96-well (2.5-100mg/L) (Valita Cell, Beckman Coulter, VAL003), as per manufacturer instructions, using IgG1 Kappa from human myeloma plasma standards (Merck, IE). Specific productivity (pg/cell/day) was calculated as described in Clarke et al (2011) (Clarke et al., 2011).

Quantitation of Ammonia Levels in Culture Media

The ammonia levels present in culture media was determined using an ammonia assay kit (Merck, MAK538-1KT) as per manufacturer's instructions. Triplicate culture media samples were collected for both time points and centrifuged (300 x g for 5 minutes) to remove cells

and cell debris. Samples were diluted using purified water to within the assay test range, as suggested by the manufacturer. Optical density was measured at 340 nm using a Multi-scan Go (Thermo Fisher Scientific, IE).

Protein extraction and Preparation for Mass Spectrometry

Cells (3 x 10^6 per sample) were harvested, subjected to three centrifugation cycles (300 x g for 5 minutes), washed with ice cold PBS between each cycle, snap-frozen, and stored at -80 °C. Proteomic sample preparation for LC-MS/MS was carried out using the PreOmics iST kit, following the manufacturer's protocol (PreOmics GmbH, P.O.00001). A total of 100 µg protein per pellet was resuspended in 50 µL of supplied lysis buffer and heated at 95 °C for 10 minutes, which also reduced and alkylated the protein cysteine residues. Proteins were enzymatically digested using a Trypsin/LysC mix at 37 °C for 3 hours with agitation at 500 RPM. Digestion was stopped, and peptides were purified using the kit's cartridge system. The eluates were dried using a vacuum evaporator at 45 °C and stored at -20 °C.

Proteomic analysis using LC-MS/MS

Proteomic profiling was carried out using an UltiMate 3000 nanoRSLC system (Thermo Scientific) coupled to an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific). Tryptic peptides (1 µg per sample) were introduced onto a C18 trapping column (PepMap100, 300 µm × 5 mm (Thermo Scientific)) via nano-flow HPLC at a rate of 25 µL/min in 2% acetonitrile (ACN) and 0.1% trifluoroacetic acid (TFA) for a 3-minute desalting step. Peptides were then separated on an analytical column (Acclaim PepMap 100, 75 µm × 50 cm, 3 µm particles (Thermo Scientific)) before entering the mass spectrometer.

A 90-minute binary gradient was applied using two solvents: solvent A (0.1% formic acid in water) and solvent B (80% ACN, 0.08% formic acid). The gradient schedule consisted of 2–32% solvent B over 75 minutes, followed by an increase to 90% B over 5 minutes, and

maintained at 90% B for an additional 5 minutes. The flow rate during peptide elution was maintained at 300 nL/min. For ionisation, a spray voltage of 2.0 kV and a capillary temperature of 320°C were used. MS spectra were acquired using the Orbitrap analyser at 120,000 resolution (m/z 200), with a 50 ms maximum injection time and an AGC target of 4×10^5. Full MS scans were recorded across a mass range of 380–1500 m/z. A data-dependent top-speed method was employed for MS/MS acquisition, selecting precursors within a charge state of 2+ to 7+. After detection, dynamic exclusion was set to 60 seconds to prevent repeated selection. Fragmentation was performed by higher-energy collision dissociation (HCD) at a normalised collision energy of 28%, and product ions were analysed using the linear ion trap at low resolution.

Peptide identification was carried out using Proteome Discoverer version 2.4 (Thermo Scientific, UK) with SEQUEST HT against a *Cricetulus griseus* protein database, derived from a UniProt repository Taxon – 10029, proteome ID UP001108280 (downloaded September 2024). Search parameters within Proteome Discoverer included a precursor mass tolerance of 20 ppm, fragment mass tolerance of 0.6 Da, trypsin as the proteolytic enzyme, allowance for up to two missed cleavages, carbamidomethylation of cysteines as a fixed modification, and methionine oxidation as a variable modification. Further filtering selected only identifications with an XCorr score of ≥1.9. The identified peptides were then subjected to label-free quantification using Progenesis QI for Proteomics (NonLinear Dynamics, Waters, UK). Total numbers of proteins identified in the triplicate samples from each condition is listed in **Supplementary Data 3.**

Label-free quantitation was performed to assess differential protein abundance between cells cultured in 0 mM, 10 mM, and 30 mM ammonia conditions at both 48 hours and 96/120 hours post ammonia addition. Prior to statistical analysis, peptide abundance data were normalised and arcsinh-transformed to satisfy the assumptions of one-way ANOVA.

Peptides with a p-value ≤ 0.05 were selected for further analysis. Proteins were considered differentially expressed if they met the following criteria: (i) ANOVA p-value ≤ 0.05, (ii) identification based on at least two unique peptides, and (iii) a fold change of ±1.5 or greater between 0 mM (control) vs 10 mM and 0 mM (control) vs 30 mM ammonia-supplemented conditions.

KEGG pathway and STRING association analysis

Differentially expressed (DE) proteins between control (no ammonia supplementation) and test samples (10 mM or 30 mM ammonia supplementation) were input into STRING (https://www.string-db.org) (Szklarczyk et al., 2021), which was used to group the proteins based on protein-protein interactions and to provide groupings within KEGG biological pathways (https://www.genome.jp/kegg) (Kanehisa et al., 2023). DAVID was also used for KEGG pathway mapping (Huang et al., 2009; Sherman et al., 2022) ER proteins were listed according to the STRING GO term "GO:0098827 Endoplasmic reticulum sub-compartment". The statistical strength of the enrichment is given by the "signal" column. According to the STRING database "*signal is defined as a weighted harmonic mean between the observed/expected ratio and -log(FDR). FDR tends to emphasise larger terms due to their potential for achieving lower p-values, while the observed/expected ratio highlights smaller terms, which have a high foreground to background ratio but cannot achieve low FDR values due to their size. The signal measure seeks to balance both metrics for more intuitive ordering of enriched terms*."

Data Visualisation

Plotting of viability, titre and ammonia level data was performed using R, https://www.R-project.org/, package *ggplot2* (Wickham, 2016).

Results and Discussion

Extracellular Ammonia Concentration in Culture Media

In this study, the concentration of ammonia (mM) was measured using a colorimetric ammonia assay, using samples acquired at the time of supplementation (0hr), at 48 hours and 120 hours post-supplementation (**Figure 1**). Unsupplemented cultures produce more ammonia than supplemented culture, but given the added ammonia, total ammonia level is always higher in supplemented cultures. This is likely due to the greater cell count in the unsupplemented cultures when compared to the elevated ammonia cultures. The 0-hour measurements were between 2.27-2.87 mM ammonia, where 0 hours represents the time of supplementation, which is 48 hours post culture seeding. Final ammonia levels in the unsupplemented cultures increases approximately 4-fold, with values between 9.33-10.93 mM. The 30 mM cultures reach the highest ammonia concentration 48 hours after ammonia supplementation (NISTCHO-32.1, CHO-DP12-34.52). Considering the added ammonia, 10 mM cultures produced more ammonia than 30 mM cultures, likely due to the higher cell number at the lower concentration, with an increase of approximately 12 mM. Cells supplemented with 10 mM reach similar ammonia concentrations by the later timepoint (NISTCHO-12.40 mM, CHO-DP12-13.60 mM). The ammonia concentration in NISTCHO cultures is lower when compared to the CHO-DP12 cultures, under the same conditions. This could be due to the slower growth of NISTCHO cells or may be an effect of the different culture media. Consistent with previous studies, the initial spike and subsequent drop in ammonia levels suggest that metabolic pathways may be adapting to the supplementation (Chitwood et al., 2021; Synoground et al., 2021).

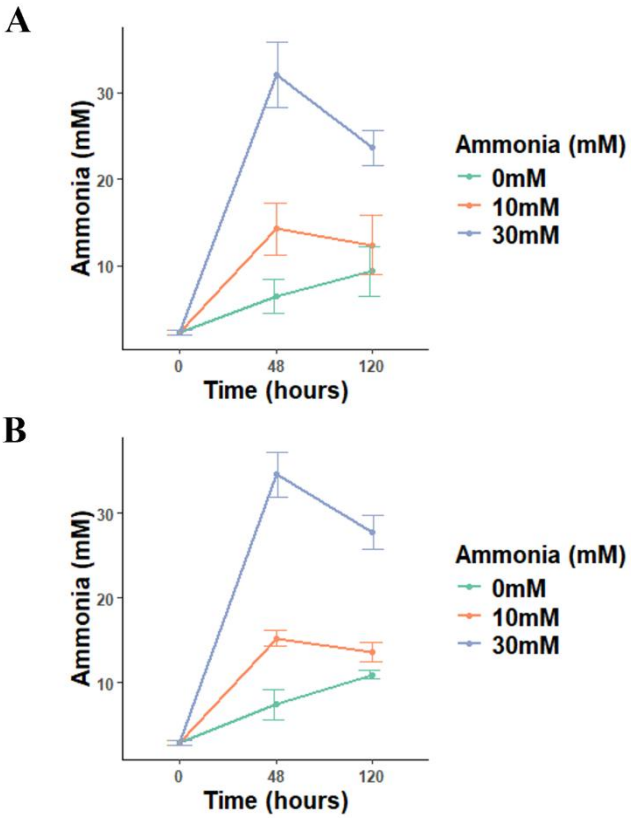


Figure 1. Ammonia concentrations present in the culture media from **(A)** NISTCHO and **(B)** CHO-DP12 cells cultured in 0 mM, 10 mM or 30 mM ammonia **where 0 hours is the time of ammonia supplementation**, which was 48 hours after culture seeding. In all plots, green represents 0 mM, orange represents 10 mM, and purple represents 30 mM ammonia supplemented culture. The data lines are the mean value of duplicate results, with error bars representing the standard deviation from the mean value at each time point.

The Glutamine Synthetase (GS) expression system is widely used in cell line development and can reduce ammonia levels within cultures by enabling glutamine synthesis. However, ammonia still accumulates from glutamine metabolism, media components, and cell lysis, even when exogenous glutamine is reduced or eliminated. Several approaches for implementing GS expression have been explored, with varying increases in expression. GS inhibitor methionine sulfoximine (MSX) selection is still utilised, but due to the ability of some CHO lines to tolerate MSX, not all non-producers will be screened out (Sanders and

Wilson, 1984). Bioinformatic analysis has shown that more than one GS gene is present in CHO cells, which may not be expressed under normal conditions. Deleting the main GS gene may lead to the activation of the second copy, so methods such as CRISPR are needed to ensure all related genes are deleted (Srila et al., 2023). GS system studies have also revealed novel target genes for further genetic optimisation (Budge et al., 2021; Cour Karottki et al., 2021), highlighting the continuing efforts to enhance the GS system and cell line selection options. Increasing our understanding of how ammonia can alter the CHO proteome can reveal elements of CHO biology that can be utilised in furthering production.

Cellular Viability and IgG Productivity

The build-up of waste products such as ammonia in cultures has long been recognised as an obstacle to culture growth and productivity. All cell lines in this study reflected this, with viable cell number decreased in the 10 mM and 30 mM ammonia cultures compared to unsupplemented cultures at both 48 hours and 120 hours post addition of ammonia, as shown in **Figure 2**. The cell number of the unsupplemented cultures (0 mM) increased steadily over time in culture, almost doubling every 24 hours in CHO-DP12, with the growth rate of NISTCHO being slower, which is in line with the manufacturer's findings with this cell line ("NISTCHO," 2016). Both concentrations of ammonia caused the growth rate of all cell lines to decrease. As time in culture increases, the 30 mM ammonia-supplemented cultures start to exhibit a faster decrease in viability compared to the 10 mM and 0 mM ammonia-supplemented cultures. However, cultures remain above 80% viability until 120 hours in both cell lines, when the viability of all cultures corresponding to the 0, 10 and 30 mM ammonia supplementation begins to decrease considerably. This indicates that higher levels of ammonia do not have an acute toxic effect on CHO cells, but the presence of elevated levels of ammonia is an unfavourable condition for cell growth and replication, as reported in other studies (Synoground et al., 2021; Yang and Butler, 2000b). While the addition of NH_4Cl can change the osmolarity of a culture, studies have found that once this factor is

"corrected" for, ammonia still exerts toxic effects (Lao and Toth, 1997; Yang and Butler, 2000b). Yang and Butler found that the addition of NaCl to culture did not affect recombinant EPO yield, and that for values less than 40 mM NaCl, there was no significant difference in cell count between NaCl and control cultures (Yang and Butler, 2000b). pH levels in cultures remained at 7 throughout the culture lifespan (data not shown).

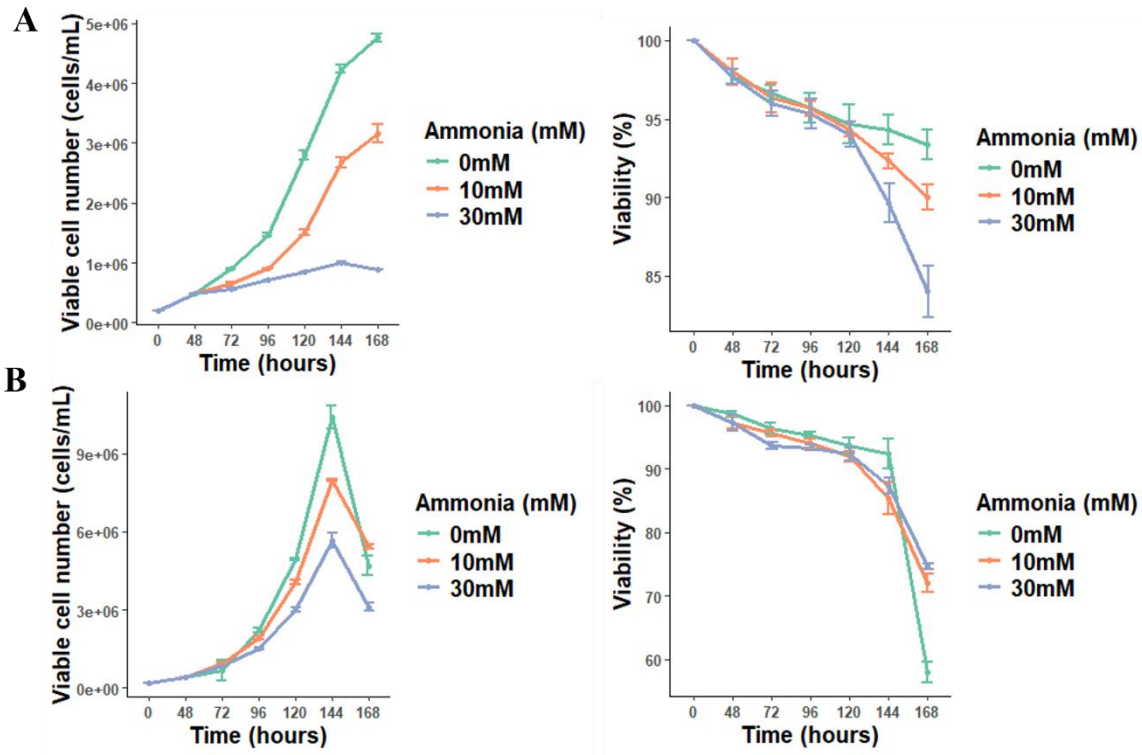
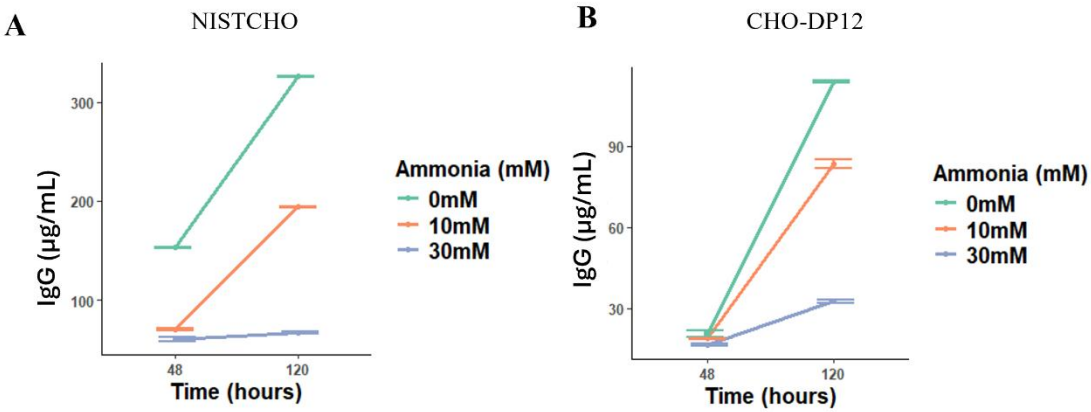


Figure 2. Viable cell density (cells/mL) and percentage cell viability for **(A)** NISTCHO and **(B)** CHO-DP12 cells cultured in 0 mM, 10 mM and 30 mM ammonia **where 0 hours is the time of ammonia supplementation**, which was 48 hours after culture seeding. In all plots, green represents 0 mM, orange represents 10 mM, and purple represents 30 mM ammonia supplemented culture. The data lines are the mean value of triplicate results, with error bars representing the standard deviation from the mean value at each time point.

Elevated ammonia levels also have an impact on product titre in producer cell lines, as shown in **Figure 3**. Ammonia supplementation (10 mM and 30 mM) reduced IgG titre in both cell lines at the 48-hour and 120-hour timepoints. Cell specific productivity (Qp) also decreased

in the 30 mM ammonia conditions; however, in both producer cell lines following 10 mM ammonia supplementation, the individual cells were more productive than the cells without ammonia supplementation (0 mM). These results align with previous studies, where higher levels of excess ammonia have reduced productivity (Elliott et al., 2020; H. A. Hansen and Emborg, 1994). A moderate increase in ammonia has been found to reduce titre but increase productivity (Yang and Butler, 2000b). In cultures with higher levels of excess ammonia, productivity is decreased. The lower overall titre, but similar specific productivity in 10 mM supplemented cultures aligns with results previously reported for this concentration (Synoground et al., 2021).



C

	QP(pg/cell/day)		
	0mM	**10mM**	**30mM**
NISTCHO	29.10	34.04	4.8
CHO-DP12	10.81	11.06	4.02

Figure 3. IgG titre from **(A)** NISTCHO and **(B)** CHO-DP12 cells following 0 mM, 10 mM, and 30 mM ammonia supplementation. **(C)** Specific cell productivity (Qp) from both cell lines measured at 120 hours post ammonia supplementation. Time points refer to 48- and 120-hours post ammonia addition, which happened 48 hours after culture seeding. In all plots, green represents 0 mM, orange represents 10 mM and purple represents 30 mM. The

data lines are the mean value of triplicate results, with error bars representing the standard deviation from the mean value at each time point.

Proteomic Analysis of CHO Cell Lines in Response to Ammonia Supplementation

Proteomic analysis was performed by comparing non-treated cells (0 mM) with those exposed to 10 mM or 30 mM ammonia supplementation at two time points: an early stage (48 hours post-ammonia addition) and a later stage of culture lifespan (120 hours). We quantified approximately 5000 proteins across all conditions, and out of these approximately 3000 were identified with two or more matched peptides. Proteins were defined as differentially expressed (DE) if they displayed a fold change in abundance of ≥1.5 between conditions, had an ANOVA p-value ≤ 0.05, and were identified by at least two unique peptides. By comparing two IgG-producing cell lines across different media, we identified conserved proteomic shifts that represent robust stress responses. These findings highlight potential ammonia-induced pathways directly implicated in the regulation of cellular productivity. By studying early and late culture stages, it is possible to observe both the immediate impact of elevated ammonia, and the adaptive changes cells undergo with prolonged exposure.

The number of upregulated and downregulated proteins for each cell line at both time points is summarised in **Table 1**. All DE proteins can be found in **Supplementary Data 1**. The high-producing IgG-producer NISTCHO shows a different proteomic response to ammonia supplementation than the low-producing IgG-producer CHO-DP12. This is seen in the total number of DE proteins between the non-treated cells and cells exposed to both ammonia concentrations (Table 1). NISTCHO is more responsive to elevated ammonia, compared to CHO-DP12, with more alterations in the proteome. This suggests that the level of ammonia has a greater impact on cellular function in NISTCHO than in CHO-DP12 early in culture lifespan, which mirrors the reduced growth seen in this cell line. The shared numbers of DE

proteins are shown in **Figure 4**. This figure demonstrates the distinct number of DE proteins found in NISTCHO 48 hours post ammonia supplementation. This is suggestive of an ammonia stress response that changes with time, with cells adapting their biological pathways the longer they are exposed to excess ammonia. This proteomic analysis detects only a small amount of transcription factors, which have been shown to be difficult to be comprehensively quantified in whole cell proteomics, where abundant proteins reduce the sensitivity for detecting less abundant proteins and their interactions (H. A. Hansen and Emborg, 1994). Fractionation techniques can be utilised to gain greater insight into other sub-cellular compartments, with nuclear fractionation allowing for greater insight into transcription factors (Dasarathy and Hatzoglou, 2018), which represent an important avenue for future studies into ammonia stress.

Table 1: The number of proteins found to be differentially expressed in NISTCHO and CHO-DP12 exposed to 10 mM and 30 mM ammonia supplementation, at 48 hours and 120 hours. Upregulated proteins are proteins more abundant under increased ammonia conditions, and downregulated proteins are more abundant under normal conditions.

	48 hours				***96/120 hours***			
	10 mM		**30 mM**		**10 mM**		**30 mM**	
	Up	Down	Up	Down	Up	Down	Up	Down
NISTCHO	42	157	162	387	379	58	591	440
CHO-DP12	69	25	142	143	192	235	332	380

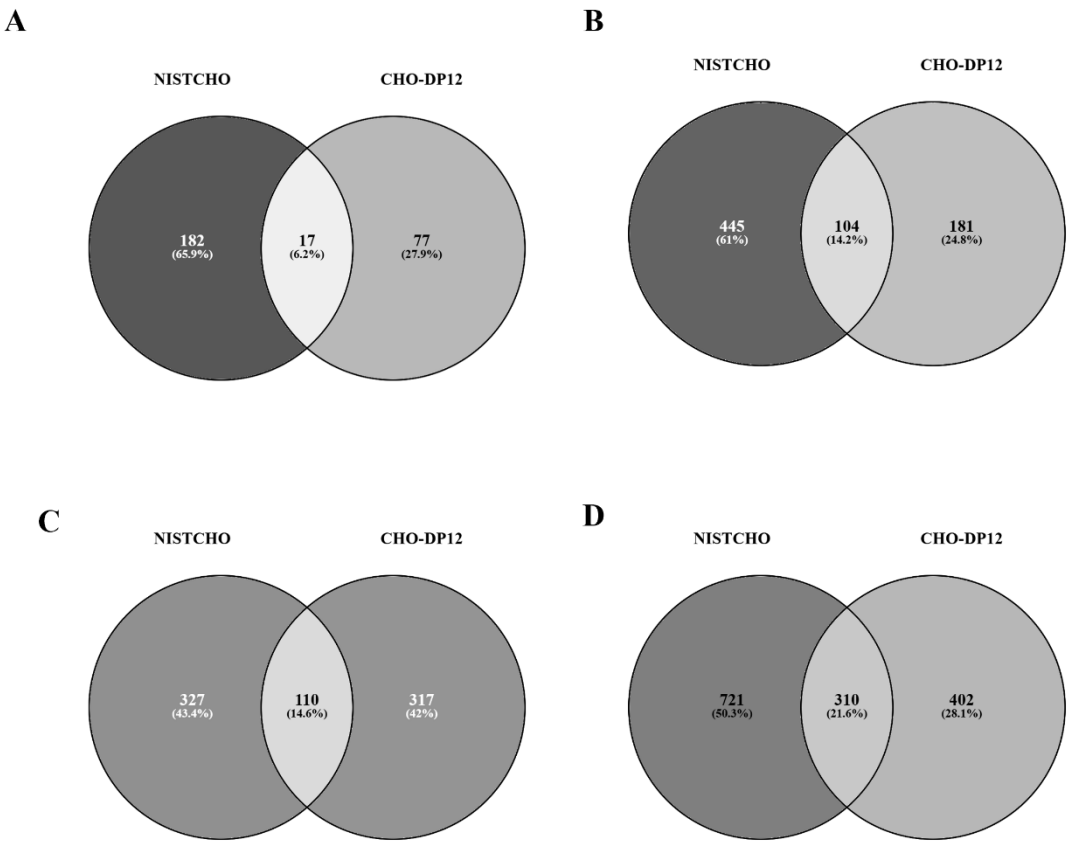


Figure 4. The number of shared differentially expressed proteins between NISTCHO and CHO-DP12 at both time points. **(A)** 48 hours post 10 Mm ammonia supplementation, **(B)** 48 hours post 30 mM ammonia supplementation, **(C)** 120 hours post 10 mM ammonia supplementation, **(D)** 120 hours post 30 mM ammonia supplementation. *Figure generated using Venny 2.1.*

Biological Processes and Pathways Affected by Elevated Ammonia Levels

Analysis of biological process enrichment was carried out using the STRING database. At 48 hours post-supplementation (10 mM), upregulated NISTCHO proteins mapped to 11 biological processes, compared to only three in CHO-DP12, with one in common, i.e. 'Cellular process (GO:0009987). At 48 hours post-10 mM ammonia supplementation, downregulated proteins mapped to 67 biological processes in NISTCHO and two in CHO-

DP12, with one of these processes 'Regulation of cellular component organisation (GO:0051128)' common to both lines.

A total of 82 biological processes were commonly enriched by downregulated proteins in both cell lines 48 hours after 30 mM ammonia supplementation. Similar to the 10 mM ammonia-supplemented cultures, several metabolic processes were affected including 'cellular metabolic process' (GO:0044237), 'nitrogen compound metabolic process' (GO:0006807), and 'organic cyclic compound metabolic process (GO:1901360).

By 120 hours a number of trends emerged. For example, several metabolic processes ('cellular metabolic process' (GO:0044237), 'nitrogen compound metabolic process' (GO:0006807), and 'organic cyclic compound metabolic process (GO:1901360)) were downregulated in both cell lines with 10 mM ammonia supplementation. Many of these processes were also downregulated in both cell lines with 30 mM ammonia supplementation, including additional pathways related to cellular component organisation, nuclear organisation and protein folding. Other elements of metabolic processes, including 'organic substance catabolic process' (GO:0071704), 'small molecule metabolic process' (GO:0044281) and 'alcohol metabolic process' (GO:0006066) were upregulated at 120 hours with 30 mM ammonia supplementation, along with several processes related to biosynthesis and cell cycle. In these cultures, biological pathways relating to cell cycle progression (cge04110), ribosome function (cge03010) and DNA replication (cge03030) were downregulated. The higher initial ammonia level may limit these important functions for cell growth and replication, leading to lower cell numbers. This, along with a reduction in metabolism pathways (cge01200, cge01212, cge00620, cge00480), may reduce the ammonia being produced by cells in these cultures. All processes, functions and pathways are listed in **Supplementary Data 2.**

The molecular functions of the downregulated proteins 48 hours after 10 mM ammonia supplementation are predominantly related to different forms of 'binding' in NISTCHO, including heterocyclic compound binding, nucleotide binding, protein-containing complex binding and nucleic acid binding. This suggests that excess ammonia may interfere with cellular processes which are reliant on these binding types, such as transcription, translation, and signal transduction. While the effect of ammonia on these cellular processes in CHO cells is not well characterised, studies in other cells and tissues have shown suppression of DNA replication and altered transcriptional and translational regulation in high ammonia environments (Dasarathy and Hatzoglou, 2018; Li et al., 2022; Zhang et al., 2015). The molecular functions that are most significantly enriched between the elevated ammonia and control (0 mM) cultures are noted in **Table 2** and **Table 3**.

Table 2: Molecular functions with the strongest enrichment signal from differentially expressed proteins from NISTCHO and CHO-DP12 cells exposed to 10 mM and 30 mM ammonia compared to 0 mM control cells, 48 hours after ammonia supplementation.

	10 mM			**30 mM**		
	GO ID	*Term*	*Signal*	*GO ID*	*Term*	*Signal*
NISTCHO	GO:0140657	ATP-dependent activity	1.74	GO:0140657	ATP-dependent activity	1.71
	GO:0016887	ATP hydrolysis activity	1.65	GO:0003724	RNA helicase activity	1.69
	GO:0003723	RNA binding	1.47	GO:0003723	RNA binding	1.61
	GO:0016462	Pyrophosphatase activity	1.43	GO:0000166	Nucleotide binding	1.49
	GO:0017111	Nucleoside-triphosphatase activity	1.39	GO:0016874	Ligase activity	1.47
CHO-DP12	**10 mM**			**30 mM**		
	GO ID	*Term*	*Signal*	*GO ID*	*Term*	*Signal*
	GO:0003725	Double-stranded RNA binding	1.03	GO:0003723	RNA binding	1.4
	GO:0005515	Protein binding	0.65	GO:0097367	Carbohydrate derivative binding	0.96
	GO:0019899	Enzyme binding	0.57	GO:0036094	Small molecule binding	0.92
	GO:0003723	RNA binding	0.56	GO:1901363	Heterocyclic compound binding	0.9
	GO:0005488	Binding	0.51	GO:0097159	Organic cyclic compound binding	0.89

Table 3. Molecular functions with the strongest enrichment signal from differentially expressed proteins from NISTCHO and CHO-DP12 cells exposed to 10 mM and 30 mM ammonia compared to 0 mM control cells, 120 hours after ammonia supplementation.

	10 mM			**30 mM**		
NISTCHO	*GO ID*	*Term*	*Signal*	*GO ID*	*Term*	*Signal*
	GO:0044877	Protein-containing complex binding	1.06	GO:0003723	RNA binding	1.59
	GO:0005515	Protein binding	0.98	GO:0140657	ATP-dependent activity	1.51
	GO:0036094	Small molecule binding	0.98	GO:0044877	Protein-containing complex binding	1.43
	GO:0019899	Enzyme binding	0.95	GO:0016462	Pyrophosphatase activity	1.34
	GO:0042802	Identical protein binding	0.95	GO:0000166	Nucleotide binding	1.33
CHO-DP12	**10 mM**			**30 mM**		
	GO ID	*Term*	*Signal*	*GO ID*	*Term*	*Signal*
	GO:0016874	Ligase activity	1.71	GO:0016874	Ligase activity	1.56
	GO:0044877	Protein-containing complex binding	1.26	GO:0003729	mRNA binding	1.44
	GO:0043531	ADP binding	1.23	GO:0036094	Small molecule binding	1.42
	GO:0036094	Small molecule binding	1.22	GO:0016853	Isomerase activity	1.41
	GO:0042802	Identical protein binding	1.19	GO:0000166	Nucleotide binding	1.4

Across all cell lines, more KEGG pathways are represented from proteins downregulated in ammonia-supplemented cultures than to upregulated proteins. KEGG pathways of interest enriched in DE proteins from 30 mM ammonia cultures are listed in **Table 4** and **Table 5**. The total number of proteins found in each pathway is included, as well as the number of DE proteins identified in each cell line. Pathways observed to be downregulated in NISTCHO 48 hours post 10 mM ammonia supplementation include 'apoptosis' (cge04210), 'carbon metabolism'(cge01200) and the 'citrate/TCA cycle' (cge00020), with 'apoptosis'

(cge04210) also being enriched by downregulated DE proteins in CHO-DP12. The TCA cycle is a crucial part of energy metabolism in the cell, and a reduction in TCA cycle activity would cause a reduction in the levels of ATP present for the cell to accomplish fundamental processes (Arnold and Finley, 2022). Previous findings have found that promoting TCA cycle metabolism with valine-fed culture leads to a reduction in ammonia accumulation (Savizi et al., 2022). The upregulated pathways include 'biosynthesis of amino acids' (cge01230), 'Glycolysis/Gluconeogenesis' (cge00010) and 'Proteasome' (cge03050). The TCA cycle is downregulated in the 30 mM ammonia supplemented cultures in both cell lines, with 'carbon metabolism' (cge00010), 'protein processing in the ER'(cge04141), 'apoptosis' (cge04210), and 'glycolysis/gluconeogenesis' (cge00010) also downregulated.

In CHO-DP12 cells 120 hours post 30 mM ammonia supplementation, 7 of 38 upregulated ER proteins involve 'ER to Golgi vesicle-mediated transport' (GO:0006888). Additionally, 8 of 19 downregulated proteins are assigned to the biological process 'Endomembrane system organization' (GO:0010256), which refers to the cell membrane, nuclear envelope, ER, Golgi apparatus, vesicles, and lysosomes. This suggests that proteins found in a variety of cellular membranes are affected by elevated ammonia, highlighting the wide-ranging effects of ammonia on CHO cells, predominantly linked to membrane-related biological pathways.

Table 4. KEGG Pathways found to be enriched in downregulated DE proteins from NISTCHO and CHO-DP12 cultures supplemented with 30 mM ammonia, 48 hours after supplementation. The total number of proteins in those pathways, as well as the number found to be differentially expressed in each cell line is shown. *All enriched pathways are listed in Supplementary Materials 2.*

KEGG Pathway	Total proteins	NISTCHO	CHO-DP12
		Downregulated	
Apoptosis	92	12	6
TCA cycle	27	9	3
Metabolic pathways	1132	65	18
RNA transport	135	12	7
Glycolysis / Gluconeogenesis	55	5	3
Cell cycle	99	6	5
		Upregulated	
Endocytosis	183	7	6
Metabolic pathways	1132	23	20

Table 5. KEGG Pathways found to be enriched in downregulated DE proteins from NISTCHO and CHO-DP12 cultures supplemented with 30 mM ammonia, 120 hours after supplementation. The total number of proteins in those pathways, as well as the number found to be differentially expressed in each cell line is shown. *All enriched pathways are listed in Supplementary Materials 2.*

KEGG Pathway	Total proteins	NISTCHO	CHO-DP12
		Downregulated	
Metabolic pathways	1132	41	64
Spliceosome	114	14	7
Aminoacyl-tRNA biosynthesis	36	6	6
Valine, leucine and isoleucine degradation	42	5	12
Aminoacyl-tRNA biosynthesis	36	6	6
		Upregulated	
Metabolic pathways	1132	102	56
Endocytosis	183	33	11
Proteasome	37	14	6
Protein processing in endoplasmic reticulum	128	20	16
Carbon metabolism	104	20	10
Ubiquitin mediated proteolysis	104	13	7
Biosynthesis of amino acids	67	10	8

Differential Expression of ER Resident Proteins in Elevated Ammonia Conditions

In this study, the proteins located within the ER were chosen as a focus point, given the crucial role played by the ER in protein folding, and the effect of other environmental stressors on ER function (Han et al., 2010; Rutkowski and Kaufman, 2004). For patient safety, recombinant products must be correctly folded, so it is important to understand the

ER pathways and the impact of culture environment changes on ER function. This was achieved by comparing the DE proteins with a curated list of 932 proteins from STRING, all of which are localised to the ER. The number of proteins from this location that were found to be DE can be found in **Table 6**. All ER localised DE proteins and relevant fold changes are listed in **Supplementary Data 4.** NISTCHO cells consistently yielded the largest number of DE proteins regardless of ammonia concentration or timepoint used in the study (**Supplementary Data 1).**

Table 6: The number of ER resident proteins found to be differentially expressed in NISTCHO and CHO-DP12 exposed to 10 mM and 30 mM ammonia supplementation, at 48 hours and 120 hours. ER resident proteins identified from 932 proteins in the STRING "Endoplasmic Reticulum" cellular location. Upregulated proteins are proteins more abundant under increased ammonia conditions, and downregulated proteins are more abundant under normal conditions.

	48 hours		***96/120 hours***	
	10 mM	**30 mM**	**10 mM**	**30 mM**
NISTCHO	12	29	40	70
CHO-DP12	7	14	20	57

The numbers of DE ER proteins shared between the cell lines are shown in **Figure 5**. More of these ER proteins are shared between cell lines at the later timepoint, suggesting that the long-term response to excess ammonia is more conserved, while the short-term response is more cell line dependent. Higher levels of ammonia can lead to increased nitrogen in culture, so proteins found in the biological process "Response to nitrogen compound" (GO:1901698) that are located in the ER were especially considered. This process contains 1149 proteins, taken from STRING. In NISTCHO, 108 individual DE proteins mapped to this process, and 96 individual DE proteins from CHO-DP12, with 53 proteins shared between both cell lines **(Figure 6)**. In these shared proteins, 17 proteins are found in the ER, including 6 proteins

linked to ERAD (PSMC6, RCN3, ERLEC1, CANX, UGGT1, PIG-O). The DE ER proteins that are common between NISTCHO and CHO-DP12 at each experimental condition are summarised in **Figure 7** and **Figure 8**. By comparing two cell lines producing different monoclonal antibodies, this study identifies proteomic shifts in the ER across varying ammonia concentrations. While some responses may be antibody-specific, the shared proteomic changes highlight robust, conserved mechanisms independent of individual cell line biology or recombinant product.

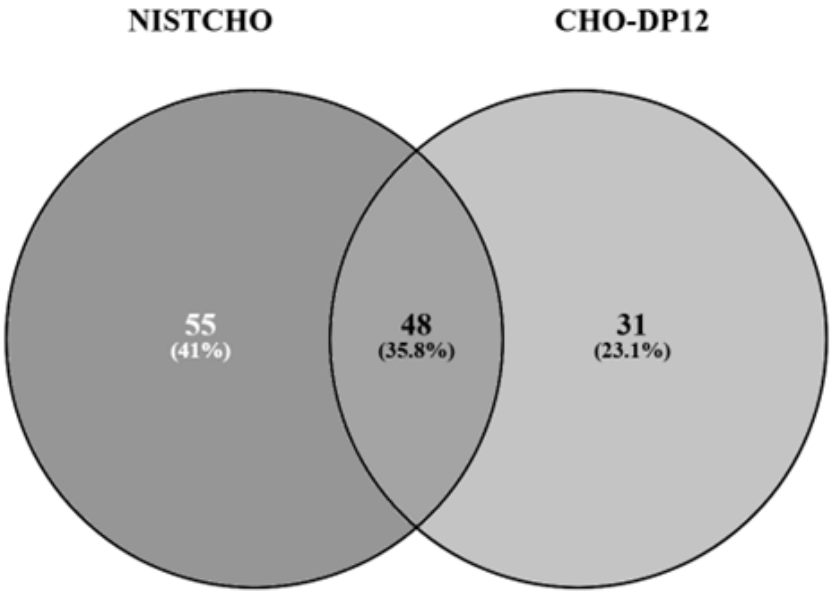


Figure 5. The number of shared differentially expressed proteins between NISTCHO and CHO-DP12 that are found to be differentially expressed in the Endoplasmic Reticulum. *Figure generated using Venny 2.1*

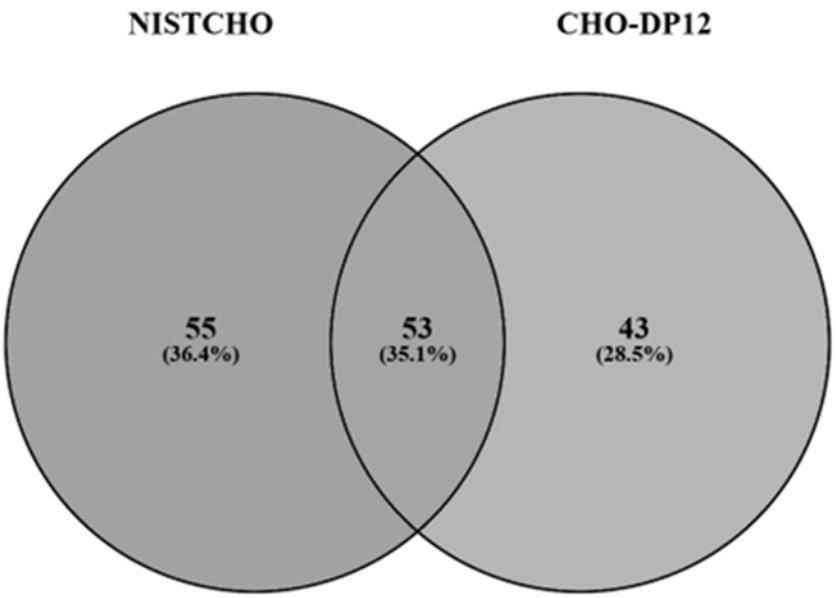


Figure 6. The number of shared differentially expressed proteins between NISTCHO and CHO-DP12 that are found to be differentially expressed in the biological process "Response to nitrogen compound". *Figure generated using Venny 2.1.*

48hrs

10mM

	NISTCHO	CHO-DP12
BiP	-1.62	-2.81
KTN1	-2.07	-2.22

30mM

	NISTCHO	CHO-DP12
SRPRA	-1.91	-2.16

120hrs

10mM

	NISTCHO	CHO-DP12
HSD17B7	+1.82	+1.63
KPNB1	+1.83	+1.51

30mM

	NISTCHO	CHO-DP12
ACSL1	+1.82	+1.60
ATP2A2	+1.58	+1.76
ATP6AP2	+1.91	+1.77
CANX	+2.75	+1.77
EPM2AIP1	-1.93	-3.44
HMOX2	+1.69	+1.85
LSS	+2.41	+2.01
MLEC	+2.12	+1.94
NCLN	-2.17	-1.64
PCYT1A	-17.70	-3.51
PDIA3	+1.70	+1.82
RAB18	+2.52	+1.60
RAB9A	+2.08	+2.02
RPN2	+1.58	+1.74
TMED1	+4.84	+3.88
TMED4	+2.24	+1.72

A negative figure (-) represents reduced expression in the ammonia supplemented cells compared to the control conditions and a positive figure (+) respresents increased expression in ammonia supplemented cells

Figure 7. The differentially expressed proteins that are common between NISTCHO and CHO-DP12 at each experimental condition, where the direction of the fold change is also shared. *All data for ER DE proteins can be found in Supplementary Data 4.*

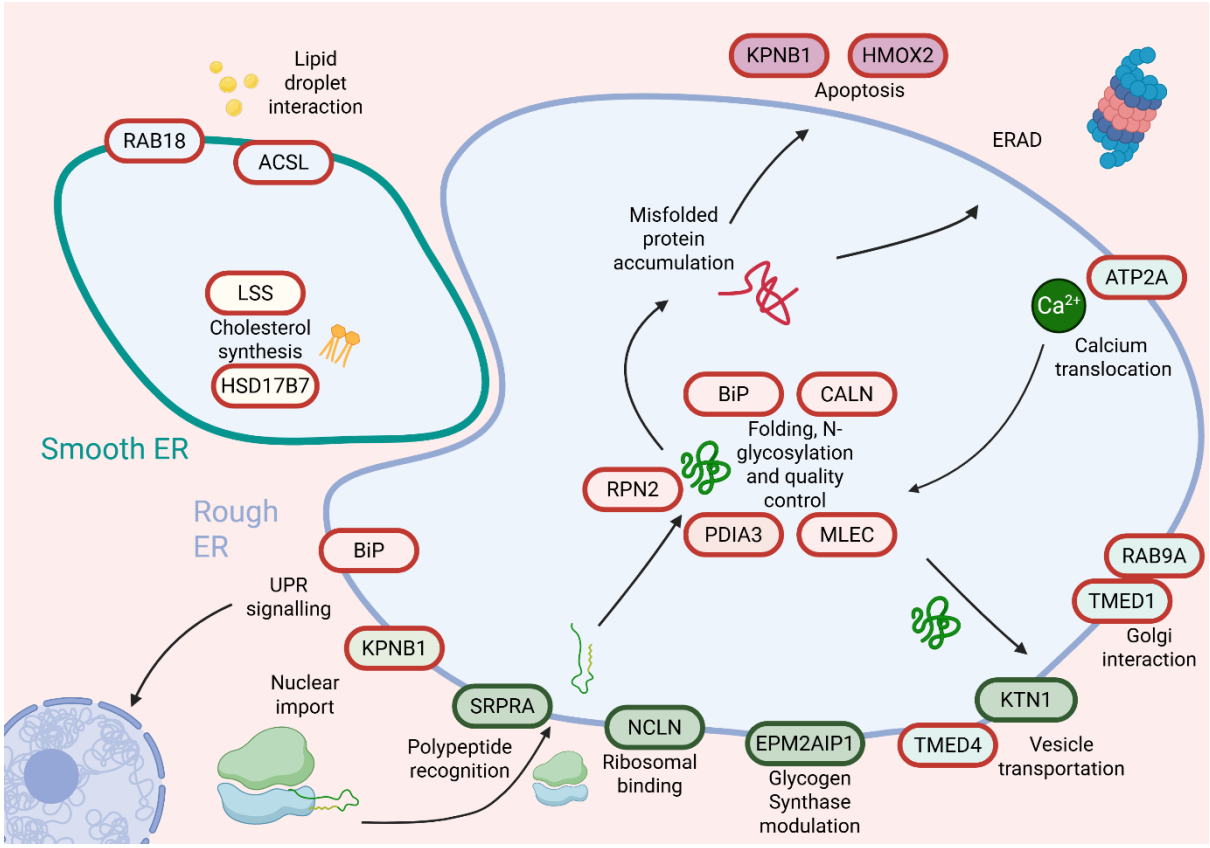


Figure 9. The differentially expressed ER proteins shared between cell lines (CHO-DP12 and NISTCHO) across both time points (48 and 120 hours) and ammonia concentrations (10 mM and 30 mM). A red circle represents a protein upregulated in elevated ammonia conditions and a green circle represents a protein downregulated in the elevated ammonia conditions. *Figure generated using BioRender.com.*

48 hours post ammonia supplementation

At 48 hours post-supplementation, DE ER proteins were predominantly upregulated in CHO-DP12 but downregulated in NISTCHO. Notably, NISTCHO showed a significant decrease in ER membrane-associated proteins. The most significantly downregulated protein is BCAP31, showing a 27.8-fold decrease in abundance. BCAP31 is a multi-pass transmembrane protein of the ER, involved in the export of secretory proteins, recognition of misfolded proteins, and their targeting to the ERAD pathway (Annaert et al., 1997). SRPR-α, a signal recognition particle receptor subunit, is downregulated in NISTCHO at

both tested concentrations, as well as in CHO-DP12 in 30 mM ammonia supplemented cultures. SEC11α is downregulated following 30 mM supplementation. SRPR-α is crucial for the correct targeting of the nascent secretory proteins to the ER membrane (Halic et al., 2006), while SEC11α catalyses the cleavage of N-terminal signal sequences from nascent polypeptides during their translocation into the ER lumen (Liaci et al., 2021). This indicates that elevated ammonia levels in culture may interfere with the ER function of NISTCHO cells by limiting the entry of unfolded polypeptides into the ER (**Figure 8**). The limited presence of ER stress markers during early ammonia stress could be explained by the reduction in polypeptides entering the ER, resulting in a reduction of proteins needing folding and a reduction in the accumulation of misfolded proteins which would trigger the UPR. In 30 mM ammonia-treated NISTCHO cells, PREB and SEC24C, two proteins involved in the assembly of the coat protein complex II/COPII in ER to Golgi vesicle-mediated transport were downregulated, while SEC31A, which has a similar function, was upregulated. PREB is involved in the activation of SAR1A and SAR1B by promoting the exchange of guanosine diphosphate (GDP) for guanosine triphosphate (GTP) in these small GTPases (Melville et al., 2020). The downregulation of PREB could therefore be related to the downregulation of many metabolic pathways seen in the GO analysis. Both SEC24C and SEC31A promote the formation of transport vesicles from the ER and the selection of cargo molecules (Tang et al., 2000; Tani et al., 1999). Luminally exposed glycosylphosphatidylinositol (GPI)-anchored membrane proteins cannot bind directly to SEC24 (Bonnon et al., 2010), so it is possible that the different regulation of these two similar proteins may be due to the cargo molecules they select for. In both cell lines, KTN1 is downregulated. This membrane protein is involved in kinesin-driven vesicle transport from the ER (Kumar et al., 1995).

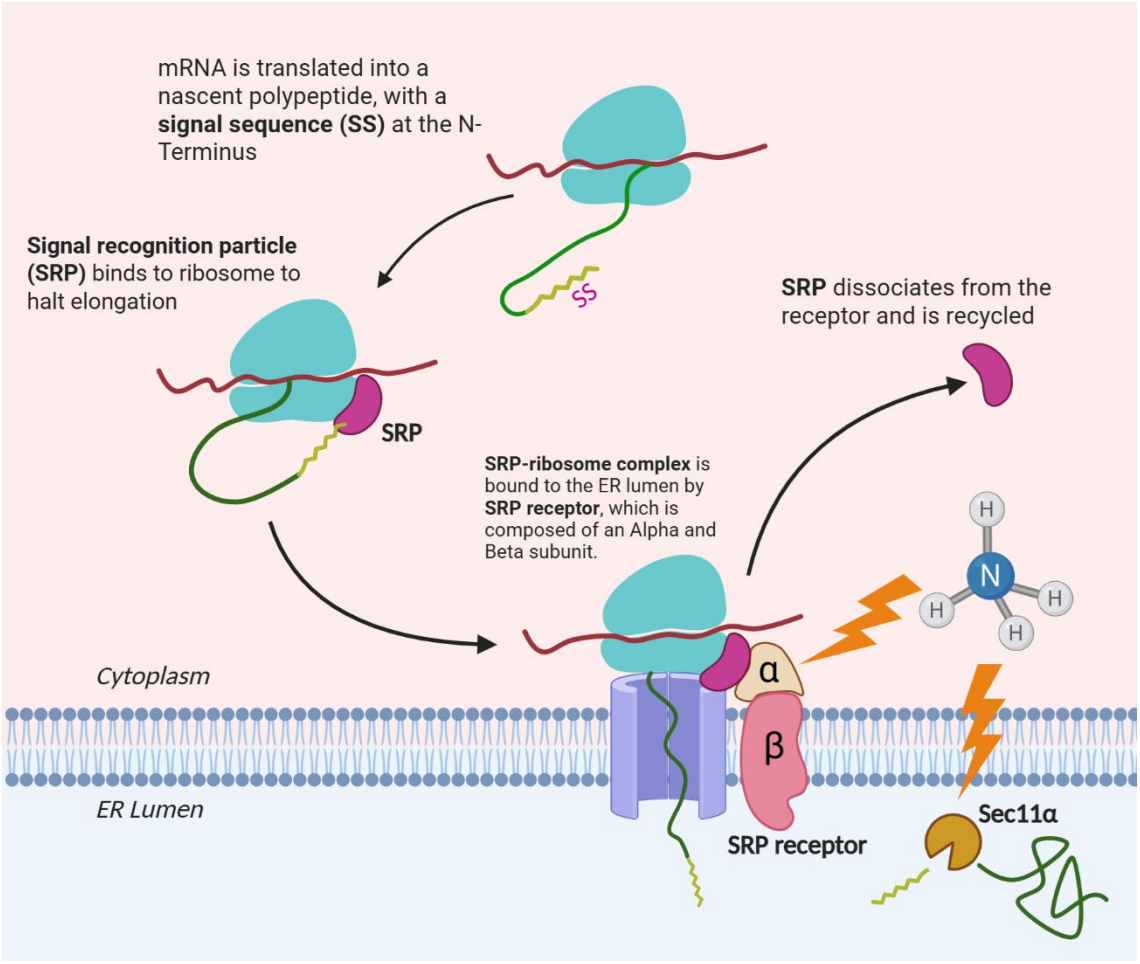


Figure 9. The presence of elevated ammonia in NISTCHO cultures results in a downregulation of the signal recognition particle (SRP) complex receptor-alpha subunit (SRPRα), which is vital in the correct targeting of nascent secretory polypeptides to the ER lumen. SEC11α is also downregulated in elevated ammonia cultures and is necessary for the proper cleavage of the N-terminus signal sequence from polypeptides in the ER lumen. Together, the downregulation of these proteins indicate that elevated ammonia interrupts the normal processing of the newly translated polypeptides in NISTCHO cells. SRPRα is also downregulated in 30 mM ammonia CHO-DP12 cultures. *Figure generated using BioRender.com*.

The ER chaperone BiP, responsible for the recognition of unfolded proteins (Dudek et al., 2009) was downregulated in both cell lines following 10 mM ammonia supplementation. This indicates a reduction in the need for nascent polypeptide recognition in cells exposed to ammonia. Other chaperone proteins essential for proper protein folding, oligomeric assembly, and quality control (Bedard et al., 2005) are differentially expressed. In the 30 mM ammonia-supplemented cultures, CALR is upregulated in NISTCHO and CANX is upregulated in CHO-DP12, along with another calcium-binding protein CALU, which is found in the ER, but has also been shown to be present in the Golgi apparatus (Vorum et al., 1999). Reticulocalbin 3 (RCN3), a molecular chaperone that facilitates calcium ion binding during protein biosynthesis [57], was upregulated under ammonia stress in both lines: at 48 hours (10 mM) in CHO-DP12 and at 120 hours (30 mM) in NISTCHO. Several members of the protein disulfide isomerase (PDI) family were differentially expressed in both cell lines. PDI proteins are crucial to protein folding and quality control in the ER through the recognition of exposed hydrophobic regions or cysteine residues that are hidden in correctly folded proteins [58]. In NISTCHO, PDIA3 is upregulated and TMX1 is downregulated, while ERP44 is upregulated in CHO-DP12. While the specific PDI proteins may not be the same, proteins of similar function are similarly regulated in both cell lines.

The differences in expression of these crucial ER chaperones between CHO-DP12 and NISTCHO highlight the importance of cell line-specific studies, with both IgG producers having differing ER responses to excess ammonia. It appears that short-term exposure to elevated ammonia elicits different responses in the two cell lines, activating stress responses in CHO-DP12, through the upregulation of proteins involved in folding and quality control. In NISTCHO most ER-related proteins are downregulated, suggesting that elevated ammonia levels can disrupt the function of ER transport pathways, rather than activating specific stress-related pathways.

120 hours post ammonia supplementation

Two Transmembrane emp24 domain-containing (TMED) proteins were upregulated following 30 mM ammonia supplementation in CHO-DP12, with three upregulated in NISTCHO. TMED1 has been linked to interleukin-33 (IL-33) signalling through the transport of the ST2 receptor to the cell membrane (Connolly et al., 2013). TMED2 and TMED4 are both involved in vesicular protein trafficking through the early secretory pathway (Aber et al., 2019). The upregulation of these transmembrane proteins possibly indicates an increase in misfolded proteins that must be trafficked out of the ER. In both cell lines, transmembrane (TMEM) proteins are potentially utilised to prevent protein aggregation and maintain product secretion rates despite the ammonia stress. This family of proteins has been linked with ER responses previously, but with limited study of their role in CHO cells (Sakabe et al., 2015; Zhang et al., 2021).

Between the two ammonia supplementation concentrations, 15 proteins involved in 'Golgi vesicle transport' (GO:0048193) were DE in NISTCHO, including three TMEM proteins and three SEC proteins (TMED1, TMED2, TMED4, SEC1, SEC23, SCE31a). Vesicle transport proteins were predominantly upregulated at 10 mM ammonia but shifted toward downregulation following 30 mM supplementation. This may indicate that higher levels of ammonia may interfere with the cell's ability to transport proteins out of the ER. For example, KPNB1 (Importin subunit beta-1), which was upregulated in both cell lines in the 10 mM supplemented cultures, is involved in import of proteins from the ER to the nuclear envelope. The inhibition of KPNB1 has been shown to impair proteostasis, induce UPR and induce apoptosis (Zhu et al., 2018). It is upregulated in both cell lines in the 10 mM supplemented cultures. Two RAB proteins, RAB9 and RAB18 were found to be upregulated in both cell lines. Both are small GTPase proteins involved in intracellular membrane trafficking and organelle communication (Vazquez-Martinez and Malagón, 2011). RAB9 acts as a regulator for the trans-Golgi network (Lombardi et al., 1993) while RAB18

regulates lipid metabolism through the ER membrane and lipid droplet interactions (Dejgaard and Presley, 2019).

Proteins related to the ubiquitin-dependent ERAD pathway (GO:0030433) were also differentially expressed at 120 hours. Following supplementation with 30 mM ammonia, UBXN4 was upregulated in NISTCHO, and UBXN7 was upregulated in CHO-DP12. UBXN proteins are a family of ubiquitin-binding proteins that play a role in regulating protein turnover and interacting with the ubiquitin-proteasome system (Rezvani, 2016). ERLEC1, an N-glycan recognition factor implicated in ERAD through its interaction with the membrane-associated ubiquitin ligase complex [69], was upregulated in NISTCHO cells across both ammonia concentrations. In CHO-DP12 cells, upregulation was observed following 30 mM supplementation. In both cell lines, elements of the glycosylation and calreticulin/calnexin cycles were upregulated in elevated ammonia conditions, including calreticulin, calnexin, BiP, RPN2, MLEC, and PDIA3. This is a further indication of a level of ER stress present at this later time point, in particular in response to 30 mM ammonia supplementation. The increase in ER chaperone expression in elevated ammonia cultures in both cell lines is suggestive of a fundamental change in ER activity in CHO cells in these conditions. ATP2A2, which catalyses the translocation of calcium from the cytosol to the sarcoplasmic reticulum lumen (Dhitavat et al., 2003), was upregulated in both cell lines following 30 mM ammonia supplementation. Several ER quality control chaperones are dependent on calcium in the ER to function (Michalak et al., 2002). This response has been well documented in ER stress studies in CHO cells using other types of ER stress induction (Prashad and Mehra, 2015; Roy et al., 2017; Samy et al., 2021), and from our study elevated ammonia also elicits an ER stress response.

HMOX1 and HMOX2 were upregulated in the 30 mM supplemented CHO-DP12 cell cultures and in the 10 mM supplemented NISTCHO cell cultures, with HMOX2 upregulated in NISTCHO in the 30 mM cultures, while HMOX1 is downregulated. These proteins are

linked to anti-apoptotic activity (Chen et al., 2018; Gozzelino et al., 2010; Zuo et al., 2025). These results may indicate that HMOX proteins play a role in mediating ammonia stress at higher concentrations. HMOX1 may play a greater role in NISTCHO adaptation to changing conditions over culture lifespan, with the unsupplemented cultures expressing a greater level of this protein at the later stage of culture. Previously, in CHO cells where apoptosis-related genes in the *bcl-2* family were overexpressed, apoptosis was not detected in cultures with concentrations up to 50 mM ammonia (Mastrangelo et al., 2000). The upregulation of HMOX proteins across cell lines and ammonia concentrations suggests that CHO cells may utilise this anti-apoptosis pathway to survive in elevated ammonia conditions.

In NISTCHO cells, long-chain fatty-acid-coenzyme A ligase family proteins were upregulated following ammonia supplementation, two in the 10 mM and three in the 30 mM cultures. This family of proteins are crucial factors in *de novo* lipid synthesis, fatty acid catabolism, and remodelling of cellular membranes (Soupene and Kuypers, 2008). ACSL1 was upregulated in both the 10 mM and 30 mM NISTCHO cultures and in the 30 mM CHO-DP12 cultures. This protein is involved in catalysing the conversion of long-chain fatty acids to an active form, allowing for both synthesis and degradation via beta-oxidation of cellular lipids (Ohkuni et al., 2013). Given the role of ACSLs in membrane maintenance and remodelling (Soupene and Kuypers, 2008), this reiterates the effect of elevated ammonia on cellular membranes, and that the producer cell lines can utilise these proteins to attempt to maintain membrane integrity. The common expression of ACSL proteins is suggestive of a CHO cell response to cell maintain membrane integrity against ammonia toxicity. Following 10 mM supplementation, both cell lines exhibited increased levels of HSD17B7, a key enzyme involved in *de novo* cholesterol synthesis and part of a larger family of proteins linked with metabolism and synthesis of lipids (Heikelä et al., 2020). EPM2AIP1 is downregulated in both cell lines in 30 mM ammonia supplemented cultures. This protein

associates with the cytosolic face of the ER, and binds to laforin, which may play a role in correct glycogen synthesis (Turnbull et al., 2013).

A greater number of DE ER proteins are common between both cell lines at the later timepoint, suggesting that long-term ammonia exposure activates a more innate CHO response, in particular in the 30 mM supplemented cultures. Elevated ammonia appears to disrupt ER membrane-associated functions and transport of proteins to the Golgi apparatus. This potentially reduces import of nascent polypeptides into the ER and induces an ERAD response at higher concentrations and over longer exposure. With longer exposure, the response of both cell lines converges on ERAD activation and the upregulation of transmembrane proteins, indicating an increased need for the trafficking of misfolded proteins out of the ER.

Conclusions

In response to elevated ammonia levels, CHO cells exhibit reduced proliferation and decreased productivity compared to cells maintained under standard batch culture conditions. This decline in growth and protein production may be associated with the downregulation of key biological pathways, including those involved in cellular metabolism, nitrogen compound metabolism, and organic cyclic compound metabolism. Proteins related to the ER membrane are amongst the ER proteins most affected by elevated ammonia levels. This is an indication that there is a limited effect of elevated ammonia on proteins in the ER lumen, directly related to protein folding, and that most interference in ER function is related to the function of the ER membrane. In both cell lines, TMEM proteins are upregulated at the later time point, most related to Golgi-mediated transport, whereas SEC proteins are downregulated, indicating cells may decrease the formation of transport vesicles in elevated ammonia cultures. The upregulation of anti-apoptosis factors HMOX1 and HMOX2 in CHO cells exposed to excess ammonia suggests that they could play an important role in cell

survival in elevated ammonia conditions. CHO cell anti-apoptosis mechanisms could be exploited to create more robust cell lines, resilient to varied environmental stressors, and potentially useful in GS expression-derived cells where ammonia may not reach inhibitory levels.

The altered expression of ER stress related proteins along with indications of ERAD activity are evident at the later time point, particularly in cells exposed to the higher concentrations of ammonia supplementation. These include UBXN4, UBXN7, FAF2 and ERLEC1 which suggest that elevated ammonia does result in some build-up of misfolded proteins in the ER, which requires ubiquitin-mediated degradation. Other hallmarks of ER stress, such as calreticulin, calnexin, BiP and PDIA3, are here linked to ammonia stress. The induction of ER stress could explain the reduction in product titre and proper glycosylation previously associated with elevated ammonia.

Based on this research, future work could focus on a variety of CHO cell lines producing different product types, as well as functional work investigating the hypotheses put forward about the role of several proteins in the ammonia stress response, using transfection to knockdown or overexpress specific proteins of interest. Future studies could utilise a gradual addition of ammonia to better reflect ammonia accumulation in cultures. Future work should also study the effects of elevated ammonia on the quality of the recombinant product. Given that glycosylation and folding pathways are affected by elevated ammonia, it is possible that product quality will also be affected, reducing the usable yield from the culture. Previous studies have indicated that stress conditions may cause changes in cell diameter (Kim et al., 2001; Schellenberg et al., 2022); therefore, future studies of ammonia stress should include an investigation of cell size and diameter.

Acknowledgements/funding

This research was funded by a Science Foundation Ireland Frontiers for the Future Award (grant no. 19/FFP/6759). The Orbitrap Fusion Tribrid mass spectrometer was funded by a

Science Foundation Ireland Infrastructure Award to Dublin City University (SFI 16/RI/3701). IgG titre data was generated by Robert Ryan from ValitaCell (now part of Beckman Coulter LS), Dublin, Ireland. Figures 8 and 9 was created with BioRender.com. Venn diagrams created using Venny 2.1 (Oliveros, 2007).

References

Aber, R., Chan, W., Mugisha, S., Jerome-Majewska, L.A., 2019. Transmembrane emp24 domain proteins in development and disease. Genetics Research 101, e14. https://doi.org/10.1017/S0016672319000090

Aebi, M., 2013. N-linked protein glycosylation in the ER. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research 1833, 2430–2437.

Annaert, W.G., Becker, B., Kistner, U., Reth, M., Jahn, R., 1997. Export of Cellubrevin from the Endoplasmic Reticulum Is Controlled by BAP31. J Cell Biol 139, 1397–1410.

Arnold, P.K., Finley, L.W.S., 2022. Regulation and function of the mammalian tricarboxylic acid cycle. J Biol Chem 299, 102838. https://doi.org/10.1016/j.jbc.2022.102838

Bedard, K., Szabo, E., Michalak, M., Opas, M., 2005. Cellular Functions of Endoplasmic Reticulum Chaperones Calreticulin, Calnexin, and ERp57, in: International Review of Cytology, A Survey of Cell Biology. Academic Press, pp. 91–121. https://doi.org/10.1016/S0074-7696(05)45004-4

Bonnon, C., Wendeler, M.W., Paccaud, J.-P., Hauri, H.-P., 2010. Selective export of human GPI-anchored proteins from the endoplasmic reticulum. J Cell Sci 123, 1705–1715. https://doi.org/10.1242/jcs.062950

Budge, J.D., Roobol, J., Singh, G., Mozzanino, T., Knight, T.J., Povey, J., Dean, A., Turner, S.J., Jaques, C.M., Young, R.J., 2021. A proline metabolism selection system and its application to the engineering of lipid biosynthesis in Chinese hamster ovary cells. Metabolic Engineering Communications 13, e00179.

Capiaumont, J., Legrand, C., Carbonell, D., Dousset, B., Belleville, F., Nabet, P., 1995. Methods for reducing the ammonia in hybridoma cell cultures. Journal of biotechnology 39, 49–58.

Chen, P., Harcum, S.W., 2007. Differential display identifies genes in chinese hamster ovary cells sensitive to elevated ammonium. Appl Biochem Biotechnol 141, 349–359. https://doi.org/10.1007/BF02729072

Chen, P., Harcum, S.W., 2006. Effects of elevated ammonium on glycosylation gene expression in CHO cells. Metabolic Engineering 8, 123–132. https://doi.org/10.1016/j.ymben.2005.10.002

Chen, P., Harcum, S.W., 2005. Effects of amino acid additions on ammonium stressed CHO cells. Journal of Biotechnology 117, 277–286. https://doi.org/10.1016/j.jbiotec.2005.02.003

Chen, X., Wang, Y., Xie, X., Chen, H., Zhu, Q., Ge, Z., Wei, H., Deng, J., Xia, Z., Lian, Q., 2018. Heme Oxygenase-1 Reduces Sepsis-Induced Endoplasmic Reticulum Stress and Acute Lung Injury. Mediators Inflamm 2018, 9413876. https://doi.org/10.1155/2018/9413876

Chitwood, D.G., Wang, Q., Elliott, K., Bullock, A., Jordana, D., Li, Z., Wu, C., Harcum, S.W., Saski, C.A., 2021. Characterization of metabolic responses, genetic variations, and microsatellite instability in ammonia-stressed CHO cells grown in fed-batch cultures. BMC Biotechnol 21, 4. https://doi.org/10.1186/s12896-020-00667-2

Chung, M.-I., Lim, M.-H., Lee, Y.-J., Kim, I.-H., Kim, I.-Y., Kim, J.-H., Chang, K.-H., Kim, H.-J., 2003. Reduction of Ammonia Accumulation and Improvement of Cell Viability by Expression of Urea Cycle Enzymes in Chinese Hamster Ovary Cells. Journal of Microbiology and Biotechnology 13, 217–224.

Clarke, C., Doolan, P., Barron, N., Meleady, P., O'Sullivan, F., Gammell, P., Melville, M., Leonard, M., Clynes, M., 2011. Predicting cell-specific productivity from CHO gene expression. Journal of Biotechnology 151, 159–165. https://doi.org/10.1016/j.jbiotec.2010.11.016

Connolly, D.J., O'Neill, L.A., McGettrick, A.F., 2013. The GOLD domain-containing protein TMED1 is involved in interleukin-33 signaling. Journal of Biological Chemistry 288, 5616–5623.

Cour Karottki, K.J., Hefzi, H., Li, S., Pedersen, L.E., Spahn, P.N., Joshi, C., Ruckerbauer, D., Hernandez Bort, J.A., Thomas, A., Lee, J.S., Borth, N., Lee, G.M., Kildegaard, H.F., Lewis, N.E., 2021. A metabolic CRISPR-Cas9 screen in Chinese hamster ovary cells identifies glutamine-sensitive genes. Metab Eng 66, 114–122. https://doi.org/10.1016/j.ymben.2021.03.017

Dahodwala, H., Hodzic, I., Slesarev, A., Cutak, B., Kuzin, A., Lal, R., Liu, J., Mahon, J., Narasimhan, R.L., Onuska, J., Ravellette, J., Reger, K., Sakurada, S., Stewart, F., Borgschulte, T., Cote, C., Lee, K.H., O'Dell, W.B., Kelman, Z., Anderson, B.A., 2025. Development and Characterization of the NISTCHO Reference Cell Line. Biotechnology Journal 20, e70012. https://doi.org/10.1002/biot.70012

Dasarathy, S., Hatzoglou, M., 2018. Hyperammonemia and proteostasis in cirrhosis. Curr Opin Clin Nutr Metab Care 21, 30–36. https://doi.org/10.1097/MCO.0000000000000426

Dejgaard, S.Y., Presley, J.F., 2019. Rab18: new insights into the function of an essential protein. Cell. Mol. Life Sci. 76, 1935–1945. https://doi.org/10.1007/s00018-019-03050-3

Dhitavat, J., Macfarlane, S., Dode, L., Leslie, N., Sakuntabhai, A., MacSween, R., Saihan, E., Hovnanian, A., 2003. Acrokeratosis verruciformis of Hopf is caused by mutation in ATP2A2: evidence that it is allelic to Darier's disease. J Invest Dermatol 120, 229–232. https://doi.org/10.1046/j.1523-1747.2003.t01-1-12045.x

Dudek, J., Benedix, J., Cappel, S., Greiner, M., Jalal, C., Müller, L., Zimmermann, R., 2009. Functions and pathologies of BiP and its interaction partners. Cellular and molecular life sciences 66, 1556–1569.

Elliott, K., Harris, G., Harcum, S., Blakeman, K., Gavin, C., Anderson, J.Y., 2020. Spent media analysis with an integrated CE-MS analyzer of Chinese hamster ovary cells grown in an ammonia-stressed parallel microbioreactor platform. BioProcessing Journal 19.

Fan, Y., Jimenez Del Val, I., Müller, C., Wagtberg Sen, J., Rasmussen, S.K., Kontoravdi, C., Weilguny, D., Andersen, M.R., 2015. Amino acid and glucose metabolism in fed-batch CHO cell culture affects antibody production and glycosylation. Biotechnology and Bioengineering 112, 521–535. https://doi.org/10.1002/bit.25450

Fischer, S., Handrick, R., Otte, K., 2015. The art of CHO cell engineering: A comprehensive retrospect and future perspectives. Biotechnology advances 33, 1878–1896.

Gawlitzek, M., Ryll, T., Lofgren, J., Sliwkowski, M.B., 2000. Ammonium alters N-glycan structures of recombinant TNFR-IgG: Degradative versus biosynthetic mechanisms. Biotechnology and Bioengineering 68, 637–646.

Gerlach, J.Q., Sharma, S., Leister, K.J., Joshi, L., 2012. A tight-knit group: protein glycosylation, endoplasmic reticulum stress and the unfolded protein response. Endoplasmic reticulum stress in health and disease 23–39.

Gozzelino, R., Jeney, V., Soares, M.P., 2010. Mechanisms of cell protection by heme oxygenase-1. Annu Rev Pharmacol Toxicol 50, 323–354. https://doi.org/10.1146/annurev.pharmtox.010909.105600

Ha, T.K., Kim, Y.-G., Lee, G.M., 2015. Understanding of altered N-glycosylation-related gene expression in recombinant Chinese hamster ovary cells subjected to elevated ammonium concentration by digital mRNA counting. Biotechnology and Bioengineering 112, 1583–1593. https://doi.org/10.1002/bit.25568

Halic, M., Gartmann, M., Schlenker, O., Mielke, T., Pool, M.R., Sinning, I., Beckmann, R., 2006. Signal recognition particle receptor exposes the ribosomal translocon binding site. Science 312, 745–747. https://doi.org/10.1126/science.1124864

Hall, J.C., Heel, K., McCauley, R., 1996. Glutamine. BJS (British Journal of Surgery) 83, 305–312. https://doi.org/10.1002/bjs.1800830306

Han, Y.K., Kim, Y.-G., Kim, J.Y., Lee, G.M., 2010. Hyperosmotic stress induces autophagy and apoptosis in recombinant Chinese hamster ovary cell culture. Biotechnology and Bioengineering 105, 1187–1192. https://doi.org/10.1002/bit.22643

Hansen, Henrik A, Emborg, C., 1994. Influence of ammonium on growth, metabolism, and productivity of a continuous suspension Chinese hamster ovary cell culture. Biotechnology progress 10, 121–124.

Hansen, H. A., Emborg, C., 1994. Influence of ammonium on growth, metabolism, and productivity of a continuous suspension Chinese hamster ovary cell culture. Biotechnol Prog 10, 121–124. https://doi.org/10.1021/bp00025a014

Hayter, P.M., Curling, E., Baines, A.J., Jenkins, N., Salmon, I., Strange, P.G., Bull, A.T., 1991. Chinese hamster ovary cell growth and interferon production kinetics in stirred batch culture. Applied microbiology and biotechnology 34, 559–564.

Heikelä, H., Ruohonen, S.T., Adam, M., Viitanen, R., Liljenbäck, H., Eskola, O., Gabriel, M., Mairinoja, L., Pessia, A., Velagapudi, V., Roivainen, A., Zhang, F.-P., Strauss, L., Poutanen, M., 2020. Hydroxysteroid (17β) dehydrogenase 12 is essential for metabolic homeostasis in adult mice. American Journal of Physiology-Endocrinology and Metabolism 319, E494–E508. https://doi.org/10.1152/ajpendo.00042.2020

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4, 44–57. https://doi.org/10.1038/nprot.2008.211

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., Ishiguro-Watanabe, M., 2023. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Research 51, D587–D592.

Kim, J.Y., Kim, Y.-G., Lee, G.M., 2012. CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Applied microbiology and biotechnology 93, 917–930.

Kim, T.K., Chung, J.Y., Sung, Y.H., Lee, G.M., 2001. Relationship between cell size and specific thrombopoietin productivity in Chinese hamster ovary cells during dihydrofolate reductase-mediated gene amplification. Biotechnology and bioprocess engineering 6, 332–336.

Kumar, J., Yu, H., Sheetz, M.P., 1995. Kinectin, an Essential Anchor for Kinesin-Driven Vesicle Motility. Science 267, 1834–1837. https://doi.org/10.1126/science.7892610

Lao, M.-S., Toth, D., 1997. Effects of Ammonium and Lactate on Growth and Metabolism of a Recombinant Chinese Hamster Ovary Cell Culture. Biotechnology Progress 13, 688–691. https://doi.org/10.1021/bp9602360

Li, D., Chen, S., Liu, C., Wei, B., Li, X., 2022. Liver transcriptome analysis reveals biological pathways and transcription factors in response to high ammonia exposure. Inhal Toxicol 34, 219–229. https://doi.org/10.1080/08958378.2022.2083275

Liaci, A.M., Steigenberger, B., Telles de Souza, P.C., Tamara, S., Gröllers-Mulderij, M., Ogrissek, P., Marrink, S.J., Scheltema, R.A., Förster, F., 2021. Structure of the human signal peptidase complex reveals the determinants for signal peptide cleavage. Mol Cell 81, 3934-3948.e11. https://doi.org/10.1016/j.molcel.2021.07.031

Lombardi, D., Soldati, T., Riederer, M.A., Goda, Y., Zerial, M., Pfeffer, S.R., 1993. Rab9 functions in transport between late endosomes and the trans Golgi network. EMBO J 12, 677–682. https://doi.org/10.1002/j.1460-2075.1993.tb05701.x

Mastrangelo, A.J., Hardwick, J.M., Zou, S., Betenbaugh, M.J., 2000. Part II. Overexpression of bcl-2 family members enhances survival of mammalian cells in response to various culture

insults. Biotechnology and Bioengineering 67, 555–564. https://doi.org/10.1002/(SICI)1097-0290(20000305)67:5%253C555::AID-BIT6%253E3.0.CO;2-T

Melville, D.B., Studer, S., Schekman, R., 2020. Small sequence variations between two mammalian paralogs of the small GTPase SAR1 underlie functional differences in coat protein complex II assembly. J Biol Chem 295, 8401–8412. https://doi.org/10.1074/jbc.RA120.012964

Michalak, M., Robert Parker, J.M., Opas, M., 2002. Ca2+ signaling and calcium binding chaperones of the endoplasmic reticulum. Cell Calcium, Endoplasmic Reticulum as a Signalling Organelle 32, 269–278. https://doi.org/10.1016/S0143416002001884

Mitchell, S., Ellingson, C., Coyne, T., Hall, L., Neill, M., Christian, N., Higham, C., Dobrowolski, S.F., Tuchman, M., Summar, M., Urea Cycle Disorder Consortium, 2009. Genetic variation in the urea cycle: a model resource for investigating key candidate genes for common diseases. Hum Mutat 30, 56–60. https://doi.org/10.1002/humu.20813

NISTCHO, 2016. . NIST.

Ohkuni, A., Ohno, Y., Kihara, A., 2013. Identification of acyl-CoA synthetases involved in the mammalian sphingosine 1-phosphate metabolic pathway. Biochem Biophys Res Commun 442, 195–201. https://doi.org/10.1016/j.bbrc.2013.11.036

Pereira, S., Kildegaard, H.F., Andersen, M.R., 2018. Impact of CHO Metabolism on Cell Growth and Protein Production: An Overview of Toxic and Inhibiting Metabolites and Nutrients. Biotechnology Journal 13, 1700499. https://doi.org/10.1002/biot.201700499

Prashad, K., Mehra, S., 2015. Dynamics of unfolded protein response in recombinant CHO cells. Cytotechnology 67, 237–254. https://doi.org/10.1007/s10616-013-9678-8

Rezvani, K., 2016. UBXD Proteins: A Family of Proteins with Diverse Functions in Cancer. International Journal of Molecular Sciences 17, 1724. https://doi.org/10.3390/ijms17101724

Roy, G., Zhang, S., Li, L., Higham, E., Wu, H., Marelli, M., Bowen, M.A., 2017. Development of a fluorescent reporter system for monitoring ER stress in Chinese hamster ovary cells and its application for therapeutic protein production. PLoS One 12, e0183694. https://doi.org/10.1371/journal.pone.0183694

Rutkowski, D.T., Kaufman, R.J., 2004. A trip to the ER: coping with stress. Trends in Cell Biology 14, 20–28. https://doi.org/10.1016/j.tcb.2003.11.001

Ryan, D., Sideri, C.-K., Henry, M., Efeoglu, E., Meleady, P., 2023. Label-Free Quantitative Proteomics Analysis of Producer and Non-producer Chinese Hamsters Ovary (CHO) Cells under ER Stress Conditions. Current Research in Biotechnology 100141.

Ryan, D., Sideri, C.-K., Henry, M., Karuppuchamy, S., Efeoglu, E., Meleady, P., n.d. Proteomic and Ubiquitinated Proteome Insights Into ER Stress Responses in Chinese Hamster Ovary

Cells Under Mild Hypothermic Conditions. Biotechnology and Bioengineering n/a. https://doi.org/10.1002/bit.70081

Sakabe, I., Hu, R., Jin, L., Clarke, R., Kasid, U.N., 2015. TMEM33: a new stress-inducible endoplasmic reticulum transmembrane protein and modulator of the unfolded protein response signaling. Breast cancer research and treatment 153, 285–297.

Samy, A., Yamano-Adachi, N., Koga, Y., Omasa, T., 2021. Secretion of a low-molecular-weight species of endogenous GRP94 devoid of the KDEL motif during endoplasmic reticulum stress in Chinese hamster ovary cells. Traffic 22, 425–438. https://doi.org/10.1111/tra.12818

Sanders, P.G., Wilson, R.H., 1984. Amplification and cloning of the Chinese hamster glutamine synthetase gene. The EMBO journal 3, 65–71.

Savizi, I.S.P., Maghsoudi, N., Motamedian, E., Lewis, N.E., Shojaosadati, S.A., 2022. Valine feeding reduces ammonia production through rearrangement of metabolic fluxes in central carbon metabolism of CHO cells. Appl Microbiol Biotechnol 106, 1113–1126. https://doi.org/10.1007/s00253-021-11755-4

Schellenberg, J., Nagraik, T., Wohlenberg, O.J., Ruhl, S., Bahnemann, J., Scheper, T., Solle, D., 2022. Stress-induced increase of monoclonal antibody production in CHO cells. Engineering in Life Sciences 22, 427–436.

Selvaprakash, K., Sideri, C.-K., Henry, M., Efeoglu, E., Ryan, D., Meleady, P., 2024. Characterization of the Ubiquitin-Modified Proteome of Recombinant Chinese Hamster Ovary Cells in Response to Endoplasmic Reticulum Stress. Biotechnology Journal 19, e202400413. https://doi.org/10.1002/biot.202400413

Sherman, B.T., Hao, M., Qiu, J., Jiao, X., Baseler, M.W., Lane, H.C., Imamichi, T., Chang, W., 2022. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic Acids Res 50, W216–W221. https://doi.org/10.1093/nar/gkac194

Soupene, E., Kuypers, F.A., 2008. Mammalian Long-Chain Acyl-CoA Synthetases. Exp Biol Med (Maywood) 233, 507–521. https://doi.org/10.3181/0710-MR-287

Srila, W., Baumann, M., Riedl, M., Rangnoi, K., Borth, N., Yamabhai, M., 2023. Glutamine synthetase (GS) knockout (KO) using CRISPR/Cpf1 diversely enhances selection efficiency of CHO cells expressing therapeutic antibodies. Sci Rep 13, 10473. https://doi.org/10.1038/s41598-023-37288-6

Synoground, B.F., McGraw, C.E., Elliott, K.S., Leuze, C., Roth, J.R., Harcum, S.W., Sandoval, N.R., 2021. Transient ammonia stress on Chinese hamster ovary (CHO) cells yield alterations to alanine metabolism and IgG glycosylation profiles. Biotechnology Journal 16, 2100098. https://doi.org/10.1002/biot.202100098

Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., Jensen, L.J., von Mering, C., 2021. The STRING database

in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 49, D605–D612. https://doi.org/10.1093/nar/gkaa1074

Tang, B.L., Zhang, T., Low, D.Y., Wong, E.T., Horstmann, H., Hong, W., 2000. Mammalian homologues of yeast sec31p. An ubiquitously expressed form is localized to endoplasmic reticulum (ER) exit sites and is essential for ER-Golgi transport. J Biol Chem 275, 13597–13604. https://doi.org/10.1074/jbc.275.18.13597

Tani, K., Oyama, Y., Hatsuzawa, K., Tagaya, M., 1999. Hypothetical protein KIAA0079 is a mammalian homologue of yeast Sec24p. FEBS Lett 447, 247–250. https://doi.org/10.1016/s0014-5793(99)00303-8

Turnbull, J., Tiberia, E., Pereira, S., Zhao, X., Pencea, N., Wheeler, A.L., Yu, W.Q., Ivovic, A., Naranian, T., Israelian, N., Draginov, A., Piliguian, M., Frankland, P.W., Wang, P., Ackerley, C.A., Giacca, A., Minassian, B.A., 2013. Deficiency of a Glycogen Synthase-associated Protein, Epm2aip1, Causes Decreased Glycogen Synthesis and Hepatic Insulin Resistance*. Journal of Biological Chemistry 288, 34627–34637. https://doi.org/10.1074/jbc.M113.483198

Vazquez-Martinez, R., Malagón, M.M., 2011. Rab Proteins and the Secretory Pathway: The Case of Rab18 in Neuroendocrine Cells. Front. Endocrinol. 2. https://doi.org/10.3389/fendo.2011.00001

Vorum, H., Hager, H., Christensen, B.M., Nielsen, S., Honoré, B., 1999. Human calumenin localizes to the secretory pathway and is secreted to the medium. Exp Cell Res 248, 473–481. https://doi.org/10.1006/excr.1999.4431

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer International Publishing.

Xu, S., Gavin, J., Jiang, R., Chen, H., 2017. Bioreactor productivity and media cost comparison for different intensified cell culture processes. Biotechnology progress 33, 867–878.

Yang, M., Butler, M., 2000a. Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation. Biotechnology and bioengineering 68, 370–380.

Yang, M., Butler, M., 2000b. Effect of ammonia on the glycosylation of human recombinant erythropoietin in culture. Biotechnology Progress 16, 751–759. https://doi.org/10.1021/bp000090b

Zhang, J., Li, C., Tang, X., Lu, Q., Sa, R., Zhang, H., 2015. High Concentrations of Atmospheric Ammonia Induce Alterations in the Hepatic Proteome of Broilers (Gallus gallus): An iTRAQ-Based Quantitative Proteomic Analysis. PLoS One 10, e0123596. https://doi.org/10.1371/journal.pone.0123596

Zhang, Z., Luo, S., Barbosa, G.O., Bai, M., Kornberg, T.B., Ma, D.K., 2021. The conserved transmembrane protein TMEM-39 coordinates with COPII to promote collagen secretion and regulate ER stress response. PLoS genetics 17, e1009317.

Zhu, Z.-C., Liu, J.-W., Li, K., Zheng, J., Xiong, Z.-Q., 2018. KPNB1 inhibition disrupts proteostasis and triggers unfolded protein response-mediated apoptosis in glioblastoma cells. Oncogene 37, 2936–2952. https://doi.org/10.1038/s41388-018-0180-9

Zuo, L., Zou, X., Ge, J., Hu, S., Fang, Y., Xu, Y., Chen, R., Xu, S., Yu, G., Zhou, X., Ji, L., 2025. The Nrf2-HMOX1 pathway as a therapeutic target for reversing cisplatin resistance in non-small cell lung cancer via inhibiting ferroptosis. Cell Death Discov. 11, 287. https://doi.org/10.1038/s41420-025-02564-z

Chapter 5: Regression and machine learning methods for the analysis of CHO cell mass spectrometry ID (mzID) files and identification of bioprocess relevant stress markers.

5.1. Introduction

As mass spectrometry becomes faster and more sensitive, the amount of information gained during a proteomic study increases beyond what can realistically be analysed by a single researcher. It is therefore important to utilise bioinformatic methods to allow for in-depth exploration of large datasets. In this chapter, the statistical software, R, is used in the analysis of mass spectrometry identify files from CHO cells exposed to mild hypothermia (Chapter 3). Using logistic regression, the correlation between specific proteins and peptides and the stress condition (31°C culture) was explored, with proteins involved in cell cycle regulation and cell structure linked to the lower temperature. Basic machine learning methods were explored to find which is best at handling MZID data for stress prediction. While limited by computational capacity, it was found that Random Forest was better at recognising stress samples than Support Vector Machines or Extreme Gradient Boost. Extreme Gradient Boost was found to be the fastest and least computationally intense, so while less accurate, it may be more accessible to researchers. These methods uncovered a set of anti-apoptosis factors that the algorithms linked with the stress condition, giving insight into how CHO cells adapt to stress conditions. These results supply several new potential genetic engineering targets to enhance cells stress resistance. In the proteomics studies detailed in the earlier chapter, the aim of understanding the ER stress mechanisms required a focused approach which meant an abundance of generated data without detailed analysis.

5.1.1 The growth of proteomics datasets

Following the development of MALDI and electrospray ionisation in the late 1980s, it was recognised that by cleaving a protein with a known enzyme, the sequence of the fragmented peptides could be determined by comparing the mass of the fragment to the known masses of amino acids. Combined with the increasing size of databases stemming from the

advancement of genomic methods (Mann et al., 1993), peptide fragment masses determined using MS could create an amino acid sequence that could be searched in a database, resulting in the identification of a protein without needing to *de novo* characterise the whole sequence.

In the past 20 years, the amount of data that is generated from a proteomic study has increased dramatically. The studies detailed in this thesis could result in approximately 5000 proteins, while newer methods are resulting in 8000-10000 identifications. Computational tools are essential for interpreting the large volumes of data generated and for controlling false discovery rates

5.1.2 The MZIdentML file format

This wealth of data represents several underexplored avenues for investigation. MS identity files, abbreviated to MZIdentML or MZID files are created from comparison of raw MS data to a protein database through a SEQUEST algorithm, and standardises how search engine outputs are stored and exchanged. This file format was launched in 2012, developed by the Proteomics Standards Initiative in collaboration with MS instrument and software vendors, and proteomics open-source projects (Jones et al., 2012).The structure includes detailed elements describing spectra, search parameters, peptides, modifications, protein groups, and scores/statistics for PSMs. The full structure of the file is explored in **Figure 1**, taken from the 2012 paper describing the file type. The structure includes peptide-spectrum matches (PSMs), which record the association between observed MS/MS spectra and the peptide sequences they most likely represent; protein inference information, which aggregates peptide identifications to determine the presence of proteins in the sample; scoring metrics from search engines that quantify the confidence of each identification; modifications present, if any, and metadata detailing experimental parameters, instrument settings and search conditions. Access to this information alongside the identifications allows for reproducibility, downstream statistical analysis, and integration of these results across research groups. Although this data format is utilised for the sharing of proteomic data

through open-source consortiums including ProteomeXchange (Perez-Riverol et al., 2025), PeptideAtlas (Desiere et al., 2006) and GPMDB (Craig et al., 2004), it is predominantly used as solely a repository for the relevant information and rarely used as a format to be analysed.

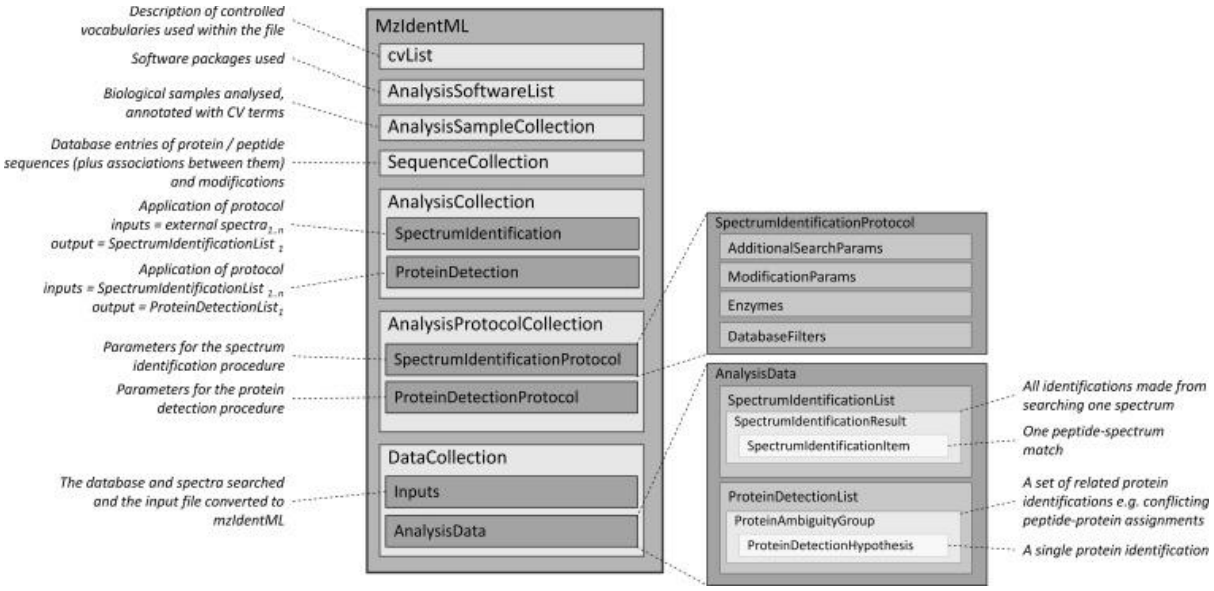


Figure 1: The overall structure of a typical mzIdentML file. Each file must contain one or more instances of SpectrumIdentificationList (the set of peptide identifications made by a search) and must contain zero or one ProteinDetectionList (the set of proteins identities inferred from peptide identifications). *Figure from "The mzIdentML Data Standard for Mass Spectrometry-Based Proteomics Results" by Andrew R Jones et al., Molecular and Cellular Proteomics, 2012.*

The data contained within an MZID file provides an opportunity for us to uncover what signatures of stress may be found within the broader data of proteomics than just the final protein expression levels. By taking a subset of the MZID, we can include MS data such as retention time, along with protein accession, modifications and peptide length to investigate any connection between these parameters and stress conditions. While the amount of data contained in these datasets may be incredibly valuable, the sheer size of these datasets means that true insight can only be gained through the use of computational methods.

5.1.3 R as a tool for protein bioinformatic analysis

R, www.r-project.org, is an open-source system for statistical computation and graphic modelling, first released in 1992, which allows for researchers to freely use a wide range of techniques, and is highly extensible (Kurt Hornik and the R Core Team, 2024). Through its dedicated bioinformatics platform, Bioconductor (Gentleman et al., 2004), R serves as an invaluable tool for life science researchers, with currently 2341 packages available on the platform's most recent update, Bioconductor 3.21 (Turner, 2024). Amongst programming languages, R's syntax is amongst the most intuitive, making it accessible for most researchers with little training needed to perform basic analysis. With greater familiarity, the possibilities for analysis in R are expanded, with packages available for dealing with almost any data type.

Several R-based packages support the efficient import, parsing, and preprocessing of MS files, enabling seamless integration of proteomics identification results into downstream statistical workflows. The mzR package can read and access raw mass spectrometry data and facilitate cross-format compatibility (Chambers et al., 2012). MSnbase builds on these capabilities by providing a high-level infrastructure for storing, manipulating, and processing proteomics data, supporting tasks such as filtering identifications by score thresholds, merging datasets, and mapping peptides to proteins (Gatto and Lilley, 2012). The mzID package provides a dedicated parser for mzIdentML, extracting peptide-spectrum matches, protein inference results, and associated metadata into accessible R objects (Thomas Pedersen et al., 2025). These tools enable reproducible, scriptable data processing pipelines, ensuring consistency across large-scale proteomics studies.

R provides a robust platform for proteomics data analysis due to its combination of statistical power, reproducibility, and extensibility. R allows researchers to explore a variety of statistical methodologies, from classical regression to advanced machine learning, and facilitates reproducibility through scripted workflows and integration with version control

systems. R's compatibility with machine learning frameworks such as caret, mlr3, and tidymodels allows for the application of classification, regression, and feature selection algorithms to proteomics datasets. The comprehensive visualisation capabilities of packages like ggplot2, enable the creation of high-quality, accessible plots for exploratory data analysis and result presentation (Wickham, 2016). This combination of statistical depth, machine learning support, visualisation power, and reproducibility makes R a particularly suitable choice for processing and analysing mzIdentML data in proteomics research.

5.1.4 Regression

Regression analysis is a statistical technique used to model the relationship between one or more explanatory variables and a dependent variable and is a fundamental tool in quantitative proteomics for identifying trends and associations within large datasets (Kutner, 2005). Linear regression models the dependent variable y as a linear combination of an independent variable x and an error term ε expressed mathematically as $y=\beta_0+\beta_1 x+\varepsilon$ where β_0 is the intercept, β_1 is the slope, and ε accounts for residual variation (Freedman, 2009). Multiple linear regression generalises this model to incorporate multiple predictors, $y=\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_n x_n+\varepsilon$, enabling the simultaneous assessment of several variables that may exist in a dataset (Kutner, 2005).

In a model which contains a binary classification, i.e. Stressed/Unstressed, linear regression will overfit the line. The assumptions of linear regression, such as normally distributed residuals and that the variance of residuals is constant across all independent variables (homoscedasticity), do not hold true when the dependent variable is binary, leading to biased standard errors and unreliable inference (Kutner, 2005). In high-dimensional datasets, such as those derived from mzID files, linear regression also risks overfitting: with many predictors relative to the number of samples, the model may capture noise rather than the true signal, reducing generalisability to new data. Logistic regression, which is based on a linear predictor, applies the logistic (sigmoid) function to model the probability of a binary

outcome making it suitable for classification problems such as distinguishing experimental conditions (Hosmer et al., 2013).

$P(y = 1) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}}$ where $P(y = 1)$ is the probability that the dependent variable y equals 1, and the exponential term e^{-z} ensures the output is transformed by the sigmoid function, which constrains predicted probabilities to the range (0, 1), with z being the linear predictor. One of the strengths of these regression methods is their interpretability: estimated coefficients directly quantify the direction and magnitude of the association between predictors and the outcome, allowing for clear biological interpretation.

5.1.5 Machine Learning

A current hot topic, in and outside the life sciences, Artificial Intelligence (AI) has the potential to fundamentally change how we look at biological datasets (Jones et al., 2018). Machine learning (ML) refers to a field within AI that focuses on developing algorithms that best represent a set of data, and the relationship between input data and output, using novel or different combinations of features and weights (James et al., 2013). There are four different learning methods that can be applied in ML, each with different strengths depending on the dataset and learning objective: supervised, unsupervised, semi supervised, and reinforcement learning (Hastie et al., 2005). The uses of these different types of ML are detailed in **Figure 2**.

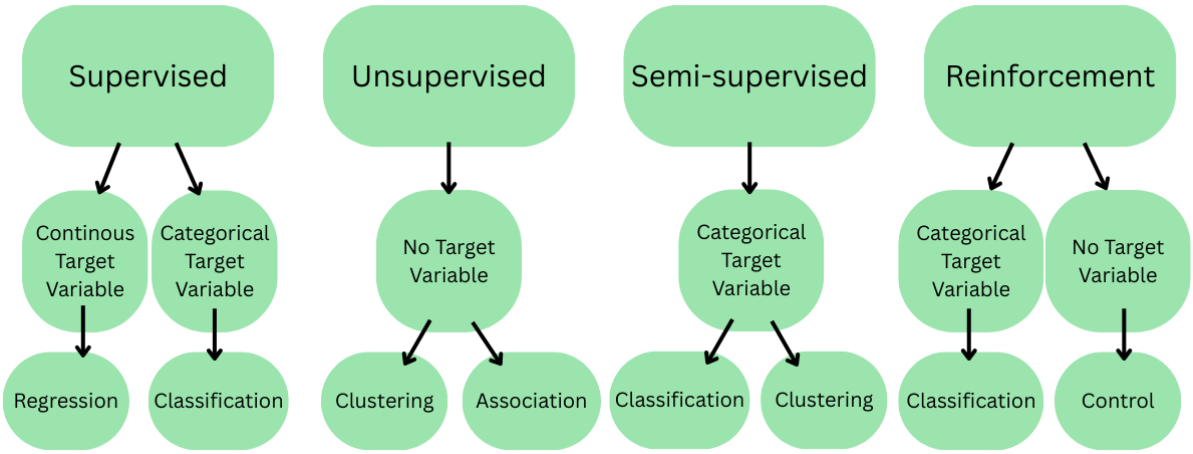


Figure 2: Types of machine learning and the tasks they can be used for. ML encompasses four types- supervised, unsupervised, semi-supervised and reinforcement. Supervised learning used models trained with labelled data for classification and regression. Unsupervised learning finds patterns in unlabelled data to use for clustering. Semi-supervised learning utilises both labelled and unlabelled data. Reinforcement learning improves and refines model performance through interaction with the data environment to maximise a reward signal.

The "target" is the feature that is to be predicted (stressed/unstressed). Datasets used for ML are generally split into training, validation, and testing datasets. Supervised learning uses patterns in the training dataset to map features onto a target variable, allowing the algorithm to make predictions on unseen datasets based on the perceived relationship between predictors and outcomes (Ohri and Kumar, 2021). This method is termed supervised because the algorithm is provided with feature-target pairs during training and is "guided" by the known target values to evaluate its predictions and adjust parameters accordingly. This requires the labelling of data, introducing additional work in dataset preparation and increases computational demands, particularly when large-scale datasets are involved (Alloghani et al., 2020). Despite this overhead, supervised learning remains one of the most widely applied models due to its ability to provide robust predictive performance across a range of problems, including classification of disease states, regression-based prediction of protein abundance, and feature selection for biomarker discovery (Libbrecht and Noble, 2015). Unsupervised learning uses unlabelled data to build its algorithms, usually requiring

larger datasets than supervised methods, aiming to find patterns in the dataset and evaluate data points according to those patterns (Naeem et al., 2023). Patterns within the data are detected based on several methods; clustering, association, anomaly detection and autoencoders, with clustering and association being the most common methods (Gao et al., 2020). Clustering in unsupervised learning groups data points based on similarity, aiming to maximise within-cluster similarity while minimising between-cluster similarity (Rodriguez et al., 2019). Algorithms such as k-means, hierarchical clustering, or density-based methods use distance metrics (e.g., Euclidean, cosine) to assign data points to clusters (Akram et al., 2022). This enables the discovery of hidden structures or natural groupings in complex datasets. Association in unsupervised learning identifies relationships between variables by finding rules that capture how items or features co-occur within datasets. Algorithms such as the Apriori or FP-growth methods generate rules in the form *if X then Y*, quantified by support, confidence, and lift measures (Zhang et al., 2018). Anomaly detection identifies observations that deviate from expected patterns, which may represent errors, rare events, or biologically meaningful variation (Chandola et al., 2009). Autoencoders, a neural network approach, works similarly, detecting anomalies by compressing data and reconstructing it, with unusually high reconstruction error signalling an anomaly, making them particularly effective for high-dimensional datasets (Hinton and Salakhutdinov, 2006). Semi-supervised learning uses both labelled and unlabelled data, training some elements of the system using a small manually labelled dataset, and based on those learnings, the system use a larger unlabelled dataset to build predictor elements. A dataset of labelled and pseudo-labelled data is then used to train the network (Van Engelen and Hoos, 2020). In reinforcement learning an algorithm learns from trial-and-error feedback, and is predominantly used for determining the best possible solution in a situation (Ladosz et al., 2022).

Supervised machine learning has been applied to proteomic data to for classification, prediction, and feature selection tasks. A prominent use is disease classification, where

algorithms such as support vector machines and random forests are trained on proteomic signatures to distinguish between healthy and diseased states, aiding in biomarker discovery for conditions such as colorectal cancer and leukaemia (Li et al., 2001; Wang et al., 2024; Yin et al., 2024). Recently, ML methods were used on non-invasive urinary proteomic samples, and proteome biomarkers were found to be the second most powerful predictor of cancer status in this data cohort, after age of patient (Wang et al., 2024). Supervised approaches are also used for protein abundance prediction and quantification, where regression-based models map spectral features to protein concentrations, enhancing the accuracy of label-free quantification methods (Neely et al., 2023). Additionally, supervised feature selection methods enable the identification of the most informative peptides or proteins, thereby reducing dimensionality and improving interpretability in large-scale datasets (Webb-Robertson et al., 2011). These applications highlight the central role of supervised learning in extracting biologically meaningful insights from complex proteomic data.

5.1.6 ML methods available in R

R provides a wide range of ML packages, with most having an easy-to-use basic function, but with far more detailed settings available for more thorough analysis.

Random Forest

Random Forest (RF) is a non-parametric machine learning approach that can accommodate different response variable types such as categorical or quantitative outcomes. It can be applied to data with predictors of various scales or distributions and can perform well in high-dimensional settings where the number of predictors may be greater than the number of observations (Breiman, 2001). RF functions by creating a cluster of decision trees, built using a random subset of the sample data, and merging these decision trees together to achieve a better prediction accuracy. Decision tree modelling predicts outcomes by

recursively splitting data into branches based on feature values, forming a tree-like structure. Each internal node represents a decision rule on a feature, branches represent outcomes of that rule, and leaf nodes correspond to final predictions (Xu et al., 2021). Because of the binary splitting of features, singular decision trees can be prone to overfitting, which can be mitigated using an ensemble model such as RF, where each decision node is validated against the same decision in other trees (Oshiro et al., 2012). In 2024, RF was used to analyse the proteomic data derived from tumour (deep and superficial) and non-tumour tissue samples from 16 patients with colorectal cancer. From the analysis of 2009 proteins, 91 proteins showed significant quantitative changes, including 23 proteins that could be novel biomarkers for this cancer. In addition, RF was able to classify the three tissue types with 98.4% accuracy using a set of 21 proteins (Contini et al., 2024). In another study, using proteins from plasma, serum, or urine-derived exosomes, a RF model was more than 90% accurate in distinguishing cancer exosomes from other exosomes, therefore allowing greater detail in classifying cancer subtypes (Li et al., 2024).

Extreme Gradient Boosting

Based on the Gradient Boosting Decision Tree algorithm, which uses a negative gradient calculation to improve model efficiency (Friedman, 2001), Extreme Gradient Boosting (XGBoost) expands the objective function and adds further regularisation to leaf weights (Chen and Guestrin, 2016). In ML, boosting refers to the training of several "weak" classifiers to better inform a "strong" model, reducing bias in a dataset, and better understanding probability distribution in a model (Kearns et al., 1994). XGBoost multiprotein risk models have been shown to out-perform classic regression models and clinical scores in the prediction of all-cause mortality in patients with increased cardiovascular risk (Matthias Unterhuber et al., 2021). XGBoost has also been shown to be

more effective at modelling protein-protein interactions than other ML methods, including RF and SVM (Beltran et al., 2019).

Support Vector Machines

Support Vector Machines (SVM) work by finding the optimal hyperplane that separates data points of different classes with the maximum margin (Suthaharan, 2016). A hyperplane is a decision boundary that divides the feature space, existing as a line in two dimensions, a plane in three dimensions, and a generalisation of both in any higher dimensions. The goal is to choose the hyperplane that not only separates the classes but also maximises the distance between the boundary and the nearest data points from each class (Pisner and Schnyer, 2020). These nearest points are called support vectors, and they define the placement of the hyperplane. SVM can also handle non-linear separation by applying kernel functions, which map data into higher-dimensional spaces where a suitable linear hyperplane can be found, aiming to balance minimising classification error with maximising margin robustness (Noble, 2006). However, SVM is a mathematically complex and computationally demanding method. SVM modelling has been used for the quantitative prediction of proteotypic peptides, using a simple descriptor space based on 35 properties of amino acid content, charge, hydrophilicity and polarity, with an average accuracy measure of ~0.83, which the authors believe could reduce computational time in searching MS results against only peptides predicted to be detectable through MS rather than the entire annotated database of a species (Webb-Robertson et al., 2010). SVM models used to distinguish tissue sample proteome data from interstitial lung disease and idiopathic pulmonary fibrosis were found to be between 77.3-83.1% accurate at distinguishing the tissue types (Huang et al., 2024).

The complexity of mzIdentML data lends itself to be a treasure trove of information, but this complexity necessitates the use of regression models, and machine learning approaches to analyse. R provides a flexible and reproducible platform for parsing, preprocessing, and

analysing mzIdentML files, while regression offers interpretable baseline models for quantifying relationships between experimental factors and proteomic outputs. Machine learning methods, including decision trees, random forests, support vector machines, and gradient boosting, extend these capabilities by capturing non-linear patterns and enhancing predictive performance. Together, these approaches enable more accurate and scalable analysis of proteomics datasets, which is particularly relevant for CHO-based bioprocessing, where understanding protein expression patterns directly informs process optimisation, product quality, and cell line engineering. This relationship between statistical modelling and machine learning sets the foundation for robust, data-driven strategies in proteomics and biomanufacturing.

5.2 Methodology

5.2.1 MzIdentML file creation

MZID files were generated from Thermo Fisher '.RAW' MS files, using Proteome Discoverer 2.4 (ThermoScientific) against a *Cricetulus griseus* UniProt database (downloaded September 2022). These files are those arising from the ubiquitin enriched CHO-K1, CHO-DP12 and CHO-SK15-EPO samples detailed in Chapter 3, shown in **Table 1**. The following search parameters were set for protein identifications: (i) MS/MS mass tolerance set at 0.6 Da; (ii) peptide mass tolerance set to 20 ppm; (iii) trypsin as the enzyme; (iv) up to two missed cleavages were allowed; (v) cysteine carbamidomethylation as a static modification; (vi) methionine oxidation as a dynamic modification; and (vii) lysine with a K-ε-GG remnant (114.0429 Da) as a variable modification, but non-modified peptides were not excluded in this case.

Table 1: The samples used for statistical analysis, using the associated MZID files from ubiquitin enriched samples. The last row signifies where a sample was identified as stressed (Y) or unstressed (N).

CHO-K1				CHO-SK15-EPO				CHO-DP12			
Day 4		Day 7		Day 4		Day 7		Day 4		Day 7	
31°C_A	37°C_A	31°C_A	37°C_A	31°C_A	37°C_A	31°C_A	37°C_A	31°C_A	37°C_A	31°C_A	37°C_A
31°C_B	37°C_B	31°C_B	37°C_B	31°C_B	37°C_B	31°C_B	37°C_B	31°C_B	37°C_B	31°C_B	37°C_B
31°C_C	37°C_C	31°C_C	37°C_C	31°C_C	37°C_C	31°C_C	37°C_C	31°C_C	37°C_C	31°C_C	37°C_C
Y	N	Y	N	Y	N	Y	N	Y	N	Y	N

5.2.2 Parsing of files into R

All R code used can be found in Appendix V.

Using the packages mzID, V1.38.0. MzID and MSnbase, V 2.26.0, files were read by R (Gatto and Lilley, 2012; Thomas Pedersen et al., 2025). The mzID package "flatten" function was used to convert these files into a R data frame, and "grepl" was used to add the stressed column to these data frames, with Y signifying rows stressed samples, and N signifying

unstressed samples. In this dataset, the stressed label is applied to cells cultured at 31°C, while unstressed cells were cultures at 37°C. lapply was then used to rbind each of these data frames together into one large composite data frame, and non-informative columns dropped (sample name, run ID etc).

5.2.3 Logistic regression

Binomial logistic regression was carried out using the glm command from the in-built R package Stats V4.3.1, regressing the stress column onto all other data. Further regression was carried out using the package glmnet V4.1-10 (Friedman et al., 2021), which allowed for punitive algorithms to combat high cardinality.

5.2.4 Random forest

The R package Ranger V0.17.0 (Wright and Ziegler, 2017) was used for random forest modelling, due to its better handling of high dimension datasets than the RandomForest package. Probability was set to TRUE and importance factor used was "impurity".

5.2.5 Extreme gradient boost

Modelling using extreme gradient boost algorithms was performed using package XGBoost V.7.11.1 (Chen and Guestrin, 2016). Parameters were set to objective="binary:logistic", eval_metric="auc", max_depth=6, eta=0.1, with nrounds=200 rounds.

5.2.6 Support vector machine

SVM modelling was performed using package e1071 V1.7-16 (Meyer et al., 2024), with the following parameters kernel="radial", probability=TRUE and scale=TRUE.

5.3 Results

This statistical analysis was performed on a database of ubiquitinated and non-ubiquitinated peptides, from three CHO cell lines (CHO-K1, CHO-DP12 and CHO-SK15-EPO) which were cultured at 31°C and 37°C and subjected to MS analysis on day 4 and day 7 of culture. In total there was 301,391 rows of data, with 5015 different accessions and 20701 different peptide sequences. Due to limited computational capacity, only a subset of the data can be used to allow for analysis. To allow for greater exploration of this dataset within these constraints, each method was repeated three times using different seed sets. This limitation could be overcome with greater computing power, following the methods described in this chapter.

Logistic regression

When performed on 5% of the whole dataset, the regression of the MZID data onto the stress column using a general logistic model resulted in no accession numbers or peptides with a statistically significant p-value. This is likely due to the high cardinality of accession and peptide sequence columns, which contain a high number of unique values. High cardinality attributes can be very informative, however the vast array of values increases the dimensionality of data to a point where it becomes difficult or even impossible for most classification techniques to build a model (Moeyersoms and Martens, 2015). Two solutions were employed in tandem to work around this problem in regression. Firstly, peptide sequences and accession numbers were given scores based on their occurrence within stressed or unstressed data, and then encoded with unique values, so the actual accession and sequence could be removed, but the influence still included. This allows for a reduction of data size, and models can more easily deal with numerical values than character strings. Secondly, a penalised maximum likelihood algorithm was applied to the logistic model. This estimation method modifies the usual maximum likelihood function by adding a penalty term to shrink or constrain coefficient values. This method reduces overfitting and allows

handling of high-dimensional data such as in ridge or lasso regression, which can improve prediction (Machado et al., 2021). The three models were tested with an alpha value of 0.5, employing an elastic net regression, using a mix of ridge and lasso to balance shrinkage and feature selection. Between the 3 replicates, there were no shared accessions, as shown in **Figure 3**. The lack of overlapping results is likely due to the small data subsets used for this method, given its computational intensity.

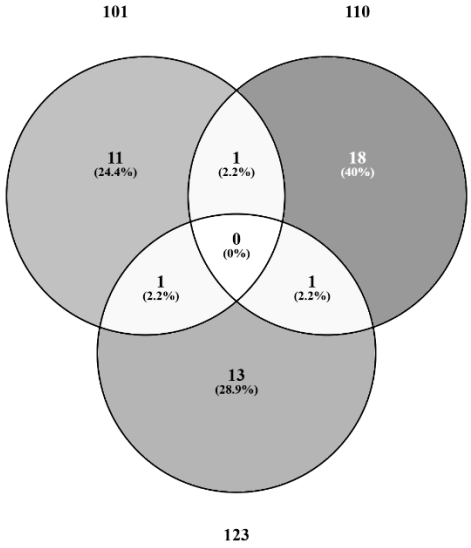


Figure 3: The shared accessions between the three Logistic Regression replicates.

Random Forest

Random Forest modelling was performed using the Ranger package in R, which is designed to better handle large data sets than the basic Random Forest package. Modelling was performed on two dataset sizes, one utilising 25% of the full data and 50% of the full data. Set for a classification task, a forest of 500 trees was used, meaning 500 different decision trees were explored. The resulting models were then tested against a validation data set of the same size as the training set. The accuracy of each model on both dataset sizes is included in **Table 2**. The protein accessions and peptide sequences that were important to stress prediction were determined by mean prediction greater than 0.9 and n>2 where n is the

number of occurrences in the dataset (**Table 3**), with 11 accessions and 16 peptide sequences common to all three replicates (**Figure 4**). These shared accessions and peptides are listed in **Table 4**.

Table 2: Accuracy of stress prediction of the triplicate RF models, using 25% of the full data and 50% of the full data.

Seed	**Accuracy** ***25% of Data***	**Accuracy** ***50% of Data***
101	0.6335029	0.7227047
110	0.6314785	0.7216569
123	0.6356213	0.7241823

Table 3: The number of accession numbers and peptide sequences found to be impactful in stress prediction, where mean prediction is greater than 0.9 and n≥2, where n is the number of occurrences in the dataset, on 50% of the full data.

Seed	**Accession**	**Peptide**
101	38	136
110	33	133
123	35	133

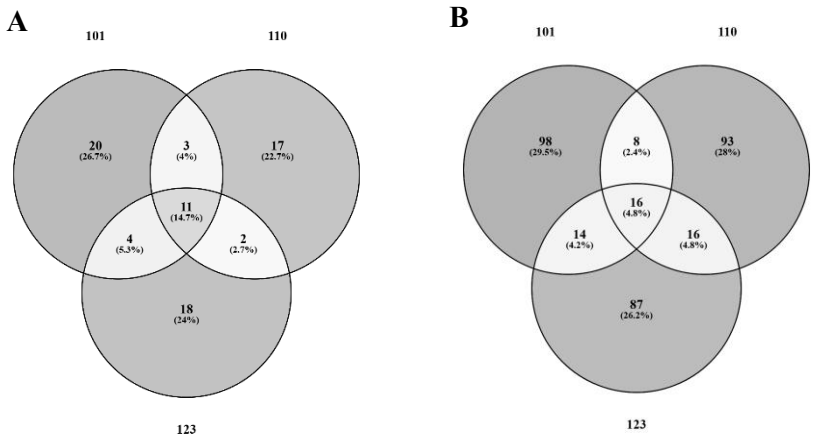


Figure 4: The shared accessions (A) and peptide sequences (B) between the three RF datasets.

Table 4: The accession numbers and peptide sequences common to all three RF seeds, found to be impactful in stress prediction, where mean prediction is greater than 0.9 and n≥2, where n is the number of occurrences in the dataset.

G3H3F8	TK1		TFSFAIPLIEK	RNA helicase
G3I180	TRAF NF-kappa-B activator		TVIVAALDGTFQR	TK1
G3H215	40S ribosomal protein SA		KINELTGIK	26S proteasome regulatory subunit 7
G3HZE5	HSP70kDA1		EQNPYEKF	PTTG1IP
G3HG79	IQGAP3		QQLLEDLK	Myomegalin
G3H8Z6	PTTG1IP		IANVSILK	VPS33B
G3HXF7	Glyceraldehyde-3-phosphate dehydrogenase		VVLDVGSGTGILCMFAAK	PRMT1
G3I4A0	ACTG2		IAQDLEMYGINYFEIKNK	EZR
G3I3X4	Fructose-bisphosphate aldolase		DIVVAETLEDLDKNK	Reticulocalbin-3
G3HLK9	Coiled-coil domain-containing protein 43		TFVQEDVYDEFVER	Aldehyde dehydrogenase
G3HP82	Myomegalin		GKLEAIITPPPAK	Histone H2A
			RLAYAIIQFLHGQLR	STGA
			VTVLFAGQHIAK	Filamin-A
			SAFPFSDKLGELSDK	Clathrin interactor 1
			KQLADETLLK	Lamin-L
			LGLVKYGEK	ABCA1

Support Vector Machine

Support Vector Machine (SVM) modelling was performed using the e1071 package in R. Training and testing was carried out on data sets that were 17% of the full dataset, with accuracy shown in **Table 5**. No proteins or peptides identified as impactful in classification were common between the three replicates. The number of proteins and peptides found to be impactful in stress prediction is included in **Table 6**.

Table 5: Accuracy of the triplicate SVM models, using 17% of the full data to train and 17% to test.

Seed	Accuracy
101	0.6209263
110	0.6137227
123	0.6242157

Table 6: The number of accession numbers and peptide sequences found to be impactful in stress prediction, where mean prediction is greater than 0.65 and n≥2, where n is the number of occurrences in the dataset.

Seed	Accession	Peptide
101	40	3
110	15	53
123	28	0

Extreme Gradient Boost

Extreme Gradient Boost (XGBoost) modelling was performed using the XGBoost R package, with 300 rounds, trained on 70% of the data and tested on the remaining 30%. The average accuracy of the XGBoost models was 0.673 from the three replicates (**Table 7**). The total number of impactful accessions and peptides where mean prediction is greater than 0.9 and n>2, from each seed is in **Table 8**. Only seed 110 gave unique impactful accessions, with 14 proteins shared between all three replicates. Greater unique values were found in the

peptide sequences, with 18 sequences common to all three replicates (**Figure 5**). These common accessions and peptides are listed in **Table 9**.

Table 7: Accuracy of the triplicate XGBoost models, using 70% of the full data to train and 30% to test.

Seed	Accuracy
101	0.671
110	0.6747
123	0.6719

Table 8: The number of accession numbers and peptide sequences found to be impactful in stress prediction, where mean prediction is greater than 0.9 and n≥2, where n is the number of occurrences in the dataset.

Seed	Accession	Peptide
101	23	68
110	32	76
123	23	55

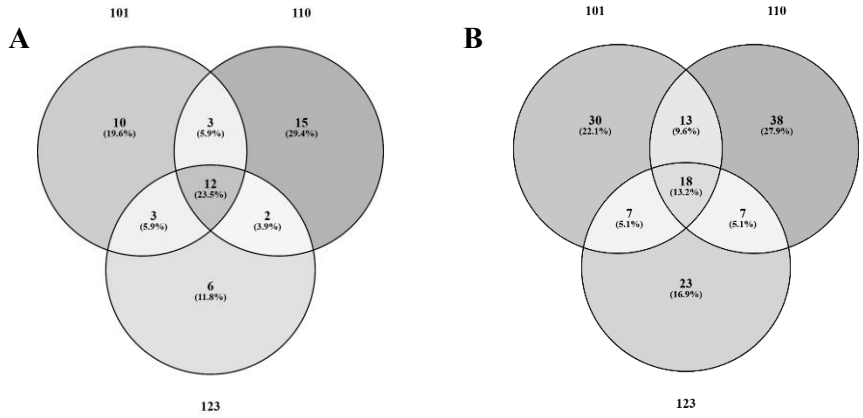


Figure 5: The shared accessions (A) and peptide sequences (B) found to be impactful in predicting stress in XGBoost models.

Table 9: The accession numbers and peptide sequences common to all three XGBoost seeds, found to be impactful in stress prediction, where mean prediction is greater than 0.9 and n≥2, where n is the number of occurrences in the dataset.

G3HP82	Myomegalin		QQLLEDLK	Myomegalin
G3H8Z6	PTTG1IP		EQNPYEKF	PTTG1IP
G3HN43	High mobility group protein B1		AQLVVIAHDVDPIELVVFLPALCRK	60S ribosomal protein L7a
G3I180	TRAF NF-kappa-B activator		IANVSILK	Vacuolar protein sorting-associated protein 33B
G3HZE5	HSP70kDA1		QLGSGLLLVTGPLALNRVPLRR	60S ribosomal protein L6
G3HXF7	Glyceraldehyde-3-phosphate dehydrogenase		GHYTEGAELVDSVLDVIR	Tubulin β chain
G3H215	40S ribosomal protein SA		TFSFAIPLIEK	Nucleolar RNA helicase 2
G3I4A0	ACTG2		TVIVAALDGTFQR	TK1
G3I3X4	Fructose-bisphosphate aldolase		VTVLFAGQHIAK	Filamin-A
G3H3F8	TK1		LGLVKYGEK	ABCA1
G3HKF3	RAI14		TFVQEDVYDEFVER	Aldehyde dehydrogenase
G3H0Q2	Cytochrome c oxidase polypeptide 7A2		RPWILTSLHNGVIQLWDYR	COPA
			DIVVAETLEDLDKNK	Reticulocalbin-3
			LDQLIYIPLPDEK	VCP
			RLAYAIIQFLHGQLR	SGTA
			AEELKTELLK	Actin

5.4 Discussion

Logistic regression

Logistic regression proved to be computationally intense, only able to process 5% of the total dataset. This was enough to be able to explore the relationship between proteins and peptide sequences and mild hypothermia. Due to the small dataset size, there are two proteins that appear in more than one of the seeds within the top twenty features correlated with stress. One of these proteins is Hypoxia up-regulated protein 1 (HYOU1), which plays a crucial role in cytoprotective cellular mechanisms (Wang et al., 2023). The other shared protein, High mobility group protein B1 (HMG1), is negatively correlated with stress. This protein acts as a DNA chaperone involved in replication, transcription, chromatin remodelling, recombination, DNA repair and genome stability (Wei et al., 2021). The corelation between this protein and non-stressed cells may indicate that expression of this protein is limited during periods of stress. High Mobility Group Box 2 (HMG2) is also found to be negatively correlated with stress in seed 123. It has similar functions to HMG1, involved in the regulation of transcription (Zimmerman and Maher, 2008). In seed 110, the protein most correlated with stress is Microsomal glutathione S-transferase 1 (MGST1). This protein is part of the MAPEG (membrane-associated proteins in eicosanoid and glutathione metabolism) superfamily, and has been found in the membranes of the mitochondria and the ER (McLellan et al., 1989). MGST1 has been proven to protects cell from oxidative stress (Morgenstern et al., 2011), so it may be possible that it also plays a role in ER stress protection. In seed 123, Cullin-associated NEDD8-dissociated protein 1 (CAND1) has a strong negative correlation with stress. This protein is linked to the inhibition of Cullin-RING ubiquitin ligases, which ubiquitinate proteins degraded by the proteasome system (Min et al., 2003). The correlation between this protein and non-stressed cells is possibly an indication of the reduced need for proteasomal degradation in these cells.

Some peptide sequences were also found to be corelated to stress, which allows for targeted MS methods. In seed 110, the peptide sequence with the greatest correlation to stress was *DSTLIMQLLR*, which comes from the 14-3-3 protein. This protein is part of a family of phospho-binding proteins that regulate most major cellular functions (Pennington et al., 2018). This protein family is the centre of a signalling hub linked to stress adaptation in cancer studies (Mhawech, 2005). 14-3-3 has been found to confer protection from apoptosis in yeast cells (Clapp et al., 2012). Few studies have been conducted investigating these proteins in CHO cells. Given their apparent role in stress signalling, they could be crucial factors in controlling stress adaptation on CHO cells.

Model performance

Random Forest performs best of all the models in predicting stress. The increase in accuracy with increased training data (**Table 2**) demonstrates the importance of large datasets in these types of models.

SVMs had the weakest performance in stress prediction. They are the most computationally intense methods, meaning the training dataset for this model had to be the smallest (RF-0.5, SVM-0.17, XGBoost-0.7 of whole data). The performance may be due the smaller dataset or may be due to the limitations of the model itself. Testing the model on a larger dataset would require greater computing power but would allow a fairer assessment of the method versus RF and XGBoost.

XGBoost is the least computationally intensive of the three ML methods tested but requires extra pre-processing steps that the others do not. Despite these extra steps, this is likely the most accessible method for a non-experienced user. This method is easily run on large dataset on a conventional desktop computer. The accuracy of this method is between that of the SVM and RF. This would likely increase with a larger database, which this method is best able to handle on a standard computer.

Proteins impactful in stress prediction

The proteins found to be impactful in predicting stress come from a range of biological functions, and in general, are not protein directed linked with ER stress. This may be reflective of the effects mild hypothermia has on the global cell environment, not just the ER. Many proteins that are ER stress predictors can be present in low levels in CHO cells without stress response activation, so without comparison of total expression levels these proteins may not be strong indicators. This includes BIP, ERP29, CANX and HSP90B1, which all have roles in normal ER function as molecular chaperones. They are crucial components of ER quality control mechanisms, not induced by stress but upregulated following UPR activation (Shaheen, 2018; Storey and Storey, 2023). Full lists of proteins impactful to prediction can be found in **Appendix V**.

In all three RF seeds, Pituitary tumour-transforming gene 1 protein-interacting protein (PTTG1IP) was found to be impactful in stress prediction. It was also found in two XGBoost seeds and one SVM seed. In a study of ER stress signatures in gastric adenocarcinoma patients, PTTG1IP expression was linked to poorer patient prognosis (Gong et al., 2022). PTTG1IP comprises p53 stability, stopping the production of pro-apoptotic proteins, with PTTG1IP-depleted cells showing increased apoptosis activity (Read et al., 2014). A recent study has found inhibitory by-products of amino acid metabolism in CHO cultures can lead to increased apoptosis, and found cell lines with greater PTTG1IP expression had better cell survival (Martínez et al., 2024). The ER protein database from STRING, used in the proteomic analysis in **Chapter 4**, contains 932 proteins, and PTTG1IP was not found on that list. This highlights the limitations in narrowing studies down using pathway or sub-cellular location lists. Although this protein can be tentatively linked to ER stress, it is too far removed from the defined processes to be considered in the relevant pathways. In this way, PTTG1IP may have been overlooked in the quest to find novel targets for cell line engineering.

Another protein identified in all three RF seeds was HSPA1A, a member of the heat shock 70KDa family. This protein aids in ER-associated degradation in the cytoplasm, binding to STUB1, which promotes ubiquitin-mediated protein degradation (Seo et al., 2016). This protein was one of 90 apoptosis regulators upregulated in a high producing CHO cell line (Orellana et al., 2018). Overexpression of HSP70 proteins has been found to extend CHO culture times up to 72h, with a corresponding 2.5-fold increase in recombinant protein production (Lee et al., 2009).

Thymidine kinase 1 is also important for stress prediction in RF, SVM and XGBoost models. Tight transcriptional regulation limits expression of this protein to the S phase of the cell cycle (Bitter et al., 2020). Previous studies have shown that mild hypothermia may arrest CHO cells in this phase (Fox et al., 2005). This demonstrates the effect of using a database with only one type of stress present. The presence of elevated TK1 does not, in this case, represent stress. It is reflective of the cell cycle phase that the stress condition in question, mild hypothermia, keeps the CHO cells in. This by-product of the dataset would possibly be less impactful in a dataset that included various other stress types. A TK1-associated reporter system or targeted MS method could be useful in detecting the level of cells in a culture that are in the S phase, so our knowledge of changes in TK1 expression could still have uses in CHO optimisation.

Myomegalin is another cell cycle progression related protein that is repeatedly linked to stress prediction (Bouguenina et al., 2017). Myomegalin, along with CAMSAP2 and AKAP9, stabilise microtubule attachment to the Golgi apparatus, and may play a part in the control of protein trafficking into the Golgi body (Wu et al., 2016).

The proteins identified as impactful in SVM models have lower prediction power than those of the other methods. This is likely a consequence of the use of a smaller dataset for SVM modelling, with less recurrence of accessions or peptides. HMOX1, which was found to be

upregulated in elevated ammonia cultures (Chapter 3), is found to be impactful in two SVM seeds. HMOX1 is an antioxidant enzyme that can act as an anti-apoptosis factor (Gozzelino et al., 2010). It is also associated with PERK activation through the PERK–Nrf2–HMOX1 signalling pathway. This is one way that cells are protected by the UPR in response to ER stress, through the activation of protective proteins, including HMOX1 (Cullinan et al., 2003; Zuo et al., 2025). In lung cancer cells experiencing ER stress, HMOX1 activation was linked to reduced levels of PERK, eIF2α, ATF4, and CHOP (Chen et al., 2018). A positive feedback loop has also been suggested, where carbon monoxide, a product of HMOX1 activity, activates PERK/Nrf2, which in turn increases HMOX1 expression (Kim et al., 2007).

Part of the Signal Receptor Particle (SRP), SRP14, is important for stress prediction in two SVM seeds. This complex mediates the co-translational targeting of secretory and membrane proteins to the ER (Weichenrieder et al., 2000). It is possible that this protein appears to be linked to the stress condition, as during mild hypothermia, the stressed cells are more productive than those at 37°C. Increased SRP activity could be due to increased transcription and translation during the arrested S phase of these cells.

Four ubiquitin-associated proteins are found to be useful in stress prediction in SVM models, USP36, HECTD3, LRSAM1 and USP4. USP36 is a de-ubiquitinase that regulates nucleolar structure and function, which is crucial to cell viability (Endo et al., 2009). It also regulates the actions of E3 ubiquitin-protein ligase NEDD4 (Anta et al., 2016). This protein demonstrates the broad impact of ubiquitination in cellular biology, outside of the proteasomal degradation. HECTD3 is an E3 ubiquitin ligase, which transfers ubiquitin from an E2 ubiquitin-conjugating enzyme to a target substrate (Bernassola et al., 2019). HECTD3 is known to ubiquitinate Caspase-8, inhibiting its ability to induce apoptosis (Yuan et al., 2022). LRSAM1is another E3 ubiquitin ligase (Mishra et al., 2019). It has been linked to the clearance of aberrant proteins thought the ubiquitin-proteasome system in neurodegenerative

diseases (Mishra et al., 2020). The overexpression of LRSAM1 could increase proteasome action, increasing the degradation of misfolded proteins, which would help maintain ER homeostasis (Dubey et al., 2022). USP4 is a deubiquitinating enzyme that removes ubiquitin from substrate proteins (Zhao et al., 2009). USP4 presence has been shown to limit p53 mediated apoptosis (Hu et al., 2021), similar to the action of PTTG1IP.

ACT2G, an actin protein, is found to be useful in stress prediction in all three models. In studies of mild hypothermia in CHO cells, it has been found that cell size increases in lower temperature cultures (Kim et al., 2001; Torres et al., 2021). Similar to TK1, ACT2G may be a hallmark of mild hypothermia on cell activity, rather than a part of an innate stress response.

ML-based proteins of interest in the broad-scope proteomic studies

When compared to the lists of differentially expressed proteins from the thapsigargin, hypothermia and ammonia studies, there are 81, 99 and 71 proteins shared with impactful identifications from ML, respectively. The fact that there is the greatest number shared with the hypothermia study is due to the fact that the MZID files used for ML analysis were from the hypothermia experiment. In the primary studies for each of the conditions, the proteins considered differentially expressed must have a fold-change of ≤ 1.5 between conditions, an ANOVA p-value of ≤ 0.05 and for thapsigargin and ammonia, at least 2 peptides identified from the protein. Given that these conditions are not applied to the MZID data, there are far more protein identifications present in these files. **Table 10** lists the fold-changes of these proteins that were identified as DE in the main studies and also discussed as ML-based identification of interest above. All shared proteins were found to be upregulated in the associated stress conditions. In situations where several timepoints or conditions were investigated, the highest value was chosen.

Table 10: The up-regulated fold-changes associated with proteins identified as impactful by ML methods that are also differentially expressed in the primary studies in this thesis.

	CHO-K1	**CHO-SK15-EPO**	**CHO-DP12**	**NISTCHO**
Thapsigargin				
HYOU1	2.7	3	4.7	*ND*
CAND1	*ND*	1.7	*ND*	*ND*
HMOX1	1.9	*ND*	12	*ND*
HECTD3	*ND*	7.3	*ND*	*ND*
Hypothermia				
MGST1	*ND*	3.3	*ND*	*ND*
TK1	3.1	*ND*	*ND*	*ND*
HMOX1	3.4	6.4	4.6	*ND*
HECTD3	13.9	*ND*	*ND*	*ND*
Ammonia				
HYOU1	4.3	*ND*	1.6	1.7
CAND1	1.7	*ND*	1.6	1.6
HMOX1	7.7	*ND*	2.4	2
HECTD3	24.3	*ND*	1.7	*ND*

5.5 Conclusions

The proteins identified within these ML models indicate the wider implications of ER stress induction of CHO cells. The decrease in culture temperature affects cellular function globally, not just in the ER. So, while we study the ER and the importance in correct recombinant protein production, these models remind us of the interconnected nature of cell biology. These models also demonstrate the ability to use big data in an accessible and manageable way. It may not be feasible to study the roles of over 5000 CHO proteins, and the selection of specific pathways to study closes off the impact of some protein-protein

interactions across pathways. Using bioinformatic methods we can assess entire proteomic datasets and find meaningful answers from all cellular functions.

The results discussed in this chapter do not form the same cohesive picture of stress as the previous chapters do. This is likely due to the segmentation of the dataset that was required to be able to run these models. While we can gather useful information, without much greater computing power, we cannot investigate the stress condition as a whole. The fact that this dataset contains proteins from the entire CHO proteome, the results discussed are not already linked in a single pathway. This makes bringing the results together far more difficult, as proteins that represent the links between pathways may not have been part of the models.

One recurring function of the proteins found to be important is the inhibition of apoptosis. This suggests that when cultured at a lower temperature, CHO cells must activate anti-apoptosis factors to prevent cell death as the cells adapt to the environmental change. PTTG1IP appearing as impactful in stress prediction in all model types is suggestive of a crucial role in ER stress adaptation. As an anti-apoptosis factor, over-expression of this protein could help create a stress resistant CHO cell line. USP4 and HMOX1, which also have anti-apoptosis roles, are also potential genetic engineering targets.

These results demonstrate the power of ML models to identify stress markers that are not typically considered in studies of ER stress. With a larger dataset, bringing together multiple stress types, the patterns attached to these stress markers could be made even clearer.

5.6 References

Akram, M., Siddique, S., Alharbi, M.G., 2022. Clustering algorithm with strength of connectedness for m-polar fuzzy network models. Mathematical Biosciences and Engineering 19, 420–455.

Alloghani, M., Al-Jumeily, D., Mustafina, J., Hussain, A., Aljaaf, A.J., 2020. A Systematic Review on Supervised and Unsupervised Machine Learning Algorithms for Data Science, in: Supervised and Unsupervised Learning for Data Science. Springer, Cham, pp. 3–21. https://doi.org/10.1007/978-3-030-22475-2_1

Anderson, N.L., Matheson, A.D., Steiner, S., 2000. Proteomics: applications in basic and applied biology. Curr Opin Biotechnol 11, 408–412. https://doi.org/10.1016/s0958-1669(00)00118-x

Anta, B., Martín-Rodríguez, C., Gomis-Perez, C., Calvo, L., López-Benito, S., Calderón-García, A.A., Vicente-García, C., Villarroel, Á., Arévalo, J.C., 2016. Ubiquitin-specific Protease 36 (USP36) Controls Neuronal Precursor Cell-expressed Developmentally Down-regulated 4-2 (Nedd4-2) Actions over the Neurotrophin Receptor TrkA and Potassium Voltage-gated Channels 7.2/3 (Kv7.2/3). J Biol Chem 291, 19132–19145. https://doi.org/10.1074/jbc.M116.722637

Beltran, J.C., Valdez, P., Naval, P., 2019. Predicting Protein-Protein Interactions based on Biological Information using Extreme Gradient Boosting, in: 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). Presented at the 2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), pp. 1–6. https://doi.org/10.1109/CIBCB.2019.8791241

Bernassola, F., Chillemi, G., Melino, G., 2019. HECT-Type E3 Ubiquitin Ligases in Cancer. Trends in Biochemical Sciences 44, 1057–1075. https://doi.org/10.1016/j.tibs.2019.08.004

Bitter, E.E., Townsend, M.H., Erickson, R., Allen, C., O'Neill, K.L., 2020. Thymidine kinase 1 through the ages: a comprehensive review. Cell & Bioscience 10, 138. https://doi.org/10.1186/s13578-020-00493-1

Bouguenina, H., Salaun, D., Mangon, A., Muller, L., Baudelet, E., Camoin, L., Tachibana, T., Cianférani, S., Audebert, S., Verdier-Pinard, P., Badache, A., 2017. EB1-binding-myomegalin protein complex promotes centrosomal microtubules functions. Proc Natl Acad Sci U S A 114, E10687–E10696. https://doi.org/10.1073/pnas.1705682114

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Chambers, M.C., Maclean, B., Burke, R., Amodei, D., Ruderman, D.L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T.A., Brusniak, M.-Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S.L., Nuwaysir, L.M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Detlev, S., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K.,

Eckels, J., Deutsch, E.W., Moritz, R.L., Katz, J.E., Agus, D.B., MacCoss, M., Tabb, D.L., Mallick, P., 2012. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 30, 918–920. https://doi.org/10.1038/nbt.2377

Chandola, V., Banerjee, A., Kumar, V., 2009. Anomaly detection: A survey. ACM Comput. Surv. 41, 15:1-15:58. https://doi.org/10.1145/1541880.1541882

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. pp. 785–794.

Chen, X., Wang, Y., Xie, X., Chen, H., Zhu, Q., Ge, Z., Wei, H., Deng, J., Xia, Z., Lian, Q., 2018. Heme Oxygenase-1 Reduces Sepsis-Induced Endoplasmic Reticulum Stress and Acute Lung Injury. Mediators Inflamm 2018, 9413876. https://doi.org/10.1155/2018/9413876

Clapp, C., Portt, L., Khoury, C., Sheibani, S., Norman, G., Ebner, P., Eid, R., Vali, H., Mandato, C.A., Madeo, F., 2012. 14-3-3 protects against stress-induced apoptosis. Cell death & disease 3, e348–e348.

Contini, C., Manconi, B., Olianas, A., Guadalupi, G., Schirru, A., Zorcolo, L., Castagnola, M., Messana, I., Faa, G., Diaz, G., Cabras, T., 2024. Combined High—Throughput Proteomics and Random Forest Machine-Learning Approach Differentiates and Classifies Metabolic, Immune, Signaling and ECM Intra-Tumor Heterogeneity of Colorectal Cancer. Cells 13, 1311. https://doi.org/10.3390/cells13161311

Craig, R., Cortens, J.P., Beavis, R.C., 2004. Open source system for analyzing, validating, and storing protein identification data. J Proteome Res 3, 1234–1242. https://doi.org/10.1021/pr049882h

Cullinan, S.B., Zhang, D., Hannink, M., Arvisais, E., Kaufman, R.J., Diehl, J.A., 2003. Nrf2 Is a Direct PERK Substrate and Effector of PERK-Dependent Cell Survival. Mol Cell Biol 23, 7198–7209. https://doi.org/10.1128/MCB.23.20.7198-7209.2003

Desiere, F., Deutsch, E.W., King, N.L., Nesvizhskii, A.I., Mallick, P., Eng, J., Chen, S., Eddes, J., Loevenich, S.N., Aebersold, R., 2006. The PeptideAtlas project. Nucleic Acids Res 34, D655-658. https://doi.org/10.1093/nar/gkj040

Dubey, A.R., Mishra, R., Sundaria, N., Jagtap, Y.A., Kumar, P., Kinger, S., Choudhary, A., Jha, H.C., Prasad, A., Gutti, R.K., Mishra, A., 2022. Resveratrol Promotes LRSAM1 E3 Ubiquitin Ligase-Dependent Degradation of Misfolded Proteins Linked with Neurodegeneration. Cellular Physiology & Biochemistry 56, 530–545.

Endo, A., Matsumoto, M., Inada, T., Yamamoto, A., Nakayama, K.I., Kitamura, N., Komada, M., 2009. Nucleolar structure and function are regulated by the deubiquitylating enzyme USP36. J Cell Sci 122, 678–686. https://doi.org/10.1242/jcs.044461

Fox, S.R., Tan, H.K., Tan, M.C., Wong, S.C.N.C., Yap, M.G.S., Wang, D.I.C., 2005. A detailed understanding of the enhanced hypothermic productivity of interferon-γ by Chinese-hamster

ovary cells. Biotechnology and Applied Biochemistry 41, 255–264. https://doi.org/10.1042/BA20040066

Freedman, D., 2009. Statistical Models: Theory and Practice. Cambridge University Press.

Friedman, J., Hastie, T., Tibshirani, R., Narasimhan, B., Tay, K., Simon, N., Qian, J., 2021. Package 'glmnet.' CRAN R Repositary 595, 874.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Annals of statistics 1189–1232.

Gao, J., Zhong, C., Chen, X., Lin, H., Zhang, Z., 2020. Unsupervised Learning for Passive Beamforming. https://doi.org/10.48550/arXiv.2001.02348

Gatto, L., Lilley, K.S., 2012. MSnbase-an R/Bioconductor package for isobaric tagged mass spectrometry data visualization, processing and quantitation. Bioinformatics 28, 288–289. https://doi.org/10.1093/bioinformatics/btr645

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A.J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J.Y.H., Zhang, J., 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5, R80. https://doi.org/10.1186/gb-2004-5-10-r80

Gong, M., Wang, J., Gao, W., Liu, Q., Chen, J., Wang, G., Zhang, Q., 2022. Establishment of an endoplasmic reticulum stress-related signature predicting outcomes of gastric adenocarcinoma patients. Front. Genet. 13. https://doi.org/10.3389/fgene.2022.944105

Gozzelino, R., Jeney, V., Soares, M.P., 2010. Mechanisms of cell protection by heme oxygenase-1. Annu Rev Pharmacol Toxicol 50, 323–354. https://doi.org/10.1146/annurev.pharmtox.010909.105600

Gygi, S.P., Corthals, G.L., 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, 9390–9395. https://doi.org/10.1073/pnas.160270797

Hastie, T., Tibshirani, R., Friedman, J., Franklin, J., 2005. The elements of statistical learning: data mining, inference and prediction. The Mathematical Intelligencer 27, 83–85.

Henzel, W.J., Billeci, T.M., Stults, J.T., Wong, S.C., Grimley, C., Watanabe, C., 1993. Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. Proc Natl Acad Sci U S A 90, 5011–5015. https://doi.org/10.1073/pnas.90.11.5011

Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the Dimensionality of Data with Neural Networks. Science 313, 504–507. https://doi.org/10.1126/science.1127647

Hosmer, D.W., Lemeshow, S., Sturdivant, R.X., 2013. Applied Logistic Regression: Third Edition. wiley. https://doi.org/10.1002/9781118548387

Hu, B., Zhang, D., Zhao, K., Wang, Y., Pei, L., Fu, Q., Ma, X., 2021. Spotlight on USP4: Structure, Function, and Regulation. Front. Cell Dev. Biol. 9. https://doi.org/10.3389/fcell.2021.595159

Huang, Y., Ma, S.-F., Oldham, J.M., Adegunsoye, A., Zhu, D., Murray, S., Kim, J.S., Bonham, C., Strickland, E., Linderholm, A.L., Lee, C.T., Paul, T., Mannem, H., Maher, T.M., Molyneaux, P.L., Strek, M.E., Martinez, F.J., Noth, I., 2024. Machine Learning of Plasma Proteomics Classifies Diagnosis of Interstitial Lung Disease. American Journal of Respiratory and Critical Care Medicine. https://doi.org/10.1164/rccm.202309-1692OC

James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An introduction to statistical learning: with applications in R. Springer.

Jiang, L., He, L., Fountoulakis, M., 2004. Comparison of protein precipitation methods for sample preparation prior to proteomic analysis. J Chromatogr A 1023, 317–320. https://doi.org/10.1016/j.chroma.2003.10.029

Jones, A.R., Eisenacher, M., Mayer, G., Kohlbacher, O., Siepen, J., Hubbard, S.J., Selley, J.N., Searle, B.C., Shofstahl, J., Seymour, S.L., Julian, R., Binz, P.-A., Deutsch, E.W., Hermjakob, H., Reisinger, F., Griss, J., Vizcaíno, J.A., Chambers, M., Pizarro, A., Creasy, D., 2012. The mzIdentML data standard for mass spectrometry-based proteomics results. Mol Cell Proteomics 11, M111.014381. https://doi.org/10.1074/mcp.M111.014381

Jones, L.D., Golan, D., Hanna, S.A., Ramachandran, M., 2018. Artificial intelligence, machine learning and the evolution of healthcare: A bright future or cause for concern? Bone Joint Res 7, 223–225. https://doi.org/10.1302/2046-3758.73.BJR-2017-0147.R1

Kearns, M., Li, M., Valiant, L., 1994. Learning boolean formulas. Journal of the ACM (JACM) 41, 1298–1328.

Kim, K.M., Pae, H.-O., Zheng, M., Park, R., Kim, Y.-M., Chung, H.-T., 2007. Carbon monoxide induces heme oxygenase-1 via activation of protein kinase R-like endoplasmic reticulum kinase and inhibits endothelial cell apoptosis triggered by endoplasmic reticulum stress. Circ Res 101, 919–927. https://doi.org/10.1161/CIRCRESAHA.107.154781

Kim, T.K., Chung, J.Y., Sung, Y.H., Lee, G.M., 2001. Relationship between cell size and specific thrombopoietin productivity in Chinese hamster ovary cells during dihydrofolate reductase-mediated gene amplification. Biotechnology and bioprocess engineering 6, 332–336.

Kurt Hornik and the R Core Team, 2024. The R FAQ [WWW Document]. The R FAQ. URL https://CRAN.R-project.org/doc/manuals/R-FAQ.html

Kutner, M.H., 2005. Applied Linear Statistical Models. McGraw-Hill Irwin.

Ladosz, P., Weng, L., Kim, M., Oh, H., 2022. Exploration in deep reinforcement learning: A survey. Information Fusion 85, 1–22. https://doi.org/10.1016/j.inffus.2022.03.003

Lee, Y.Y., Wong, K.T.K., Tan, J., Toh, P.C., Mao, Y., Brusic, V., Yap, M.G.S., 2009. Overexpression of heat shock proteins (HSPs) in CHO cells for extended culture viability

and improved recombinant protein production. Journal of Biotechnology 143, 34–43. https://doi.org/10.1016/j.jbiotec.2009.05.013

Li, B., Kugeratski, F.G., Kalluri, R., 2024. A novel machine learning algorithm selects proteome signature to specifically identify cancer exosomes. eLife 12, RP90390. https://doi.org/10.7554/eLife.90390

Li, L., Weinberg, C.R., Darden, T.A., Pedersen, L.G., 2001. Gene selection for sample classification based on gene expression data: study of sensitivity to choice of parameters of the GA/KNN method. Bioinformatics 17, 1131–1142. https://doi.org/10.1093/bioinformatics/17.12.1131

Libbrecht, M.W., Noble, W.S., 2015. Machine learning applications in genetics and genomics. Nat Rev Genet 16, 321–332. https://doi.org/10.1038/nrg3920

Machado, R.J.M., van den Hout, A., Marra, G., 2021. Penalised maximum likelihood estimation in multi-state models for interval-censored data. Computational Statistics & Data Analysis 153, 107057. https://doi.org/10.1016/j.csda.2020.107057

Mann, M., Højrup, P., Roepstorff, P., 1993. Use of mass spectrometric molecular weight information to identify proteins in sequence databases. Biol Mass Spectrom 22, 338–345. https://doi.org/10.1002/bms.1200220605

Martínez, V.S., Rodriguez, K., McCubbin, T., Tong, J., Mahler, S., Shave, E., Baker, K., Munro, T.P., Marcellin, E., 2024. Amino acid degradation pathway inhibitory by-products trigger apoptosis in CHO cells. Biotechnology Journal 19, 2300338. https://doi.org/10.1002/biot.202300338

Matthias Unterhuber, M.D., Karl-Patrik Kresoja, M.D., Karl-Philipp Rommel, M.D., Christian Besler, M.D., Andrea Baragetti, P., Nora Klöting, D., Uta Ceglarek, D., Matthias Blüher, M.D., Markus Scholz, D., Alberico L. Catapano, P., Holger Thiele, M.D., Philipp Lurz, M.D., 2021. Proteomics-Enabled Deep Learning Machine Algorithms Can Enhance Prediction of Mortality. Journal of the American College of Cardiology. https://doi.org/10.1016/j.jacc.2021.08.018

McLellan, L.I., Wolf, C.R., Hayes, J.D., 1989. Human microsomal glutathione S-transferase. Its involvement in the conjugation of hexachlorobuta-1,3-diene with glutathione. Biochem J 258, 87–93. https://doi.org/10.1042/bj2580087

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., Leisch, F., Chang, C.C., Lin, C.C., 2024. Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien [R package e1071 version 1.7-16]. CRAN Contrib. Packag.

Mhawech, P., 2005. 14-3-3 proteins—an update. Cell research 15, 228–236.

Min, K.-W., Hwang, J.-W., Lee, J.-S., Park, Y., Tamura, T., Yoon, J.-B., 2003. TIP120A associates with cullins and modulates ubiquitin ligase activity. J Biol Chem 278, 15905–15910. https://doi.org/10.1074/jbc.M213070200

Mishra, R., Amanullah, A., Upadhyay, A., Dhiman, R., Dubey, A.R., Singh, S., Prasad, A., Mishra, A., 2020. Ubiquitin ligase LRSAM1 suppresses neurodegenerative diseases linked aberrant

proteins induced cell death. The International Journal of Biochemistry & Cell Biology 120, 105697. https://doi.org/10.1016/j.biocel.2020.105697

Mishra, R., Upadhyay, A., Prajapati, V.K., Dhiman, R., Poluri, K.M., Jana, N.R., Mishra, A., 2019. LRSAM1 E3 ubiquitin ligase: molecular neurobiological perspectives linked with brain diseases. Cell Mol Life Sci 76, 2093–2110. https://doi.org/10.1007/s00018-019-03055-y

Moeyersoms, J., Martens, D., 2015. Including high-cardinality attributes in predictive models: A case study in churn prediction in the energy sector. Decision Support Systems 72, 72–81. https://doi.org/10.1016/j.dss.2015.02.007

Morgenstern, R., Zhang, J., Johansson, K., 2011. Microsomal glutathione transferase 1: mechanism and functional roles. Drug Metabolism Reviews 43, 300–306. https://doi.org/10.3109/03602532.2011.558511

Naeem, S., Ali, A., Anam, S., Ahmed, M.M., 2023. An unsupervised machine learning algorithms: Comprehensive review. International Journal of Computing and Digital Systems.

Neely, B.A., Dorfer, V., Martens, L., Bludau, I., Bouwmeester, R., Degroeve, S., Deutsch, E.W., Gessulat, S., Käll, L., Palczynski, P., Payne, S.H., Rehfeldt, T.G., Schmidt, T., Schwämmle, V., Uszkoreit, J., Vizcaíno, J.A., Wilhelm, M., Palmblad, M., 2023. Toward an Integrated Machine Learning Model of a Proteomics Experiment. J. Proteome Res. 22, 681–696. https://doi.org/10.1021/acs.jproteome.2c00711

Noble, W.S., 2006. What is a support vector machine? Nature biotechnology 24, 1565–1567.

Ohri, K., Kumar, M., 2021. Review on self-supervised image recognition using deep neural networks. Knowledge-Based Systems 224, 107090.

Orellana, C.A., Marcellin, E., Palfreyman, R.W., Munro, T.P., Gray, P.P., Nielsen, L.K., 2018. RNA-Seq Highlights High Clonal Variation in Monoclonal Antibody Producing CHO Cells. Biotechnol J 13, e1700231. https://doi.org/10.1002/biot.201700231

Oshiro, T.M., Perez, P.S., Baranauskas, J.A., 2012. How Many Trees in a Random Forest?, in: Machine Learning and Data Mining in Pattern Recognition. Presented at the International Workshop on Machine Learning and Data Mining in Pattern Recognition, Springer, Berlin, Heidelberg, pp. 154–168. https://doi.org/10.1007/978-3-642-31537-4_13

Pennington, K.L., Chan, T.Y., Torres, M.P., Andersen, J.L., 2018. The dynamic and stress-adaptive signaling hub of 14-3-3: emerging mechanisms of regulation and context-dependent protein–protein interactions. Oncogene 37, 5587–5604. https://doi.org/10.1038/s41388-018-0348-3

Perez-Riverol, Y., Bandla, C., Kundu, D.J., Kamatchinathan, S., Bai, J., Hewapathirana, S., John, N.S., Prakash, A., Walzer, M., Wang, S., 2025. The PRIDE database at 20 years: 2025 update. Nucleic Acids Research 53, D543–D553.

Pisner, D.A., Schnyer, D.M., 2020. Support vector machine, in: Machine Learning. Elsevier, pp. 101–121.

Read, M.L., Seed, R.I., Fong, J.C., Modasia, B., Ryan, G.A., Watkins, R.J., Gagliano, T., Smith, V.E., Stratford, A.L., Kwan, P.K., 2014. The PTTG1-binding factor (PBF/PTTG1IP) regulates p53 activity in thyroid cells. Endocrinology 155, 1222–1234.

Rodriguez, M.Z., Comin, C.H., Casanova, D., Bruno, O.M., Amancio, D.R., Costa, L. da F., Rodrigues, F.A., 2019. Clustering algorithms: A comparative approach. PloS one 14, e0210236.

Seo, J.H., Park, J.-H., Lee, E.J., Vo, T.T.L., Choi, H., Kim, J.Y., Jang, J.K., Wee, H.-J., Lee, H.S., Jang, S.H., Park, Z.Y., Jeong, J., Lee, K.-J., Seok, S.-H., Park, J.Y., Lee, B.J., Lee, M.-N., Oh, G.T., Kim, K.-W., 2016. ARD1-mediated Hsp70 acetylation balances stress-induced protein refolding and degradation. Nat Commun 7, 12882. https://doi.org/10.1038/ncomms12882

Shaheen, A., 2018. Effect of the unfolded protein response on ER protein export: a potential new mechanism to relieve ER stress. Cell Stress and Chaperones 23, 797–806.

Storey, J.M., Storey, K.B., 2023. Chaperone proteins: universal roles in surviving environmental stress. Cell Stress and Chaperones, Organisms in a Changing Environment 28, 455–466. https://doi.org/10.1007/s12192-022-01312-x

Suthaharan, S., 2016. Support Vector Machine, in: Suthaharan, S. (Ed.), Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning. Springer US, Boston, MA, pp. 207–235. https://doi.org/10.1007/978-1-4899-7641-3_9

Thomas Pedersen, Vladislav Petyuk, Laurent Gatto, 2025. mzID: An mzIdentML parser for R.

Torres, M., Akhtar, S., McKenzie, E.A., Dickson, A.J., 2021. Temperature Down-Shift Modifies Expression of UPR-/ERAD-Related Genes and Enhances Production of a Chimeric Fusion Protein in CHO Cells. Biotechnology Journal 16, 2000081. https://doi.org/10.1002/biot.202000081

Turner, S., 2024. R 4.5.0 and Bioconductor 3.21 [WWW Document]. URL https://blog.stephenturner.us/p/r-450-bioconductor-321 (accessed 8.12.25).

Van Engelen, J.E., Hoos, H.H., 2020. A survey on semi-supervised learning. Machine learning 109, 373–440.

Wang, S., Wei, D., Zhao, Y., Pang, X., Zhang, Z., 2024. Development and validation of machine learning models for diagnosis and prognosis of cancer by urinary proteomics, based on the FLEMENGHO cohort. Am J Cancer Res 14, 643–654. https://doi.org/10.62347/COTC6463

Wang, W., Jiang, X., Xia, F., Chen, X., Li, G., Liu, L., Xu, Q., Zhu, M., Chen, C., 2023. HYOU1 promotes cell proliferation, migration, and invasion via the PI3K/AKT/FOXO1 feedback loop in bladder cancer. Mol Biol Rep 50, 453–464. https://doi.org/10.1007/s11033-022-07978-x

Webb-Robertson, B.-J.M., Cannon, W.R., Oehmen, C.S., Shah, A.R., Gurumoorthi, V., Lipton, M.S., Waters, K.M., 2010. A support vector machine model for the prediction of proteotypic

peptides for accurate mass and time proteomics. Bioinformatics 26, 1677–1683. https://doi.org/10.1093/bioinformatics/btq251

Webb-Robertson, B.-J.M., Matzke, M.M., Jacobs, J.M., Pounds, J.G., Waters, K.M., 2011. A statistical selection strategy for normalization procedures in LC-MS proteomics experiments through dataset-dependent ranking of normalization scaling factors. Proteomics 11, 4736–4741. https://doi.org/10.1002/pmic.201100078

Wei, J., Alfajaro, M.M., DeWeirdt, P.C., Hanna, R.E., Lu-Culligan, W.J., Cai, W.L., Strine, M.S., Zhang, S.-M., Graziano, V.R., Schmitz, C.O., Chen, J.S., Mankowski, M.C., Filler, R.B., Ravindra, N.G., Gasque, V., de Miguel, F.J., Patil, A., Chen, H., Oguntuyo, K.Y., Abriola, L., Surovtseva, Y.V., Orchard, R.C., Lee, B., Lindenbach, B.D., Politi, K., van Dijk, D., Kadoch, C., Simon, M.D., Yan, Q., Doench, J.G., Wilen, C.B., 2021. Genome-wide CRISPR Screens Reveal Host Factors Critical for SARS-CoV-2 Infection. Cell 184, 76-91.e13. https://doi.org/10.1016/j.cell.2020.10.028

Weichenrieder, O., Wild, K., Strub, K., Cusack, S., 2000. Structure and assembly of the Alu domain of the mammalian signal recognition particle. Nature 408, 167–173. https://doi.org/10.1038/35041507

Wickham, H., 2016. ggplot2: Elegant Graphics for Data Analysis. Springer International Publishing.

Wright, M.N., Ziegler, A., 2017. ranger: A fast implementation of random forests for high dimensional data in C++ and R. Journal of statistical software 77, 1–17.

Wu, J., de Heus, C., Liu, Q., Bouchet, B.P., Noordstra, I., Jiang, K., Hua, S., Martin, M., Yang, C., Grigoriev, I., Katrukha, E.A., Altelaar, A.F.M., Hoogenraad, C.C., Qi, R.Z., Klumperman, J., Akhmanova, A., 2016. Molecular Pathway of Microtubule Organization at the Golgi Apparatus. Dev Cell 39, 44–60. https://doi.org/10.1016/j.devcel.2016.08.009

Xu, S., Liu, S., Wang, H., Chen, W., Zhang, F., Xiao, Z., 2021. A Hyperspectral Image Classification Approach Based on Feature Fusion and Multi-Layered Gradient Boosting Decision Trees. Entropy 23, 20. https://doi.org/10.3390/e23010020

Yates, J.R., Speicher, S., Griffin, P.R., Hunkapiller, T., 1993. Peptide mass maps: a highly informative approach to protein identification. Anal Biochem 214, 397–408. https://doi.org/10.1006/abio.1993.1514

Yin, H., Xie, J., Xing, S., Lu, X., Yu, Y., Ren, Y., Tao, J., He, G., Zhang, L., Yuan, X., Yang, Z., Huang, Z., 2024. Machine learning-based analysis identifies and validates serum exosomal proteomic signatures for the diagnosis of colorectal cancer. Cell Reports Medicine 5, 101689. https://doi.org/10.1016/j.xcrm.2024.101689

Yuan, L., Li, P., Zheng, Q., Wang, H., Xiao, H., 2022. The Ubiquitin-Proteasome System in Apoptosis and Apoptotic Cell Clearance. Front. Cell Dev. Biol. 10. https://doi.org/10.3389/fcell.2022.914288

Zhang, Yingjie, Zhang, Ying, Chen, Q., Ai, Z., Gong, Z., 2018. True-link clustering through signaling process and subcommunity merge in overlapping community detection. Neural Computing and Applications 30, 3613–3621.

Zhao, B., Schlesiger, C., Masucci, M.G., Lindsten, K., 2009. The ubiquitin specific protease 4 (USP4) is a new player in the Wnt signalling pathway. J Cell Mol Med 13, 1886–1895. https://doi.org/10.1111/j.1582-4934.2008.00682.x

Zimmerman, J., Maher, L.J., 2008. Transient HMGB protein interactions with B-DNA duplexes and complexes. Biochem Biophys Res Commun 371, 79–84. https://doi.org/10.1016/j.bbrc.2008.04.024

Zuo, L., Zou, X., Ge, J., Hu, S., Fang, Y., Xu, Y., Chen, R., Xu, S., Yu, G., Zhou, X., Ji, L., 2025. The Nrf2-HMOX1 pathway as a therapeutic target for reversing cisplatin resistance in non-small cell lung cancer via inhibiting ferroptosis. Cell Death Discov. 11, 287. https://doi.org/10.1038/s41420-025-02564-z

Chapter 6: Overview of major findings and conclusions

This thesis explores the proteomic effects of ER stress on CHO cell culture viability and productivity. Using the artificial ER stress inducer, thapsigargin and bioprocess relevant conditions, lower temperature and excess ammonia, this research profiles how the ER adapts to changes in the cell environment. This was done in pursuit of expanding our existing understanding of CHO cell biology and identifying proteins and pathways that could be potential genetic engineering targets. As the variety of CHO-derived products increases, it is necessary to further adapt the secretory system of these cells. The ER stress effect that recombinant product load exerts on the host cell line is a limiting factor in production. By increasing our understanding of the CHO ER stress response, we can try to reduce the negative effects of complex modalities on productivity. Through artificial stress induction in three CHO cell lines using thapsigargin, an overview of ER stress was created, comparing two producer cell lines and one non-producer. These three cell lines were also cultured under mild hypothermia (31°C), to better our understanding of this industrially relevant method of increasing productivity. As part of this investigation, the ubiquitinated proteome was also explored, to better understand the role of this critical PTM in the stress response. This study represents the first time the ubiquitinated proteome of CHO cells has been studied under these conditions. The role of PTMs has been well documented for decades, but traditional proteomic studies do not include this level of detail. We cannot expect to accurately genetically engineer the pathways of CHO cells without knowledge of the cell signalling and regulation involved in these pathways. Continuing the investigation of bioprocess induced stress, two IgG1 producers and one non-producer cell line were cultured in elevated ammonia environments. In batch-cultures, excessive ammonia levels can inhibit culture growth and limit productivity. At present, the most effective way to combat ammonia stress is using media-demanding continuous of perfusion cultures. By furthering our knowledge of the cellular response to elevated ammonia, we can potentially adapt CHO cells to remain

productive in batch cultures without the need for expensive culture filtering equipment. This study provides the most current, comprehensive study of the CHO proteome in response to ammonia stress.

Comparing the ER stress response between the different stressors allows us to identify proteins which are intrinsic to any change in ER function, and which are stressor dependent. The proteomic response to thapsigargin exposure shows greater similarity to that observed under hypothermic conditions than to the response elicited by ammonia stress. The most common overlap between the artificially induced and hypothermia-induced stress is in polypeptide folding and glycosylation pathways in the ER. GANAB is downregulated in both cell lines, in both thapsigargin exposed cells, and temperature shifted cells at day 4 and day 7. UGGT1 is downregulated in both cell lines in thapsigargin exposed cells and at 31°C on day 4, but not at day 7. Both proteins are part of the glycosylation pathway, a crucial part of ensuring the correct protein folding. The protein disulfide isomerases, PDIA4 and PDIA6 are also downregulated in both thapsigargin and hypothermia exposed cells, which are also crucial elements of protein folding processes. The abundance of shared downregulated proteins between cells exposed to thapsigargin and cells cultured at 31°C indicates that both stressors impact the cells' ability to fold proteins correctly, but the absence of many shared upregulated ERAD genes suggests that the extent of this issue is lesser in temperature shifted cells than in thapsigargin exposed cells. The data suggests that the build-up of misfolded proteins in cells cultured at 31°C is not so drastic as to require significant upregulation of the proteins involved in the degradation of the misfolded proteins, when compared to the upregulation of these proteins seen in thapsigargin exposed cells. This may indicate that cells can adapt to a lower temperature, but that the continued presence of thapsigargin does not allow the recovery of ER homeostasis. In both conditions, producer cells showed a greater level of expression of proteins linked to protein targeting to ubiquitin-ligase complex and ERAD following ER stress induction when compared to the non-producer cell line,

suggesting adapted mechanisms for the restoration of homeostasis within the ER. This may be due to the level of ER stress already place on CHO cells which have been engineered to produce recombinant products.

The increase in expression of apoptosis-related proteins without a corresponding increase in dead cells within the culture may be due to activation of these proteins following the shock of temperature reduction, which is then mitigated as cell adjust to the lower temperature. While this finding may seem paradoxical, there are some possible explanations, which could form the basis for future, in depth exploration of apoptotic protein activity. It has been found that lower temperatures lead to a reduction in metabolic activity, along with cell cycle arrest, so it is possible that while these apoptosis trigger proteins are upregulated, the corresponding cellular pathways and cascades may be slowed, delaying cell death. Often apoptosis-linked proteins do not act in isolation, so while these proteins are upregulated, the other mechanisms may not have been activated to a level that commits the full cell to apoptosis.

Exposure to excess ammonia seems to affect cells on a more global scale, interfering with metabolism and RNA transport processes rather than directly on ER function. This wider interference does have a knock-on effect on ER function, diminishing the protein load requiring folding, leading to the downregulation of many ER chaperones, but also some ER stress markers such as BIP. Proteins within the ER lumen seem to be protected from elevated ammonia levels, while many ER membrane bound proteins are downregulated. This includes several secretory proteins, which may limit the productivity of producer CHO cells in elevated ammonia cultures.

This thesis utilises proteomic analysis to gain insight into the changes in protein expression in CHO cells experiencing different stress conditions. This was done with the aim of identifying potential pathways for increasing bioprocess yield. The study was subject to some limitations based on the experimental designs used and the analytical approach. The

stress culture methods used in this research do not always best reflect bioprocess conditions but instead act as models for that stress type in CHO cells. Thapsigargin exposure does not accurately replicate the ER stress induced by recombinant protein production but allows us to study the hallmarks of ER stress in these cell lines, in a defined, controlled manner. The varied results often seen from -omic based CHO studies highlight the abundance of cell line specific changes, which this thesis aimed to account for by using different producer and non-producer cell lines in these studies. This made the identification of any common changes more robust and generalisable. The "spiking-in" of ammonia also does not directly mirror the real-world accumulation of ammonia in a bioreactor. This method, used often in published studies, ensures that both cell lines experience the same stress level at the same time, improving reproducibility and making it easier to attribute observed effects to ammonia exposure rather than culture variability.

The role of ubiquitin in the degradation of proteins that are no longer required by the cell means that the ubiquitinated proteome is a transient and dynamic environment. The study of ubiquitinated proteins during mild hypothermia described in chapter 3 of this thesis only looks at the ubiquitinated proteins at two times points. This restricts the insight gained into the short-lived nature of ubiquitin attachment. Given the cost of the large-scale culturing and enrichment for ubiquitinated proteins studies, there was little scope to include a greater number of sampling timepoints. The chosen sampling times of 48- and 120-hours after temperature downshift reflect both early and later stage culture conditions, allowing some comparison of the temporal changes in ubiquitination of proteins.

The use of mass spectrometry proteomics allows for high-throughput insight into the changes in the CHO proteome, this method is biased towards more abundant proteins, potentially leading to an underrepresentation of low-abundance ER stress regulators. This cell lysis, protein digestion and separation methods used can also miss out on highly hydrophilic peptides, very small proteins or membrane proteins. The use of fractionation

steps prior to MS analysis, as well as target MS methods would allow future works to bridge the gaps in the methodology.

Consequently, the findings described within this thesis should be interpreted with the acknowledgement of these technical and biological limitations, but the chosen approaches remain appropriate for the stated goal of generating clearer insight into the proteomic changes during periods of ER stress in CHO cells.

This thesis expands our understanding of CHO cell response to stress, providing clarity and new insight into ER function regulation. This allows us to consider clearly the methods we use to adapt CHO cells to produce complex therapeutics. While genomic studies of CHO cells provide information on the blueprint of these cells, proteomic studies provide a dynamic, functional view of a cell, including how proteins are modified and interact. This is crucial for understanding how cells adapt to stress and how biological processes change with time in culture. Including ubiquitination analysis in a proteome study allows an even clearer understanding of the dynamic regulation of the CHO cell response to lower culture temperature. This thesis emphasises how the role of PTMs cannot be overlooked when we consider pathway alteration for enhanced production. The importance of using bioinformatic methods to handle the large datasets arising from such proteomic studies is also explored, utilising machine learning methods to supply proteins of interest in stress adaptation overlooked in the main studies.

The conclusions specific to each chapter are summarised in the following pages.

Chapter 2: Paper 1-Whole cell proteomics of CHO cells experiencing artificially induced endoplasmic stress through thapsigargin exposure

- Thapsigargin-induced ER stress elicits a robust and cell line-dependent proteomic response in CHO cells, demonstrating some shared and distinct molecular adaptations among non-producer and producer lines.
- Core UPR and ERAD pathways are consistently upregulated across all lines, indicating conserved mechanisms of ER stress sensing and mitigation in CHO cell biology.
- Recombinant protein-producing CHO lines (CHO-DP12 and CHO-SK15-EPO) show enhanced proteostasis mechanisms, with significant upregulation of ERAD-related components, ubiquitin ligases, and protein quality control compared to the non-producing CHO-K1 line.
- Thapsigargin exposure induces not only ER-specific stress pathways but also broad cellular reprogramming, including alterations in: Amino acid biosynthesis, Aminoacyl-tRNA synthetase expression, RNA transport and processing, suggesting a global stress response encompassing transcriptional, translational, and metabolic adjustments.
- Increased expression of protein disulphide isomerases (PDIs), molecular chaperones, and glycoprotein processing enzymes reinforces the role of protein folding and glycosylation in adaptive stress response.
- Identification of multiple putative engineering targets for enhancing cellular stress resilience and secretory performance in recombinant CHO cells, including UPR regulators (BIP), ERAD components (SEL1L, HRD1), and PDIs.

Chapter 3: Paper 2-Proteomic and Ubiquitinated Proteome Insights into ER Stress Responses in Chinese Hamster Ovary Cells Under Mild Hypothermic Conditions

- Temperature downshift to 31°C enhances both overall titre and specific cell productivity, with CHO-K1, CHO-DP12 and CHO-SK15-EPO cells having reduced proliferation but longer culture lifespans.
- Several proteins involved in the glycosylation of nascent polypeptides are downregulated at 31°C, including DDOST, P4HB, PRKSCH and LMAN1, in all cell lines studied, suggesting that mild hypothermic shock disrupts the cell's normal ability to fold new proteins, leading to ER stress as the misfolded proteins build up.
- The upregulation of transcription factor eIF2α in producer cells at 31°C, where it is downregulated in non-producer cells suggests a more adapted ER stress response in recombinant protein-producing CHO cells which allows more efficient function under hypothermic conditions.
- Type of recombinant product can impact the ER stress response, with the IgG producer and EPO producer having different expression of transport proteins at 31°C, including several secretory proteins-SEC12, SEC23, SEC24.
- Enrichment of the ubiquitinated proteome indicates that pathways related to protein binding, enzyme binding and organic cyclic compound binding are under tight cellular regulation during mild hypothermic stress.
- Both ubiquitinated and non-ubiquitinated forms of proteins are present in cells at the same time, sometimes being differently regulated, suggesting both forms play a role in the cellular response to hypothermic stress.
- Proteins that are needed for correct ubiquitin regulation in the ER are themselves, at times, ubiquitinated, and given that this was seen in all three cell lines, it is likely a vital part of the stress response.

Chapter 4: Paper 3-Proteomic Analysis of Endoplasmic Reticulum (ER) Protein Expression under Ammonia-Induced Stress in Chinese Hamster Ovary (CHO) Cell Cultures

- Elevated ammonia reduces CHO cell proliferation and recombinant protein titre, likely due to the downregulation of cellular metabolism, nitrogen compound metabolism, and organic cyclic compound metabolism pathways.
- When focusing on ER protein expression, ER membrane proteins are most affected, with ER lumen proteins involved in folding less impacted. Downregulation of some ER chaperones in elevated ammonia cultures is possibly linked to a reduction in translation and trafficking of polypeptides into the ER.
- Downregulation of SRPRα and SEC11α in NISTCHO cells suggests reduced protein folding capacity due to limited polypeptide entry into the ER.
- In producer cell lines, transmembrane proteins related to Golgi-transport are upregulated, while several SEC proteins are downregulated, indicating Golgi-mediated protein transport is more stable in cells exposed to long term ammonia stress.
- Upregulation of anti-apoptosis factors HMOX1 and HMOX2 suggests a protective role in adapting to ammonia-induced stress.
- ERAD-related proteins (UBXN4, UBXN7, FAF2) indicate misfolded protein degradation occurs at later culture timepoints under higher excess ammonia, while overall ER homeostasis and cell viability are largely maintained.
- Even low ammonia levels impact multiple cellular processes, highlighting potential for engineering CHO cells to tolerate ammonia and improve recombinant protein production.

Chapter 5: Regression and machine learning methods for the analysis of mass spectrometry ID (mzID) files.

- Machine learning methods can be applied to mass spectrometry identity files and used to identify proteins and peptides that are linked with stress in the algorithm.
- Anti-apoptosis factors are repeatedly linked to stress prediction. These include HMOX1, USP4 and PTTG1IP, which could be potential targets for genetic engineering for stress resilience in CHO cells.
- Given the use of mild hypothermia data in these models, proteins that are indicators of the changed cell physiology are also linked to the prediction of this stress condition. TK1 is hallmark of the S phase of the cell cycle. Finding markers of specific stages of cell activity could help better our understanding of all aspects of cell biology in these conditions, which could lead to enhanced production.
- Ubiquitin ligases were also found to be important for stress prediction, supporting the research of Chapter 3, which explores the crucial role of ubiquitin in the ER stress response.
- Regression and ML methods can generate novel targets for genetic engineering, as they are not confined to considering specific pathways.

Chapter 7: Future Work

Based on the findings of this thesis, several avenues for future research are available that would expand on the results and have an impact on CHO cell production.

The results described within this thesis show the dynamic nature of endoplasmic reticulum protein expression in stress conditions. The cell lines studied show distinct, but overlapping, responses to thapsigargin, hypothermia and elevated ammonia. Thapsigargin stress induces a strong UPR/ERAD activation, along with alteration of cell metabolism and RNA processing. In mild hypothermia cultures, cells exhibit decreased proliferation and increased productivity, with a downregulation of glycosylation machinery and prominent ubiquitin-related regulation. Elevated ammonia conditions cause reduced proliferation and titre, with ER membrane protein expression most affected and the activation of anti-apoptotic factors. Producer cell lines generally show more robust proteostasis mechanisms than non-producer cells. Thapsigargin exposure and mild hypothermia both elicit a similar disruption to glycosylation mechanism proteins in the ER.

The priority moving forward with the results of this work should be based around the anti-apoptosis proteins found to be upregulated in stress conditions, such as HMOX1, HMOX2, USP4 and PTTG1IP. Knock-downs of these proteins could confirm their roles in cell survival, and overexpression studies could investigate if increased expression of these proteins allows for longer cultures with greater titre. If these studies prove useful, stable cell lines could be developed, which could be utilised in industrial bioprocessing settings. Characterisation studies would have to be carried out in conjunction, to ensure these adapted cell lines and extended culture times would be producing useable products.

Further investigation of ER stress

Repeating the analysis across additional CHO cell lines with more complex product types (e.g. other monoclonal antibodies, fusion proteins, antibody-drug-conjugates, bispecific

antibodies, trispecific antibodies) would test the broad applicability of the findings and identify cell/product-specific optimisation strategies. Due to issues like protein misfolding, aggregation, or degradation, complex modalities are considered difficult-to-express (DTE) products. This thesis has shown that ER adaptation varies between high and low IgG producing cell lines (NISTCHO and CHO-DP12, respectively, and between IgG and EPO (CHO-SK15-EPO) producers. To engineer host cell lines for DTE products, an understanding of how that modality interacts with ER and secretory processes must be established. The experimental design set out in this thesis shows how this interaction can be profiled using proteomic analysis.

Targeted mass spectrometry methods could be designed to monitor ER stress by selecting a panel of peptides representing UPR and ERAD proteins discussed in this thesis. Targeted proteomics monitors proteotypic peptides with a unique amino acid sequence to identify a specific protein of interest during mass spectrometry analysis. The use of mzID files, as explored in chapter 5 of this thesis, allows quick identification of specific peptides, their experimental mass and their retention time, thus making the design of targeted methods easier. Targeted methods can be performed with shorter run times than tradition shotgun proteomics, meaning results can be gained much faster than by studying the whole cell proteome.

Chapter 3 of this thesis explored the role of ubiquitination in the CHO stress response. This highlighted the importance of including PTM characterisation in proteomic studies. Ubiquitin is not the only PTM involved in regulation of cell signalling, and it is likely that other PTMs interact with ubiquitination during times of cell stress. Previous studies have shown how the phosphoproteome of CHO cells changes significantly during the growth phases (Kaushik et al., 2018). Given the arrest of CHO cell cycle following temperature downshift, it is possible that phosphorylation may play an important role in cell signalling

during this stress adaptation. Further work could investigate the relationship between different PTMs, i.e. ubiquitination and phosphorylation, during periods of ER stress.

Multi-omic analysis (transcriptome, metabolome) would provide a systems-level view of ER stress adaptation and uncover regulatory mechanisms not detectable at the protein level alone. Time-course proteomic and transcriptomic profiling following ER stress induction could uncover the kinetics of adaptive responses and reveal early-stage regulators with potential for modulation. While each omic study provides crucial insight, looking at each part of CHO cell biology in insolation ignores the intrinsic connection between these processes.

Adaptation of CHO cells to increase productivity and culture longevity

This thesis has identified many differentially expressed proteins in CHO cells related to ER stress. It has been shown in previous studies that the altered expression of UPR and ERAD proteins, such as BIP and ERO1, can increase CHO productivity and culture viability (Cain et al., 2013; Jiang et al., 2016). Functional overexpression or knockout studies of these proteins (e.g. SEL1L, HRD1, PDIA4) could validate their specific roles in enhancing secretion efficiency in producer CHO lines. Transient transfections through lipofectamine-mediated delivery of target protein siRNAs could provide insight into the value of changing the expression of these proteins. Should these exploratory experiments prove beneficial to increasing productivity or stress resistance, more stable transfection methods, such as CRISPR-Cas9, could be used to develop an optimised producer CHO host cell line. Over-expression of anti-apoptosis factors identified as important in the CHO stress response, such as HMOX1, HMOX2 and USP4, could increase stress resilience, increase culture longevity and reduce the need for expensive online culture interventions.

At 10 mM ammonia supplementation, NISTCHO cells show a minor increase in specific cell productivity. This suggests that similar to mild hypothermia, low level ammonia stress

may increase productivity while viable cell density is negatively affected. This increase in productivity was not found in the 30 mM ammonia supplemented cultures. Future work could investigate ammonia concentrations, around a 10 mM baseline, which could lead to greater productivity.

Creating reporter cell lines with fluorescent or luminescent UPR markers could enable real-time monitoring of ER stress in manufacturing conditions and support efficient screening of engineered variants. This method has previously been used successfully for the monitoring of BIP expression in CHO cells. Using flow cytometry to measure the increased expression of mGFP tagged BIP after thapsigragin treatment or batch culture, it was shown that real-time detection of ER stress was possible (Kyeong and Lee, 2022). It is possible that the using a system such as this for other UPR proteins, could allow better understanding of the dynamic nature of ER stress responses.

Downstream analysis of the effect of elevated ER stress conditions, focusing on the effects of these conditions on the structure of recombinant products, including post-translational modifications, such as N-glycosylation would need to be carried out. Product degradation is problematic as it could result in reduced biological functions or elicit an immune response in a patient. Proteases involved in the enzymatic degradation of proteins can cleave monoclonal antibodies in conserved hinge regions (Geng et al., 2024). While conditions such as lower temperatures or low-level ammonia elevation may lead to an increase in productivity, it is possible that the increased protein degradation involved in adaptation to these conditions may affect the recombinant product quality. Top-down mass spectrometry methods, along with PTM-based bottom-up proteomics, could be used to characterise any changes in product structure between culture conditions.

Advanced statistical modelling of protein expression changes

In chapter 5 of this thesis, investigation of machine learning methods demonstrates the potential for novel perspectives on large datasets using these methods. Given that Random Forest models showed the strongest predictive power, these models could be adapted to use other types of MS data, gene ontology data and possibly included some encoded viability/productivity data. With greater computing power, larger numbers of trees can be used on larger datasets, with the aim of characterising all aspects of the stress response.

High-function computer servers and robust Python/Bash coding would allow for the expansion of the deep data analysis. The huge amount of data acquired in omic studies such as the ones carried out in this thesis requires not only specific skills, but also vast memory and computing capacity. Using this increased power, a compiled dataset of all MZID files could be analysed at once, rather than the segmented study detailed in this thesis. This would increase the statistical power of the analysis performed and give greater insight into potential biomarkers found within these datasets.

References

Cain, K., Peters, S., Hailu, H., Sweeney, B., Stephens, P., Heads, J., Sarkar, K., Ventom, A., Page, C., Dickson, A., 2013. A CHO cell line engineered to express XBP1 and ERO1-Lα has increased levels of transient protein expression. Biotechnology Progress 29, 697–706. https://doi.org/10.1002/btpr.1693

Geng, S.-L., Zhao, X.-J., Zhang, X., Zhang, J.-H., Mi, C.-L., Wang, T.-Y., 2024. Recombinant therapeutic proteins degradation and overcoming strategies in CHO cells. Appl Microbiol Biotechnol 108, 182. https://doi.org/10.1007/s00253-024-13008-6

Jiang, Q., Sun, Y., Guo, Z., Fu, M., Wang, Q., Zhu, H., Lei, P., Shen, G., 2016. Overexpression of GRP78 enhances survival of CHO cells in response to serum deprivation and oxidative stress. Eng Life Sci 17, 107–116. https://doi.org/10.1002/elsc.201500152

Kaushik, P., Henry, M., Clynes, M., Meleady, P., 2018. The Expression Pattern of the Phosphoproteome Is Significantly Changed During the Growth Phases of Recombinant CHO Cell Culture. Biotechnology Journal 13, 1700221. https://doi.org/10.1002/biot.201700221

Kyeong, M., Lee, J.S., 2022. Endogenous BiP reporter system for simultaneous identification of ER stress and antibody production in Chinese hamster ovary cells. Metab Eng 72, 35–45. https://doi.org/10.1016/j.ymben.2022.02.002

Appendix I – Declarations of authorship

Declaration of Authorship for PhD by Publication

Section 1: Candidate's Details	
Candidate's Name	David Ryan
DCU Student Number	21261088
School	School of Biotechnology
Principal Supervisor	Dr Paula Meleady
Title of PhD by Publication Thesis	Proteomic Analysis of Endoplasmic Reticulum Stress in Chinese Hamster Ovary Cells in Culture
Section 2: Paper Details	
Title of co-authored paper included in the thesis under examination	Label-Free Quantitative Proteomics Analysis of Producer and Non-producer Chinese Hamsters Ovary (CHO) Cells under ER Stress Conditions
List of authors as in order given on accepted/published version	Ryan, David, Sideri, Christiana-Kondylo, Henry, Michael, Efeoglu, Esen, & Meleady, Paula. *D Ryan and CK Sideri are co-first authors.
Publication Details (e.g. Year of Publication, Journal Name, Issue Number, Page Numbers, Place of Publication, DOI, ISSN, URL etc.)	David Ryan, Christiana-Kondylo Sideri, Michael Henry, Esen Efeoglu, Paula Meleady, Current Research in Biotechnology, Volume 6, 2023, 100141, ISSN 2590-2628, https://doi.org/10.1016/j.crbiot.2023.100141.
Publication Status (e.g. Published, Accepted)	Published-22 August 2023. This paper was received 19 July 2023, revised 10 August 2023, accepted 11 August 2023.
Provide a brief statement here on the disciplinary ranking of this publication and the peer-review process involved (e.g. how many reviewers, single blinded or other, initial decision, revision timeline etc and any other details	Current Research in Biotechnology is ranked within the second quartile (Q2) in the field of Biotechnology, with an Impact Factor of 4.0. Anonymous (single-blind) peer review is employed, with two reviewers involved in review of this paper.

that speak to the rigour of the peer review process)	

This paper (Chapter 1) is one of **3** co-authored papers to be submitted as part of the PhD by publication thesis submitted for examination.

Section 3: Candidate's Contribution to the Paper

Provide details below of the **nature** and **extent of your contribution** to the paper (include both your intellectual and practical contributions) and your overall contribution in **percentage terms**. Note that these categories can be customised according to disciplinary norms

Along with C.-K. Sideri, I was a lead author and key contributor of this paper, involved in conception, data curation, investigation, methodology and writing the original draft. I conducted the cell culture work, western blotting and proteomic analysis of CHO-K1 and CHO-SK15-EPO, with the CHO-DP12 cell work performed by CK Sideri. I contributed extensively to the original draft of the manuscript, including abstract, methodology, results and discussion, as well as figure creation. Overall contribution: 50%.

Section 4: Signature and Validation

I confirm that the following statements are true:

the information I have provided in this form is correct

this paper is based on research undertaken during my candidature at DCU

Signature of PhD Candidate: David Ryan Date**:** 27th November2025

I confirm that the information provided by the candidate is correct:

Signature of Principal Supervisor: Paula Meleady **Date:** 27th November 2025

Declaration of Authorship for PhD by Publication

Section 1: Candidate's Details	
Candidate's Name	David Ryan
DCU Student Number	21261088
School	School of Biotechnology
Principal Supervisor	Dr Paula Meleady
Title of PhD by Publication Thesis	Proteomic Analysis of Endoplasmic Reticulum Stress in Chinese Hamster Ovary Cells in Culture
Section 2: Paper Details	
Title of co-authored paper included in the thesis under examination	Proteomic and Ubiquitinated Proteome Insights into ER Stress Responses in Chinese Hamster Ovary Cells Under Mild Hypothermic Conditions
List of authors as in order given on accepted/published version	Ryan, David, Sideri, Christiana-Kondylo, Henry, Michael, Karuppuchamy, Selvaprakash, Efeoglu, Esen, & Meleady, Paula.
Publication Details (e.g. Year of Publication, Journal Name, Issue Number, Page Numbers, Place of Publication, DOI, ISSN, URL etc.)	Ryan, D., Sideri, C.-K., Henry, M., Karuppuchamy, S., Efeoglu, E. and Meleady, P. (2025), Proteomic and Ubiquitinated Proteome Insights Into ER Stress Responses in Chinese Hamster Ovary Cells Under Mild Hypothermic Conditions. Biotechnology and Bioengineering. https://doi-org.dcu.idm.oclc.org/10.1002/bit.70081
Publication Status (e.g. Published, Accepted)	Published: 28 October 2025, manuscript accepted: 09 October 2025, manuscript revised: 19 August 2025, manuscript received: 28 May 2025
Provide a brief statement here on the disciplinary ranking of this publication and the peer-review process involved (e.g. how many reviewers, single blinded or other, initial decision, revision timeline etc and any other details that speak to the rigour of the peer review process)	Biotechnology and Bioengineering is a Q2 journal in biotechnology, with an impact factor of 3.6. It employs single blind peer review, with 2 reviewers reading this manuscript.

This paper (Chapter 2) is one of **3** co-authored papers to be submitted as part of the PhD by publication thesis submitted for examination.

Section 3: Candidate's Contribution to the Paper

Provide details below of the **nature** and **extent of your contribution** to the paper (include both your intellectual and practical contributions) and your overall contribution in **percentage terms**. Note that these categories can be customised according to disciplinary norms

I was the lead author and key contributor of the paper, leading conception, methodology, investigation, data curation and visualisation, writing-original draft. I conducted the cell culture work and proteomic analysis of CHO-K1 and CHO-SK15-EPO, with the CHO-DP12 cell work aided by CK Sideri, based on experiments of my own design. I performed ubiquitin enrichment study on all cell lines. I wrote the manuscript as well as creating all figures and tables. Overall: 80%.

Section 4: Signature and Validation

I confirm that the following statements are true:

the information I have provided in this form is correct

this paper is based on research undertaken during my candidature at DCU

Signature of PhD Candidate: ____David Ryan___ **Date:** 27th November 2025

I confirm that the information provided by the candidate is correct:

Signature of Principal Supervisor: _Paula Meleady__ **Date:** 27th November 2025

Declaration of Authorship for PhD by Publication

Section 1: Candidate's Details	
Candidate's Name	David Ryan
DCU Student Number	21261088
School	School of Biotechnology
Principal Supervisor	Paula Meleady
Title of PhD by Publication Thesis	Proteomic Analysis of Endoplasmic Reticulum Stress in Chinese Hamster Ovary Cells in Culture
Section 2: Paper Details	
Title of co-authored paper included in the thesis under examination	Proteomic Analysis of Ammonia-Induced Stress in Chinese Hamster Ovary (CHO) Cell Cultures
List of authors as in order given on accepted/published version	Ryan, David, Sideri, Christiana-Kondylo, Henry, Michael, Efeoglu, Esen, & Meleady, Paula.
Publication Details (e.g. Year of Publication, Journal Name, Issue Number, Page Numbers, Place of Publication, DOI, ISSN, URL etc.)	Journal of Proteomics, Volume 328, 2026, 105655, ISSN 1874-3919, https://doi.org/10.1016/j.jprot.2026.105655.
Publication Status (e.g. Published, Accepted)	Received 26 November 2025, Revised 27 March 2026, Accepted 31 March 2026, Available online 1 April 2026, Version of Record 8 April 2026.
Provide a brief statement here on the disciplinary ranking of this publication and the peer-review process involved (e.g. how many reviewers, single blinded or other, initial decision, revision timeline etc and any other details that speak to the rigour of the peer review process)	Biotechnology Journal is a Q2 journal in the field of biotechnology, with an impact score of 2.8. It employs a double-blind peer-review process.
This paper (Chapter 3) is one of **3** co-authored papers to be submitted as part of the PhD by publication thesis submitted for examination.	

Section 3: Candidate's Contribution to the Paper

Provide details below of the **nature** and **extent of your contribution** to the paper (include both your intellectual and practical contributions) and your overall contribution in **percentage terms**. Note that these categories can be customised according to disciplinary norms

I was the lead author and key contributor of the paper, leading conception, methodology, investigation, data curation and writing-original draft. I conducted the cell culture work and proteomic analysis of all three cell lines, based on experiments of my own design. I wrote the manuscript as well as creating all figures and tables. Overall: 85%.

Section 4: Signature and Validation

I confirm that the following statements are true:

the information I have provided in this form is correct

this paper is based on research undertaken during my candidature at DCU

Signature of PhD Candidate: David Ryan **Date:** 27th November 2025

I confirm that the information provided by the candidate is correct:

Signature of Principal Supervisor: Paula Meleady **Date:** 27th November 2025

Appendix I I– Chapter 4 supplementary materials

```
library(tidyverse)
library(caret)
library(randomForest)
library(ranger)
library(xgboost)
library(pROC)
library(mlr)
library(e1071)
library(class) #KNN
library(ggplot2)
library(glmnet)
#read in both stressed and unstressed files
mzid_files<-list.files(pattern = "\\.mzid$")
#Loop through each file
for (file in mzid_files) {
    mzid_data<-mzID(file)
     psm_data<-mzID::flatten(mzid_data)
     if (grepl("_ST", file)){
    psm_data$Stress<- "Y"
  }
  else {
   psm_data$Stress<- "N"
  }
  var_name<-gsub("\\.mzid$", "_db", file)
     assign(var_name, psm_data, envir = .GlobalEnv)
  }
##### Data frame structure #####
db_datasets<-ls(pattern = "_db$")
combined_db<-do.call(rbind, lapply(db_datasets, get))
combined_db[,c(1,4,10,11,12,13,17,22:23)]<-NULL
write.csv(combined_db, file="Ubiq_pep_DB.csv")
combined_clean<-read.csv("Ubiq_MZID_regress.DB.csv", header = T)
```

```
combined_clean<-combined_clean %>%
 select(-c(X,post, pre))
# Drop ID and redundant columns
combined_clean<-combined_db %>%
 select(-c(spectrumid, acquisitionnum, passthreshold, description, sequence, pepseq))
# Remove rows with NA/Inf
combined_clean[is.na(combined_clean)]<-0
# Target to factor
combined_clean$Stress<-ifelse(combined_clean$Stress=="Y", 1, 0)
combined_clean$Stress<-as.factor(combined_clean$Stress)
# Remove character variables not used in numeric modelling or convert appropriately
combined_clean<-combined_clean %>%
 mutate_if(is.character, as.factor)
colnames(combined_clean)[1]<-"retentiontime"
colnames(combined_clean)[7]<-"PDdeltascore"
write.csv(combined_clean, file="Ubiq_MZID_regress.DB.csv")
Acc_DF<-combined_clean[,c(1,7,10,14)]
Inc_pep_DF<-combined_db[,c(1,4,6,8,11,12,15)]
colnames(Inc_pep_DF)[1]<-"RetenTime"
###### target encode accession####
accession_enc<-Inc_pep_DF %>%
 group_by(accession) %>%
 summarise(mean_stress=mean(Stress=="Y"))
Inc_pep_DF<-Inc_pep_DF %>%
 left_join(accession_enc,by="accession") %>%
 rename(accession_score=mean_stress)

#### target encoding peptide sequence ####
pepseq_enc<-Inc_pep_DF %>%
 group_by(pepseq) %>%
 summarise(mean_stress=mean(Stress=="Y"))
Inc_pep_DF<-Inc_pep_DF %>%
 left_join(pepseq_enc,by="pepseq") %>%
```

```
 rename(pepseq_score=mean_stress)


# Partition data
set.seed(101) #or 110,123
train<-sample(nrow(Inc_pep_DF), 0.01*nrow(Inc_pep_DF), replace=FALSE)
test<-sample(nrow(Inc_pep_DF), 0.01*nrow(Inc_pep_DF), replace=FALSE)
TrainSet<-Inc_pep_DF[train,]
ValidSet<-Inc_pep_DF[test,]
TrainSet$Stress<-factor(TrainSet$Stress)
ValidSet$Stress<-factor(ValidSet$Stress)
#Logistic Regression -Repeated for Seed 101, 110, 123
###Penalised regression####
x <- model.matrix(Stress ~ . -1, data = TrainSet)
y <- as.numeric(TrainSet$Stress)
cvfit <- cv.glmnet(x, y, family = "binomial", alpha=.5)
plot(cvfit)
coef_min <- coef(cvfit, s = "lambda.min")

coef_df <- as.data.frame(as.matrix(coef_min))
coef_df$Feature <- rownames(coef_df)
colnames(coef_df)[1] <- "Coefficient"
# Filter non-zero features (excluding intercept)
significant_features<-subset(coef_df, Coefficient != 0 & Feature != "(Intercept)")
significant_features$OddsRatio <- exp(significant_features$Coefficient)
significant_features <- significant_features[, c("Feature", "Coefficient", "OddsRatio")]

#### Random Forest #### Repeated for Seed 101, 110, 123
set.seed(101)
train<-sample(nrow(Inc_pep_DF), 0.5*nrow(Inc_pep_DF), replace=FALSE)
test<-sample(nrow(Inc_pep_DF), 0.3*nrow(Inc_pep_DF), replace=FALSE)
TrainSet<-Inc_pep_DF[train,]
ValidSet<-Inc_pep_DF[test,]
TrainSet$Stress<-factor(TrainSet$Stress)
```

```
ValidSet$Stress<-factor(ValidSet$Stress)
rf_model_101<-ranger(
 Stress~.,
 data=TrainSet,
 probability=TRUE,
 importance="impurity",
 num.trees=500
)
rf_pred<-predict(rf_model_101,ValidSet)$predictions[,2]
rf_auc<-pROC::auc(ValidSet$Stress,rf_pred)
cat("RF AUC:",rf_auc,"\n")
rf_importance<-as.data.frame(rf_model_101$variable.importance)

impact_accession<-data.frame(accession=ValidSet$acc_code,
                  pred=predict(rf_model_101,ValidSet)$predictions[,2]) %>%
 group_by(accession) %>%
 summarise(mean_pred=mean(pred),n=n()) %>%
 arrange(desc(mean_pred))
impact_pep_seq<-data.frame(pep_seq=ValidSet$pep_code,
                  pred=predict(rf_model_101,ValidSet)$predictions[,2]) %>%
 group_by(pep_seq) %>%
 summarise(mean_pred=mean(pred),n=n()) %>%
 arrange(desc(mean_pred))
sign_acc_RF101<-impact_accession %>%
 filter(mean_pred>=0.9, n>2)
sign_pep_RF101<-impact_pep_seq %>%
 filter(mean_pred>=0.9, n>2)

#### XGBoost ##### Repeated for Seed 101, 110, 123
train_matrix<-model.matrix(Stress ~ . -1, data=TrainSet)
train_label<-as.numeric(TrainSet$Stress) - 1
test_matrix<-model.matrix(Stress ~ . -1, data=ValidSet)
test_label<-as.numeric(ValidSet$Stress) - 1
```

```
dtrain<-xgb.DMatrix(data=train_matrix, label=train_label)
dtest<-xgb.DMatrix(data=test_matrix, label=test_label)
xgb_params<-list(objective="binary:logistic", eval_metric="auc")
xgb_model<-xgb.train(params=xgb_params, data=dtrain, nrounds=100)
pred_xgb<-predict(xgb_model, newdata=test)
roc_xgb<-roc(test_label, pred_xgb)
auc(roc_xgb)
xgb_importance<-xgb.importance(model=xgb_model)

######### SVM #######
SVMTrainSet<-TrainSet[,-c(4,5)]
SVMValidSet<-ValidSet[,-c(4,5)]
SVMTrainSet$Stress<-factor(SVMTrainSet$Stress,levels=c("N","Y"))
svm_model<-svm(
 Stress~.,
 data=SVMTrainSet,
 kernel="radial",
 probability=TRUE,
 scale=TRUE
)
svm_pred<-attr(predict(svm_model,ValidSet,probability=TRUE),"probabilities")["Y"]
svm_auc<-pROC::auc(ValidSet$Stress,svm_pred)
cat("SVM AUC:",svm_auc,"\n")
pred_class_svm<-ifelse(svm_pred>0.5,1,0)
SVMValidSet$pred_prob<-svm_pred

# Summarise by pepseq
pepseq_importance<-SVMValidSet %>%
 group_by(pepseq) %>%
 summarise(mean_pred=mean(pred_prob),
        n=n(),
        stress_rate=mean(Stress=="Y")) %>%
 arrange(desc(mean_pred))
```

```
head(pepseq_importance,20)
####### KNN #######
scale_cols<-TrainSet %>% select(-Stress,-accession,-pepseq) %>% names()
preProc<-preProcess(TrainSet[,scale_cols],method=c("center","scale"))
train_scaled<-predict(preProc,TrainSet[,scale_cols])
test_scaled<-predict(preProc,ValidSet[,scale_cols])
knn_pred<-knn(
 train=train_scaled,
 test=test_scaled,
 cl=TrainSet$Stress,
 k=5,
 prob=TRUE
)
knn_prob<-ifelse(knn_pred=="Y",attr(knn_pred,"prob"),1-attr(knn_pred,"prob"))
knn_auc<-pROC::auc(ValidSet$Stress,knn_prob)
cat("KNN AUC:",knn_auc,"\n")
knn_results<-data.frame(
 accession=ValidSet$accession,
 pepseq=ValidSet$pepseq,
 stress=ValidSet$Stress,
 pred_prob=knn_prob
)

acc_summary<-knn_results %>%
 group_by(accession) %>%
 summarise(
  mean_pred=mean(pred_prob),
  stress_rate=mean(as.numeric(stress=="Y")),
  n=n()
 ) %>%
 arrange(desc(mean_pred))
pep_summary<-knn_results %>%
 group_by(pepseq) %>%
```

```
  summarise(
    mean_pred=mean(pred_prob),
    stress_rate=mean(as.numeric(stress=="Y")),
    n=n()
  ) %>%
  arrange(desc(mean_pred))
pred_class_svm<-ifelse(svm_pred>0.5,1,0)
KnnValid<-ValidSet
KnnValid$pred_prob<-knn_pred
```

Log Regre					SVM			
Seed 101					Seed 101			
Coefficient	Odds Ratio				accession	mean_pred	n	
0.050548	1.051848	G3H4J1			2678	0.706987	3	G3GYA2
0.106081	1.111912	G3HB62			4033	0.689055	3	G3GX55
0.213612	1.238142	G3HQB7			2064	0.684093	3	G3HES6
0.230741	1.259532	G3IL48			2180	0.681916	3	G3HJG6
-0.175	0.839457	G3HGW6			3601	0.67787	4	G3HC04
-0.12598	0.881634	G3I4H5			1530	0.677264	4	G3HAI3
0.731091	2.077345	G3H7F1			1480	0.674653	4	G3I518
-0.02255	0.977706	G3HN02			2310	0.670922	3	G3GYR1
0.016966	1.017111	G3GSY8			1208	0.668586	3	Q9NPD3
-0.01	0.99	G3IH63			1242	0.666632	4	G3HCE7
0.13	1.14	G3I973			2338	0.666584	4	G3GT65
-0.05	0.95	G3HKG8			2220	0.6661	3	G3IKI0
-0.02	0.98	G3GWG3			1955	0.665309	3	G3IK12
0.151275	1.163316	AWELYGSANAVVLLIAQEK			1161	0.665306	4	G3IF44
-0.23689	0.78908	LIALLEVLSQK			2579	0.662923	3	G3GR68
0.050295	1.051582	DNHLLGTFDLTGIPPAPR			2482	0.662666	6	G3H8G1
-0.01736	0.982789	LSVLGAITSVQQR			2573	0.659897	5	G3GWY3
0.164034	1.178254	AYHEQLSVAEITNACFEPANQMVK			2142	0.65856	3	G3HKP3

-0.04856	0.952596	FQSLGVAFYR			1433	0.657409	5	G3GWX0
-0.02699	0.973376	VDSPVTVCGDIHGQFYDLK			1413	0.656132	5	G3IAI6
-0.15307	0.858074	VIHDNFGIVEGLMTTVHAITATQK			1675	0.656049	5	G3HI96
0.180455	1.197763	SVEGWILFVTGVHEEATEEDIHDK			2266	0.655989	3	G3HV67
0.105425	1.111183	RVIISAPSADAPMFVMGVNQDK			1193	0.655849	3	G3HV17
0.071655	1.074285	DGWSLWYAEYR			2621	0.65569	3	G3H8Z6
0.128912	1.13759	LYQEVEIASVDAFQGR			3253	0.654722	3	G3I5D4
0.105481	1.111244	KSDIYVCMISFAHNVAAQGK			1250	0.653802	6	G3GZG5
-0.42338	0.65483	DLVSATALVTAEMEGTQQLR			1862	0.653636	6	G3IJF3
0.346768	1.414489	VEPGSTCAVFGLGGVGLAVIMGCK			2012	0.652913	4	G3GRU3
0.062461	1.064453	GETVLIHSGSGGVGQAAISIALSLGCR			2738	0.652529	3	G3HSU8
Seed 110					2538	0.65251	15	G3I4A0
Feature					2650	0.651931	6	Q13509
Coefficient	Odds Ratio				2066	0.651823	8	G3HAC4
0.243759	1.276037	G3H1D5			2438	0.651186	5	G3HU20
-0.07391	0.928757	G3HZV0			2112	0.650963	3	G3HLP4
0.199763	1.221113	Q9BVA1			2185	0.65096	3	G3HP82
0.053087	1.054522	G3IGZ8			3583	0.650937	3	G3I9H0
-0.05469	0.946775	G3HBD3			4893	0.650927	3	G3I3X4
0.232784	1.262109	G3HTF7			2043	0.650838	3	G3I2P8
0.57	1.77	G3H9Z1			1395	0.650752	5	G3HHU4
0.51	1.66	G3I306			2048	0.650126	21	G3HSM4
0.31	1.37	G3HYU7			pep_seq	mean_pred	n	
0.25	1.29	G3I8R9			6280	0.718112	3	YYYECVLFVVR
0.24	1.27	Q9BVA1			8057	0.706051	3	DVLLPWYNAYRFFVQNVFR
0.22	1.25	G3HYA9			12932	0.710867	3	DAVSGMGVIVHIIEK
0.15	1.16	G3H412			Seed 110			
0.14	1.15	G3HRM8			accession	mean_pred	n	
-0.14	0.87	G3HPT8			409	0.70447	4	G3HLU5
-0.17	0.84	G3I3K5			807	0.708395	3	G3HV90
-0.19	0.83	G3GU92			1013	0.702814	3	G3GU41

-0.26	0.77	G3HID3			1088	0.716608	3	G3GSC9
-0.29	0.75	G3HBD3			1756	0.763815	4	G3HHH6
-0.37	0.69	G3GUP7			1919	0.713778	5	G3HYC4
-0.43	0.65	G3I973			1974	0.704282	5	G3IJJ7
-0.52	0.59	G3GY17			2062	0.70165	7	G3HWI4
0.292339	1.339557	VLILGSGGLSIGQAGEFDYSGSQAIK			2149	0.722294	3	G3H195
0.313089	1.367644	LAQSNGWGVMVSHR			2248	0.727696	5	G3IET4
-0.04713	0.953963	SSFLVLCYLHDVK			2339	0.726429	3	G3IAL6
0.026069	1.026412	FEDGDLTLYQSNAILR			2593	0.700281	3	G3GSE2
0.13	1.14	AIMTYVSSFYHAFSGAQK			2716	0.712266	4	G3HS03
-0.14	0.87	AVFPSIVGRPR			2739	0.710706	3	G3I011
0.16	1.17	GMADIDIQDDK			3200	0.72374	10	G3H0Q2
0.54	1.71	DSTLIMQLLR			pep_seq	mean_pred	n	
Seed 123					7033	0.793198	3	HMTLVGIISWGIGCGQK
Coefficient	Odds Ratio				6536	0.756815	3	SSDQALSYEYITLSETK
0.42	1.52	G3HZV0			2310	0.746665	3	FIYMAVQHYIETLQR
0.2	1.22	G3HRM9			12278	0.736444	3	VHNDAQNFDYDHDAFLGAEEAK
0.2	1.22	G3GYY6			6944	0.726429	3	SLHTLFGDK
0.1	1.1	G3HAF4			1609	0.725492	5	EGSRTDDEVVQREEEAIQLDGLNASQIR
0.1	1.1	G3I905			2128	0.725379	4	LVGQGATAVLLDLPNSGGEAQAKK
0.01	1.01	G3HGV7			10733	0.72374	10	GGASDALLYR
-0.07	0.94	P55072			764	0.722719	3	AGLPCQDLEFVQFHPTGIYGAGCLITEGCR
-0.15	0.86	G3HVP2			5972	0.722294	3	IAAQGFTIAALLLGLAVSAMK
-0.18	0.84	G3GVM9			7991	0.718861	3	QITQVYGFYDECLRK
-0.21	0.81	G3I2B6			1644	0.717411	4	FNEEHIPDSPFVVPVASPSGDAR
-0.25	0.78	G3HES3			15888	0.716004	4	ATIELKALR
-0.32	0.73	G3HXY1			3350	0.715993	10	RPLWFICSGMGTQWR
-0.34	0.71	G3IBG3			2169	0.715756	12	NVFIIGATNRPDIIDPAILRPGR
-0.71	0.49	G3HKY0			638	0.71429	3	NICQFLIEIGLAK
-0.82	0.44	G3GWG3			9821	0.714014	4	LGLVKYGEK
0.28	1.33	ENHGIFFSGDSYLVLHNGPEEVSHLHLWIGQQSSR			15360	0.713958	3	SQEMVHLVNKEPSETPDQFMTADETR
0.16	1.18	ELEEIVQPIISK			2089	0.713644	7	RGYSLVSGGTDTHLVLVDLRPK
0	1	GMAAAGNYAWVNR			8682	0.712266	4	QSDENNLKDPGGSGFDSISK
-0.09	0.91	LQAVEVVITHLAPGTK			15824	0.710427	3	SIAQTTPVAGSQSLSSGTVAK

-0.18	0.84	VHIGQVIMSIR			4696	0.710234	3	KLDVSIEMPEELDISQLR
Random Forest					4191	0.709405	5	QETEVELYNEFPEPIKLDK
Seed 101					4720	0.709352	3	IEWLESHQDADIEDFKAK
accession	mean_pred	n			3490	0.709276	6	LIVAGTSCYSR
3429	0.999269	5	G3H3F8		7775	0.709122	3	TLVLLDNLNVR
2515	0.998611	3	G3GU91		4974	0.709046	3	YVATLGVEVHPLVFHTNRGPIK
408	0.99605	18	G3I180		6612	0.708996	3	SAEFLLHMLK
4892	0.994995	3	G3H215		5284	0.708891	3	FDYVFAENGTVQYK
19	0.99399	3	G3HZE5		5559	0.70888	7	RLAYAIIQFLHGQLR
4252	0.993285	3	G3HR78		4556	0.708681	3	ANEPTHFTVDCTEAGEGDVSVGIK
3642	0.992404	3	G3H4F5		3311	0.70806	6	GLVETAAHLGYSVVAINHIVDFK
3752	0.988395	4	G3HG79		15963	0.706837	6	GKLEAIITPPPAK
2621	0.987659	15	G3H8Z6		11196	0.706694	10	SNLKIAELK
3612	0.985772	3	G3HN43		4314	0.706489	3	QGAHIIVTPEDGIYGVQFTR
2520	0.974662	3	G3HM63		7617	0.705802	6	LNLYELKEGR
4294	0.970474	3	G3IKK6		15986	0.70572	5	KQLADETLLK
3876	0.96952	9	G3HXF7		16874	0.70562	3	FNVLKVTK
2641	0.969123	3	G3HD22		2080	0.705273	4	HPTFTQFSFMVYNSLR
2538	0.959422	48	G3I4A0		6172	0.704716	5	VQGGALEESQLVAGVAFKK
1546	0.95941	8	G3HNW8		2119	0.703743	14	ATEHADGGVAAGVAVLDNPYPV
4893	0.959232	6	G3I3X4		12672	0.703549	4	EAVEAIQNSTSIK
3835	0.95918	3	G3I9A7		762	0.703379	3	DLVSATALVTAEMEGTQQLR
2574	0.95636	3	G3HLK9		249	0.702835	3	GVQVLACLSVNDVFVTEEWGR
1653	0.954499	10	G3INX6		2373	0.701704	8	AHLQGPTGASVAGPQTSEAFAITVR
3626	0.948448	3	G3I1G2		8149	0.701315	3	YADTLFDILVAGSMLAPGGTR
2185	0.945899	11	G3HP82		893	0.7013	8	CIYCIHALSLYLFK
3200	0.942906	16	G3H0Q2		14910	0.701244	7	TFSFAIPLIEK
2678	0.941607	3	G3GYA2		7806	0.701069	9	DIVVAETLEDLDKNK
2216	0.939956	3	G3HB55		14863	0.700672	3	TLKLTTPTYGDLNHLVSATMSGVTTCLR
2724	0.939849	3	G3HF44		4049	0.70062	6	GVTQYYAYVTER
1504	0.933005	4	G3HW66		4479	0.700311	3	SQSSNDTFPTAMHIAAAVEVHEVLLPGLQK
4622	0.929785	3	G3H740		5054	0.700159	8	LPIVSSVACGASVGYAVSK
409	0.929708	9	G3HLU5		Seed 123			

2441	0.927137	3	G3HZ57		accession	mean_pred	n	
328	0.92635	15	G3HN14		3200	0.69016	10	G3H0Q2
4364	0.918371	4	G3HHK3		286	0.687394	3	G3HC71
1256	0.915706	3	G3I4C0		3429	0.684175	4	G3H3F8
1702	0.91196	3	G3IFS3		4258	0.681001	4	G3H3E7
3631	0.909807	6	G3I6L2		538	0.680968	3	G3HLW7
4422	0.909779	4	G3IFR3		2791	0.680706	3	G3HPK0
1992	0.907673	4	G3IBN9		408	0.671272	3	G3I180
2248	0.903992	7	G3IET4		1169	0.669631	9	G3IKW4
pep_seq	mean_pred	n			2621	0.6687	5	G3H8Z6
14910	0.999649	10	TFSFAIPLIEK		1643	0.665896	5	G3HN49
11649	0.999269	5	TVIVAALDGTFQR		807	0.662448	3	G3HV90
15719	0.999232	3	AEELKTELLK		927	0.661253	11	G3HDS3
18541	0.999076	3	KINELTGIK		2871	0.660926	3	Q99PC2
8240	0.998826	12	EQNPYEKF		2650	0.657683	11	Q13509
7752	0.998611	3	GVGNATACVLHEVIYVIGGHCGYR		1298	0.657629	5	G3I9W6
15827	0.998337	4	GHYTEGAELVDSVLDVIR		1878	0.65722	6	G3HRA3
7225	0.998325	3	ECVNFASFNFLGLLANPR		2248	0.656516	3	G3IET4
15020	0.997851	9	QQLLEDLK		2277	0.654604	13	Q60456
14854	0.997145	3	EILVGDVGQTVDDPYTTFVKMLPDK		2206	0.653943	3	G3IHT0
3832	0.995133	4	AQLVVIAHDVDPIELVVFLPALCRK		1346	0.653494	3	G3HCY7
18952	0.995007	6	IANVSILK		2185	0.652791	3	G3HP82
19699	0.994386	3	ETGVDLTK		1069	0.652677	3	G3GSJ5
6157	0.994288	3	YTMGDAPDFDR		2266	0.652399	5	G3HV67
12600	0.993582	4	VVLDVGSGTGILCMFAAK		2538	0.65184	18	G3I4A0
13410	0.993448	3	TCQALAAHGFTELSTLEVLPQVYNVR		1391	0.651832	6	G3HM98
6153	0.993325	3	KFDQLLAEEK		645	0.650435	6	G3I7X5
878	0.993267	5	AVVYSNTIQSIMAIVK		1395	0.65015	9	G3HHU4
12887	0.992404	3	FIQVNYLESGIVR		3876	0.650099	5	G3HXF7
15360	0.991657	3	SQEMVHLVNKEPSETPDQFMTADETR		pep_seq	mean_pred	n	
2416	0.991468	4	EVMIVGIGDK		10733	0.69016	10	GGASDALLYR
8339	0.991381	3	IAQDLEMYGINYFEIKNK		433	0.687394	3	LFMILWLK
5745	0.991329	4	FDDGAGGDNEVQR		2035	0.686754	4	ACLISLGYDVENDRQGDAEFNR
3329	0.991153	4	FNGGGHINHSIFWTNLSPK		11649	0.68396	3	TVIVAALDGTFQR

1851	0.991082	3	LRASILNTWISLK		6064	0.682893	3	EVLADRPLFPHVLCHNCAVEFNFGQK
4772	0.990982	3	SWIEEQEMGSFLSVAK		9860	0.68195	3	YFQISLEAEK
12337	0.990827	3	TLIDQQVVYIHPSSALFNR		16131	0.681265	4	DVLQAAWENEQDIIEKK
9290	0.990701	3	MFIGGLSWDTSK		15010	0.681001	4	FANFSGGKITK
2310	0.990651	6	FIYMAVQHYIETLQR		688	0.680861	5	GFLLWYSGR
2018	0.990638	5	AIYEVLFR		5139	0.679975	6	GSLLGCSIDITSAADSEAVTYQK
12295	0.99057	3	AAQLCGAGMAAVVEK		1737	0.679387	4	GISNMLDVNGLFSLAHITQLVLSHNK
19601	0.990226	3	GELLEAIK		12329	0.679312	4	GAMGIILVYDITDEK
3964	0.989865	8	RPWILTSLHNGVIQLWDYR		2373	0.679259	3	AHLQGPTGASVAGPQTSEAFAITVR
12278	0.989581	3	VHNDAQNFDYDHDAFLGAEEAK		620	0.679031	15	TISVILFLNK
12241	0.989511	3	AQMQEAMTQEVSDVFSDTTTPIK		845	0.67782	10	LTGMAFRVPTPNVSVVDLTCR
3785	0.989293	3	KADIGVAMGIVGSDVSK		857	0.67704	7	LGTLSALDILIK
6536	0.988326	5	SSDQALSYEYITLSETK		4556	0.676887	3	ANEPTHFTVDCTEAGEGDVSVGIK
15278	0.987331	3	IKIIAPPER		4731	0.675288	6	AAVGMAEEHCGFVVGFISGSR
6591	0.986485	4	FSTETTFLVDKYEIL		5109	0.675195	3	KAAAPAPEEEMDECEQALAAEPK
12521	0.986477	3	VTWDSSFCAVNPR		2872	0.674907	3	DALNQATSQVESK
3813	0.986087	3	GRFYAFGRVFSGVVSTGLK		6185	0.674347	3	RNDSIIVDIFHGLFK
16214	0.985315	3	KMAAGHELQPLAIVDQRPSSR		3736	0.67422	3	TYALFIGDTVLVDEDGPATVLTSVK
8390	0.984837	3	LAPITNDPTEAAAVGAVEASFKCCSGAIIVLTKSGR		4191	0.673655	3	QETEVELYNEFPEPIKLDK
227	0.98393	3	SADTLWGIQKELQF		2008	0.672684	3	DLPITEAVFSALVTGHAR
8159	0.983486	4	ERVETLSQVEVILQQSAADIAR		2728	0.672393	7	IYFLLVR
5384	0.982991	6	RIVGCICEEDNCTVIWFWLHK		3324	0.672232	3	IREDLPNLESSEETEQINK
12284	0.98273	3	LVGEEGFVVTEAGFGADIGMEK		4255	0.672157	7	NLEGYVGFANLPNQVYR
7530	0.978919	5	CVVTGEDGTQSEATVNVK		5340	0.671698	4	EALHPDTFIQLALQLAYYR
7991	0.977923	3	QITQVYGFYDECLRK		5317	0.671465	6	DTLIHHGVHHGVNQAGSEAGR
4336	0.975221	4	QVICLSGDDSSTCIGILAK		5704	0.671226	3	TLLILFKPLISFK
14886	0.974932	3	MATVKQGHRPIAPPPYCSKEEVGGPGCSTFCPEVWR		15888	0.670684	3	ATIELKALR
6969	0.974141	8	SLGLYGKDQR		3468	0.670338	3	EYYVGAAHGLSGIYYYLMQPSLQVSQAK
638	0.972007	6	NICQFLIEIGLAK		2922	0.669922	18	HLIVLINK
2012	0.971475	4	RSSFICVILSHGEEGVIFGTDGPVDLK		3365	0.669746	4	MREIVHLQAGQCGNQIGAK
8278	0.971011	5	LPLEGNTMK		2528	0.669631	9	LFIWVLR
8831	0.971006	3	ADMVIEAVFEDLTVK		3964	0.669451	3	RPWILTSLHNGVIQLWDYR
12715	0.970615	4	TYNIVLGSGQVVLGMDMGLR		3350	0.669338	3	RPLWFICSGMGTQWR
5183	0.970054	3	RIALTDNSLIAR		8240	0.6687	5	EQNPYEKF

7806	0.968339	18	DIVVAETLEDLDKNK		2119	0.668658	11	ATEHADGGVAAGVAVLDNPYPV
3773	0.967987	5	GGVSAVAGGVTAVGSAVVNK		4563	0.66853	7	GFHLVEVTYDDVPIPNSPFK
4815	0.96697	6	WWQLPYYWVGIK		1461	0.668203	5	ALGYFAVVTGK
6885	0.965617	3	GHYTEGAELVDSVLDVIRK		1051	0.668185	5	NVVGDHGDEVYSVFGAPILR
12741	0.963808	4	NVNIFKFIIPNVVK		4761	0.667523	3	LMVIFSK
6075	0.962501	3	IVSPSGAAVSCK		3140	0.667414	3	TLGECGFTSQTARPQAPATVGLAFR
8026	0.962005	4	SGDAPLTVNELGTAYVSATTGAVATALGLNALTK		5559	0.667179	5	RLAYAIIQFLHGQLR
7218	0.961268	4	YSVDGTLLAVGSHDNFIYLYSVSENAR		6058	0.666775	4	NENTFLDLTVQQIEHLNK
12554	0.961156	4	QQIQSIQQSIER		5248	0.666411	3	NSWGSSWGMGGFFHIAK
3220	0.961152	6	IGAFSYGSGLAASFFSFR		14910	0.666395	4	TFSFAIPLIEK
7047	0.959914	5	DDFFIKIETWHKPDLGTLENVHGLDPNTWK		1330	0.666338	4	DNYVPEVSALDQEIIEVDPDTK
12700	0.959719	8	TFVQEDVYDEFVER		5393	0.666315	4	GFCFITYTDEEPVK
3040	0.959324	3	EYYFAITMER		1779	0.666299	7	KYVIYIER
4961	0.955558	3	LFIVDARPAMAATSFVLMTTFPNK		7300	0.66626	3	VIFKPTVVQQAK
2327	0.953819	18	GLAEVAVK		7806	0.666169	4	DIVVAETLEDLDKNK
10	0.950546	6	GAGSYTIMVLFADQATPTSPIR		1977	0.66615	3	FFHLAFEEEFGR
12354	0.950389	4	LEQQVPVNQVFGQDEMIDVIGVTK		2622	0.666137	6	AAAASAAEAGIATSGTEGERDTDDALLKMTISQQEFGR
7428	0.94857	3	HQIVEVAGDDKYGR		893	0.665953	13	CIYCIHALSLYLFK
12430	0.948448	3	NFAVWLR		933	0.665927	3	LHDVNNDGFLDEQELEALFTR
1975	0.948243	3	DPTVNVWYSVNGER		7166	0.665759	4	KTQEQLASEMAELTSR
15489	0.947949	3	ADFVLAANSYDLAIK		4987	0.665584	3	NNEDISVIPPLFTVSVDHR
447	0.94737	5	NVFDEAILAALEPPEPKK		2128	0.665566	3	LVGQGATAVLLDLPNSGGEAQAKK
5379	0.946614	3	IDLRPVLGEGLPILASFLR		4164	0.665278	3	IPVTDEEQTNVPYIYAIGDILEGK
5474	0.946603	3	LPITVLNGAPGFINLCDALNAWQLVK		3742	0.665224	11	VAVLISGTGSNLQALIDSTK
3774	0.945115	3	GAVGATIGGVAWIGGK		12417	0.664964	3	FNACFESVATNIDEIYK
15963	0.945103	9	GKLEAIITPPPAK		6198	0.664455	3	QVMMFSATLSK
3120	0.945009	8	WSLALGLPLPLGEHAGR		11285	0.664337	3	ISATSIFFESMPYK
5559	0.944556	13	RLAYAIIQFLHGQLR		290	0.664326	4	DGILSDEIYCPPETAVLLGSYAVQAK
10733	0.942906	16	GGASDALLYR		3239	0.663947	5	MRECISIHVGQAGVQIGNACWELYCLEHGIQPDGQMPSDK
7555	0.941914	16	VTVLFAGQHIAK		6738	0.66385	4	EAALTSEEVGADLEQVEVLQK
6371	0.939956	3	ALAESWNAAFLESSAK		1435	0.663656	3	NLDYVAASIHEAVTK
8706	0.939849	3	AELVVIEDAATPR		8870	0.663553	6	EVGKVGEWAIR
15881	0.937996	7	SAFPFSDKLGELSDK		2130	0.663476	7	GVIINTASVAAFDGQVGQAAYSASK
1711	0.934164	4	SLDFLIELLHK		4172	0.663343	4	AFVEFLTDEIKEEK

4731	0.934039	8	AAVGMAEEHCGFVVGFISGSR		7257	0.663151	8	VLDLIEVLVTK
12451	0.933528	3	AVLTNILLR		6527	0.66289	3	LFIGGLNTETNEKALEAVFGK
6618	0.932833	3	LVGPEVR		7074	0.662726	3	VVMALGDYMGASCHACIGGTNVR
2564	0.932312	6	AEVEEEVEEQK		3598	0.662589	3	KGSITSVQAIYVPADDLTDPAPATTFAHLDATTVLSR
12499	0.931026	3	QLFVLAGSAEEGVMTSELAGVIK		12690	0.662546	3	EVIIMATNCENCGHR
16557	0.930629	3	DQLIVNLLKEEQTPQNK		1359	0.662448	3	TLVLIGASGVGR
16239	0.929321	8	EDLRDTFAAMGR		2516	0.662121	5	GIQDVYVLSEQQGLLLR
12121	0.929035	3	TFFSFPAVVAPFK		15341	0.661898	3	TPVLFDVYEIK
12417	0.928492	3	FNACFESVATNIDEIYK		3954	0.661798	31	VEIIANDQGNR
15986	0.928257	16	KQLADETLLK		3175	0.661642	9	DSHFAIFIIPQGR
8756	0.927201	3	LSQTSNVDKEEEAVTIEMNEPVQLTFALR		654	0.661559	5	TNVANFPGHSGPITSIAFSENGYYLATAADDSSVK
5739	0.927069	3	FPWLFSFGYFSK		4442	0.66133	3	TNDQMVVVYLASLIRSVVALHNLINNK
8775	0.927042	4	ENKPSIIFIDEIDSLCGSR		5320	0.661313	5	VAQGVYTGVNQAEK
3947	0.926735	3	RVDTWFNQPAR		6606	0.661256	3	DYEEVGVDSVEGEGEEEGEEY
6407	0.924658	4	DAEVVLCGGTESMSQAPYCVR		3828	0.661182	6	TPFLLVGTQIDLRDDPSTIEK
9821	0.924211	10	LGLVKYGEK		12700	0.661127	7	TFVQEDVYDEFVER
11013	0.922646	5	LHIVKVVAK		9339	0.660926	3	YWLYYLHPVLQR
8172	0.921909	3	FYFVGDEDLLEIIGNSK		1609	0.660625	3	EGSRTDDEVVQREEEAIQLDGLNASQIR
12342	0.921669	3	ALLEVVQSGGK		5054	0.660567	8	LPIVSSVACGASVGYAVSK
8823	0.920507	3	MKETAEAYLGK		2254	0.66054	4	MLAAYLYEVSQLK
15397	0.919867	4	GLANMFSVFTK		3669	0.660495	5	ILSEAEIDAHLVALAERD
1621	0.919154	4	LNTLLQR		2203	0.660075	5	FGVEQGVDMVFASFIRK
12428	0.918345	14	IPVDTYNNILTVLK		4758	0.659984	4	DSLTLVFVETKK
1435	0.918177	9	NLDYVAASIHEAVTK		886	0.659431	7	DFVVGALILR
14901	0.917017	11	LSLVSLLVR		5055	0.659346	4	VFLWGSFR
15814	0.915293	3	FYEAFSKNLK		4393	0.659155	7	EIEIDIEPTDKVER
19731	0.914788	3	GDNVAVIGEIDEETDSALDLGNIRAEPLNSVAH		7100	0.659106	6	FEELNMDLFR
12299	0.914429	3	DRINIVLSR		2499	0.658737	4	FFGILYENCYHILGK
3670	0.914125	3	SQPISSPVILQFGHAETLLPLLSLMGYFK		4590	0.65873	3	LNSNTQVVLLSATMPSDVLEVTKK
12164	0.913804	3	LAHGGQVNLDMEDHRDEDFVKPK		14918	0.658672	3	LPSGVFSSDFQEFVNK
8178	0.912621	3	FAGDKDYLTKEDLR		7858	0.658528	4	DPFAPSSSAKPPKASSSGFADFTSFGNEEQQLAWAK
4258	0.91196	3	KVLGSSTFVTNSSSCSSR		151	0.658459	9	LIALLEVLSQK
15575	0.909807	6	LEIGVQSVYEDVAR		6649	0.658077	6	QTQLNLSTLLVNPEGPTLTR
16120	0.909779	4	FIGDQSCTDILDTYNKIK		762	0.657993	6	DLVSATALVTAEMEGTQQLR

2515	0.909606	5	LTYTAEVSVPK		5983	0.657984	7	FYEAELEQLDFQGDTEQSR
12386	0.906483	3	AVELLVELASLQTSFVTLDEAIK		2970	0.657565	4	LAFIAHPK
16874	0.906149	5	FNVLKVTK		1759	0.657499	3	LLEAQACTGGIIDPSTGERFPVTEAVNK
12762	0.904907	3	YWGYPSEEHMIQTEDGYILGVHR		3493	0.657413	7	GLSEDVSISK
10582	0.904754	6	ETLVYLTHLDYVDTEIIMTK		2632	0.657258	16	QLEDGRTLSDYNIQKESTLHLVLR
175	0.904449	6	VIHDNFGIVEGLMTTVHAITATQKTVDGPSGK		4880	0.657251	3	AENGKLVINGK
526	0.903345	3	FALNYLQLLVHPNPPSVLFGADK		4262	0.657189	4	VDSPVTVCGDIHGQFYDLK
14947	0.901541	3	SQEQLAAELAEYTAK		2200	0.657036	7	FDEILEASDGIMVAR
6269	0.901524	3	HGIESTLER		3484	0.656871	4	DPNKLVFCEVFK
6695	0.901434	4	KHPDASVNFSEFSKK		2491	0.656853	3	ELEDLEGYQNTIVAGGLITK
Seed110					2722	0.656797	3	SLTNDWEEHLAVK
pep_seq	mean_pred	n			3364	0.656675	4	MREIVHIQAGQCGNQIGAK
8240	0.99995	13	EQNPYEKF		3287	0.656504	8	SSGSVYITLK
18541	0.999364	3	KINELTGIK		1715	0.656359	5	VIEVSHWGNIAVEENVDLK
15020	0.999013	5	QQLLEDLK		583	0.656176	5	EVFEDALEIR
19592	0.998549	3	GVDIVMDPLGGSDTAK		249	0.656128	3	GVQVLACLSVNDVFVTEEWGR
14910	0.997973	13	TFSFAIPLIEK		6319	0.655771	6	ATSPSTLVSTGPISR
6376	0.997839	3	ISIEMHGTLEDQLSHLR		667	0.655625	6	LIGQIVSSITASLR
18959	0.997607	3	MAEVEAPTAAETDMK		4512	0.655569	3	TYPRPHEYLSPSDIPK
12600	0.997328	3	VVLDVGSGTGILCMFAAK		4941	0.655506	5	LAFYLCGR
6157	0.996613	4	YTMGDAPDFDR		XGBOOST			
7475	0.996431	3	ILHEDGSPNLDFKPEDQPHFDIKDEF		Seed101			
15346	0.996235	3	SIHLFIDSLLNEENPSK		acc_code	mean_pred	n	
7170	0.995726	3	GSGMENPVAVTVFFGANDATLK		1196	0.999979	3	G3H8G0
4858	0.994911	7	VNVNVDYIR		4886	0.999966	3	G3HYN9
18952	0.994333	7	IANVSILK		2185	0.999929	9	G3HP82
19601	0.994272	3	GELLEAIK		430	0.9998	3	G3HN89
8450	0.993808	3	SGEEDFESLASQFSDCSSAKAR		2567	0.999799	3	G3ICK8
4268	0.993731	3	ATDFVVPGPGKVEITYTPK		4890	0.999749	3	G3I6P1
15281	0.993283	4	VKDLESLFHR		2621	0.999749	14	G3H8Z6
9234	0.992903	4	FDVNTSAVQVLIEHIGNLDR		2284	0.999749	3	G3H450
4092	0.9914	3	VLDNAIETEK		3612	0.999661	5	G3HN43
8052	0.990063	5	AVYGTYILGVLQEEEEEEK		3752	0.999634	3	G3HG79
7950	0.990013	3	LMIWDTR		408	0.999595	14	G3I180

11649	0.989221	10	TVIVAALDGTFQR		19	0.999561	13	G3HZE5
12418	0.989158	3	EFPGLAEVTVAK		3876	0.999449	4	G3HXF7
4043	0.988813	5	QTIDNSQGAYQEAFDISKK		4892	0.99912	7	G3H215
15888	0.988789	3	ATIELKALR		2538	0.998661	47	G3I4A0
9522	0.988535	3	SVTQQPGPEGTATPK		1992	0.99836	3	G3IBN9
5970	0.988234	4	QLGSGLLLVTGPLALNRVPLRR		4893	0.996032	5	G3I3X4
16073	0.988221	3	EYAEDDNIYQQKIK		3429	0.994536	7	G3H3F8
13217	0.987906	5	GHYTIGKEIIDLVLDR		3667	0.989871	4	G3IG24
4777	0.987702	4	YEGFFGLYR		2248	0.951743	7	G3IET4
13062	0.985883	3	EKLCYVALDFEQEMATAASSSSLEK		2593	0.94692	4	G3GSE2
2105	0.98579	3	QTSSLVPPYLGMILSALVQGLAGR		2747	0.925906	4	G3HKF3
1899	0.984433	7	EAGVFVPR		2149	0.912463	3	G3H195
8298	0.984347	5	KEDLVFIFWAPESAPLKSK		2574	0.909296	4	G3HLK9
4332	0.983462	3	AFAEKNNLSFIETSALDSTNVEEAFK		2058	0.906061	6	G3GUG1
15683	0.983008	6	TVSFKLGDLEEAPER		2371	0.903242	3	G3GV42
5718	0.982515	4	HESGASIKIDEPLEGSEDR		3200	0.902854	12	G3H0Q2
5297	0.980071	3	YAFTTVSCR		807	0.901334	5	G3HV90
2052	0.979953	3	EELMFFLWAPEQAPLK		pep_code	mean_pred	n	
8246	0.978372	4	NQLDQEVEFLSTSIAQLK		2613	0.999979	3	GLVVLGFPCNQFGHQENAK
8907	0.97663	5	DSEAQTANCLDLDFVTR		18959	0.999966	3	MAEVEAPTAAETDMK
8628	0.976251	4	TVSLGAGAKDELHIVEAEAMNYEGSPIK		15719	0.999951	3	AEELKTELLK
10022	0.975348	3	ATGNAVSIFVYDVKPGAEEQTQVAK		15020	0.999932	8	QQLLEDLK
4097	0.975083	3	DAGLGVYSLALLNNVSYNVVEFSK		8240	0.999911	13	EQNPYEKF
2842	0.974052	4	KITISDCGQL		526	0.999908	3	FALNYLQLLVHPNPPSVLFGADK
5737	0.973402	3	LLYGFLVR		5734	0.999891	3	DREQSSLLITLLLPFLLR
8292	0.97269	3	KPGMFFNPEESELDLTYGNR		12585	0.999889	3	ACGEFVFLMDR
5497	0.972511	3	TYNFLPEFLDSTQK		12600	0.999885	3	VVLDVGSGTGILCMFAAK
7555	0.97229	17	VTVLFAGQHIAK		3729	0.999885	3	EAAWAISNLTISGRK
12715	0.971561	4	TYNIVLGSGQVVLGMDMGLR		7855	0.999883	4	EINQMEISGDITLTSAAPTMI
3613	0.969973	3	RYNVPFLTFINK		3832	0.999881	4	AQLVVIAHDVDPIELVVFLPALCRK
7093	0.968367	5	NLEPEWATAATEVKEQTK		14866	0.999876	3	GTSELLR
527	0.96777	6	LEGSEVQLVEYEASAAGLIR		7225	0.999874	4	ECVNFASFNFLGLLANPR
3336	0.966542	3	SITVIGGGFLGSELACALGRK		15489	0.999873	3	ADFVLAANSYDLAIK
4857	0.966067	3	EYGMIYLGK		18492	0.999869	3	ILGVDKNASEDEIK

6674	0.965926	3	YGISLCQAILDETKGDYEK		4268	0.999863	3	ATDFVVPGPGKVEITYTPK
6979	0.965443	3	EIIRNDGVLLLQALTR		12295	0.999862	3	AAQLCGAGMAAVVEK
2080	0.964488	11	HPTFTQFSFMVYNSLR		19530	0.999861	3	KMDLIDLEDETIDAEVMNSLAVTMDDFR
12417	0.964279	3	FNACFESVATNIDEIYK		6376	0.999856	3	ISIEMHGTLEDQLSHLR
5792	0.963867	3	RIAEFAFEYAR		12968	0.999854	3	TAMNVNEIFMAIAK
642	0.96354	4	AAGVCLMLLSTCCEDDIVPHVLPFIK		18952	0.999854	5	IANVSILK
15986	0.963323	21	KQLADETLLK		16183	0.999843	3	AVLCPPPVKK
12482	0.963306	3	SETSGSFEDALLAIVR		9290	0.999843	3	MFIGGLSWDTSK
5734	0.963115	4	DREQSSLLITLLLPFLLR		15824	0.99983	3	SIAQTTPVAGSQSLSSGTVAK
744	0.961051	3	NDANPETHAFVTSPEIVTALAIAGTLK		8126	0.999828	3	DKTLEGPDSSGISQAFPLR
9821	0.960584	12	LGLVKYGEK		6646	0.999825	3	SFYGTALFLCRR
4993	0.960505	3	VQQLLHICSEHFDSK		6153	0.999821	3	KFDQLLAEEK
7891	0.960119	4	LIDLHSPSEIVK		16073	0.999813	3	EYAEDDNIYQQKIK
1633	0.959528	6	AGVAPLQVK		9234	0.999811	4	FDVNTSAVQVLIEHIGNLDR
5987	0.957995	3	AEVIMNYHVGETVLSLQK		12910	0.999808	3	HSSIFDQLDVVSYEEVVR
6246	0.957221	3	LLAVIEEQHK		5745	0.999783	4	FDDGAGGDNEVQR
9932	0.956354	3	LILWYYEHQLK		12518	0.999781	3	QSVVVLMGSLAK
638	0.953937	11	NICQFLIEIGLAK		5970	0.999766	4	QLGSGLLLVTGPLALNRVPLRR
5165	0.953072	5	TYIISILFK		6917	0.999741	3	TVGYTVTAPEDTR
14850	0.952893	4	VTEGLVDVILYHQPDDK		15827	0.99974	4	GHYTEGAELVDSVLDVIR
175	0.952679	4	VIHDNFGIVEGLMTTVHAITATQKTVDGPSGK		3792	0.99974	3	TYHALSNLPK
12361	0.94945	7	GALQYLVPILTQTLTK		14910	0.999711	10	TFSFAIPLIEK
5374	0.948527	4	GYLISGSSYAR		8397	0.999693	4	YLTVATVFR
6691	0.947865	3	ATVHIQVNDVNEYAPVFK		6470	0.999662	3	DPNAAKPEDWDERAK
429	0.947196	6	SLQDIIAILGMDELSEEDK		6536	0.999573	4	SSDQALSYEYITLSETK
12955	0.946977	5	EYQVETIVDTLCTNMLSDK		11649	0.993636	6	TVIVAALDGTFQR
4209	0.946858	3	IVSVGDDQEIHVYDCPI		7555	0.96917	10	VTVLFAGQHIAK
6682	0.945583	4	RISGLIYEETR		9821	0.968625	7	LGLVKYGEK
2690	0.944089	10	NLANTVTEEILEK		878	0.963862	4	AVVYSNTIQSIMAIVK
8339	0.944058	3	IAQDLEMYGINYFEIKNK		12715	0.953805	3	TYNIVLGSGQVVLGMDMGLR
12515	0.943731	3	LNEGVVEFASYGDLK		12700	0.953205	6	TFVQEDVYDEFVER
8906	0.943364	5	LTQTLQTLGDAQSK		3964	0.953182	5	RPWILTSLHNGVIQLWDYR
6887	0.943274	3	RLFEYILLYK		4777	0.941528	4	YEGFFGLYR
4739	0.941943	3	GYAFINFASFDASDAAIEAMNGQYLCNRPITVSYAFK		3988	0.932913	4	VDQSILTGESVSVIK

7806	0.941093	23	DIVVAETLEDLDKNK		10	0.925978	4	GAGSYTIMVLFADQATPTSPIR
7737	0.940985	6	EAASILQELQVETYGSMEK		12741	0.923937	3	NVNIFKFIIPNVVK
5499	0.93977	4	YSDNWEAVTR		8159	0.922457	6	ERVETLSQVEVILQQSAADIAR
1539	0.93965	3	DIFSATASQLHSNVVNYVQQILVPK		15986	0.921046	21	KQLADETLLK
5189	0.936844	3	NLELTAVGGSDVHVR		7806	0.92025	27	DIVVAETLEDLDKNK
15963	0.935166	4	GKLEAIITPPPAK		8246	0.917431	4	NQLDQEVEFLSTSIAQLK
11196	0.933901	12	SNLKIAELK		15575	0.917274	4	LEIGVQSVYEDVAR
12451	0.933861	4	AVLTNILLR		9124	0.913633	3	VSYFEIYLDK
1872	0.92989	3	TLQVETLVQPPESCTESAAFGDTLHIHYTGSLVDGR		15963	0.913574	7	GKLEAIITPPPAK
7504	0.928168	3	EQQEAIEHIDEVQNEIDRLNEQASEEILK		4739	0.913402	4	GYAFINFASFDASDAAIEAMNGQYLCNRPITVSYAFK
12700	0.927488	9	TFVQEDVYDEFVER		638	0.912579	10	NICQFLIEIGLAK
7626	0.923964	3	VAEKLEALSVK		5972	0.912463	3	IAAQGFTIAALLLGLAVSAMK
8025	0.92272	3	TLQDAFPGAFGDLYTQNTR		8628	0.906104	4	TVSLGAGAKDELHIVEAEAMNYEGSPIK
3819	0.921142	3	IQSIPLVIGQFLEAVDQNTAIVGSTTGSNYYVR		15095	0.903233	7	LDQLIYIPLPDEK
4950	0.92077	4	AFSSSSFSFNTFLTR		10733	0.902854	12	GGASDALLYR
3359	0.920683	4	IASCMEVLDLFLNQPYAVLSSFVLAEK		8218	0.901241	10	DLAGSIIGKGGQR
4060	0.920428	13	AFIITINSFGTELSK		5559	0.900955	13	RLAYAIIQFLHGQLR
6204	0.919943	3	LNLQVLLSYVLGK		3094	0.900132	4	DGSEASLEWSSDRQDIAVISDSYFPR
10338	0.918831	3	NMLLYFAYADYEESR		Seed 110			
5608	0.918733	4	TCGFDFSGAIEDISK		acc_code	mean_pred	n	
2008	0.918402	4	DLPITEAVFSALVTGHAR		967	0.999959	5	G3GTZ4
7912	0.918395	5	TLFLAVR		4886	0.999939	3	G3HYN9
3847	0.918099	5	MDENQFVAVTSTNAAK		2456	0.999917	3	G3I9A3
7066	0.918011	8	YMVADKFTELQK		3811	0.999868	3	G3I7N4
12455	0.917068	6	GSDIIVDLEVTLEEVYAGNFVEVVR		4456	0.999836	3	G3H3L4
12486	0.91667	3	NLHQSGFSLSGAQIDDNIPR		1650	0.999804	4	G3I1H7
935	0.915603	4	IVAVIGAVVDVQFDEGLPPILNALEVQGR		4892	0.999778	6	G3H215
15881	0.915132	3	SAFPFSDKLGELSDK		2621	0.999673	11	G3H8Z6
7091	0.912118	4	EGTQASEGYFSQSQEEEFAQSEEPCAK		3612	0.999413	3	G3HN43
5559	0.909839	11	RLAYAIIQFLHGQLR		408	0.999267	11	G3I180
6925	0.909315	3	AVLMAHDCYQEAVTHFR		2538	0.999157	44	G3I4A0
5248	0.908562	10	NSWGSSWGMGGFFHIAK		19	0.998955	7	G3HZE5
6499	0.90758	8	LISVETDIVVDHYK		3876	0.99877	7	G3HXF7
3085	0.906808	3	KVADLESEAR		2884	0.998735	3	G3H565

10534	0.905913	3	STVTIGLVQALTAHLNVNSFACLR		4893	0.998611	7	G3I3X4
13403	0.904828	3	APVDFGYVGIDSILEQMR		2185	0.998229	11	G3HP82
12576	0.904138	3	VITDVYYVDGTEVER		3003	0.994266	3	G3I9X3
12121	0.904096	3	TFFSFPAVVAPFK		2574	0.989917	3	G3HLK9
7410	0.903982	5	VFQFLNAK		1653	0.986969	6	G3INX6
16051	0.903462	3	APPTWVKAVASVLGR		3678	0.969549	3	G3H3F9
2651	0.902789	6	ALVNCQYSSATFSTGER		807	0.956263	5	G3HV90
8526	0.902779	3	FCLDNGAK		3429	0.944243	6	G3H3F8
7174	0.901484	3	SIYGERFPDENFK		409	0.93179	11	G3HLU5
Seed 123					1974	0.931089	3	G3IJJ7
accession	mean_pred	n			3592	0.928917	6	G3HGQ5
2185	0.999562	6	G3HP82		984	0.927915	3	G3I6J6
4385	0.999425	3	G3I540		3626	0.925775	3	G3I1G2
3675	0.998965	3	G3I0V9		353	0.923079	3	G3H1Q6
2621	0.998535	14	G3H8Z6		2641	0.911468	6	G3HD22
3664	0.997157	3	G3H4A9		3200	0.907115	19	G3H0Q2
2250	0.996439	3	G3H421		2406	0.904837	4	G3HEA9
4892	0.995902	5	G3H215		2747	0.901418	5	G3HKF3
408	0.993849	15	G3I180		pep_code	mean_pred	n	
2149	0.99253	4	G3H195		18959	0.999939	3	MAEVEAPTAAETDMK
2753	0.9885	3	G3GZE7		18952	0.999812	6	IANVSILK
3876	0.988196	8	G3HXF7		15937	0.999858	3	SQQYSSIYGASVSDEELK
2574	0.987238	3	G3HLK9		15827	0.999927	5	GHYTEGAELVDSVLDVIR
3657	0.986922	3	G3I3X9		15824	0.999924	3	SIAQTTPVAGSQSLSSGTVAK
1992	0.985629	3	G3IBN9		15719	0.999929	3	AEELKTELLK
1144	0.98555	3	G3H4E2		15575	0.91053	6	LEIGVQSVYEDVAR
1166	0.98551	3	G3H436		15360	0.999869	3	SQEMVHLVNKEPSETPDQFMTADETR
4890	0.980765	3	G3I6P1		15281	0.999904	3	VKDLESLFHR
3626	0.979705	5	G3I1G2		15278	0.999809	3	IKIIAPPER
2538	0.978889	48	G3I4A0		15255	0.999863	3	NTEISFKLDQEFEEITADDRK
4893	0.974082	7	G3I3X4		15095	0.932549	5	LDQLIYIPLPDEK
2339	0.972778	4	G3IAL6		15029	0.999873	3	SESPGGNQALLLLR
2159	0.971189	5	G3I8L4		15020	0.99996	10	QQLLEDLK
3752	0.961274	4	G3HG79		14910	0.999888	11	TFSFAIPLIEK

2082	0.961225	3	G3H7X9		12941	0.99982	3	ILACHGTDSVLEVFCVLSK
19	0.959738	4	G3HZE5		12910	0.999872	3	HSSIFDQLDVVSYEEVVR
2593	0.959395	3	G3GSE2		12741	0.905622	3	NVNIFKFIIPNVVK
1216	0.943276	4	G3IHB4		12700	0.936189	12	TFVQEDVYDEFVER
807	0.941926	3	G3HV90		12683	0.969549	3	AAITFSFFSIFSWGVLAFLAYQR
1087	0.932385	4	G3H636		12662	0.999887	3	TQEAIIQIMKMR
4408	0.9301	3	G3HRC6		12521	0.90495	4	VTWDSSFCAVNPR
3562	0.928012	3	G3I0Y0		12482	0.922931	5	SETSGSFEDALLAIVR
1907	0.919231	4	G3GXS2		12451	0.908581	4	AVLTNILLR
3200	0.913962	15	G3H0Q2		12430	0.925775	3	NFAVWLR
3429	0.910847	3	G3H3F8		12410	0.999889	3	AMTHVINQGMAMYWGTSR
3631	0.909466	7	G3I6L2		12383	0.999897	3	GESLADSVQTMSCYADVVVLR
pep_seq	mean_pred	n			12295	0.999869	3	AAQLCGAGMAAVVEK
15020	0.999988	5	QQLLEDLK		12285	0.928917	6	DLEIAAEFVLSASVQELLDVLK
12671	0.998965	3	ESLVLQVSVLTDQVEAQGEK		12278	0.999863	3	VHNDAQNFDYDHDAFLGAEEAK
8240	0.998535	14	EQNPYEKF		12275	0.999831	3	LPVGTTATLFFR
14910	0.998265	7	TFSFAIPLIEK		11649	0.944243	6	TVIVAALDGTFQR
8339	0.997646	4	IAQDLEMYGINYFEIKNK		11196	0.921746	19	SNLKIAELK
12608	0.997157	3	EAALLTIVCVEEMQDMEK		10733	0.907115	19	GGASDALLYR
12700	0.99689	11	TFVQEDVYDEFVER		9821	0.974618	6	LGLVKYGEK
18952	0.99655	6	IANVSILK		8756	0.905583	4	LSQTSNVDKEEEAVTIEMNEPVQLTFALR
3964	0.996448	3	RPWILTSLHNGVIQLWDYR		8628	0.900504	5	TVSLGAGAKDELHIVEAEAMNYEGSPIK
6285	0.996427	3	TFLLDGDEVIITGHCQGDGYR		8298	0.999832	4	KEDLVFIFWAPESAPLKSK
6044	0.996289	3	YAIQLITAASLVCRK		8240	0.999908	10	EQNPYEKF
12600	0.995741	4	VVLDVGSGTGILCMFAAK		8218	0.936476	14	DLAGSIIGKGGQR
7806	0.99568	10	DIVVAETLEDLDKNK		8159	0.904594	4	ERVETLSQVEVILQQSAADIAR
18541	0.994818	3	KINELTGIK		8052	0.989917	3	AVYGTYILGVLQEEEEEEK
5970	0.993734	4	QLGSGLLLVTGPLALNRVPLRR		8043	0.999868	3	TFTYLLGSDAAALLFNSK
2729	0.993495	4	YFLNLVLVDEEDRR		7806	0.965509	13	DIVVAETLEDLDKNK
12413	0.993407	4	NNECCMAIPLSQIVFIEEQAAGIGKSAK		7555	0.96888	20	VTVLFAGQHIAK
12383	0.992549	3	GESLADSVQTMSCYADVVVLR		7475	0.999917	3	ILHEDGSPNLDFKPEDQPHFDIKDEF
5972	0.99253	4	IAAQGFTIAALLLGLAVSAMK		7066	0.93856	10	YMVADKFTELQK
13217	0.992021	3	GHYTIGKEIIDLVLDR		6845	0.999903	3	TFISYSVTFK
2310	0.991532	3	FIYMAVQHYIETLQR		6220	0.99992	3	ASFNHFDKDHVGALGPEEFK

12295	0.990162	3	AAQLCGAGMAAVVEK		6157	0.999908	4	YTMGDAPDFDR
8052	0.987238	3	AVYGTYILGVLQEEEEEEK		5970	0.999842	3	QLGSGLLLVTGPLALNRVPLRR
15084	0.986955	4	KAYGQALAK		5745	0.999724	3	FDDGAGGDNEVQR
6648	0.986554	4	VLEALLSLK		5734	0.999896	3	DREQSSLLITLLLPFLLR
18948	0.986368	3	KQDFPLVK		5559	0.909578	15	RLAYAIIQFLHGQLR
3729	0.986355	3	EAAWAISNLTISGRK		5497	0.999873	4	TYNFLPEFLDSTQK
15932	0.985644	3	SSEFYKVTTEEYQK		5229	0.908214	3	LEITMLSQGVSMLYSFFMPAAK
2456	0.98555	3	LLLLLNK		4777	0.945684	3	YEGFFGLYR
2286	0.985529	3	RNVGVVVVGFPATPIIESR		4602	0.999894	3	ALEHFTDLYDIKR
3203	0.985499	4	SYEAAASALQIAAHTAFVAK		4097	0.999772	3	DAGLGVYSLALLNNVSYNVVEFSK
15360	0.985225	4	SQEMVHLVNKEPSETPDQFMTADETR		4092	0.999895	3	VLDNAIETEK
2270	0.984958	3	QIFILLFQRLQNSK		3964	0.936985	6	RPWILTSLHNGVIQLWDYR
4332	0.984215	4	AFAEKNNLSFIETSALDSTNVEEAFK		3832	0.999928	3	AQLVVIAHDVDPIELVVFLPALCRK
3832	0.982041	4	AQLVVIAHDVDPIELVVFLPALCRK		3763	0.999893	3	DLFPYEESKEK
12985	0.980707	3	DVNTAVMELLIMAYALK		3329	0.999867	3	FNGGGHINHSIFWTNLSPK
12430	0.979705	5	NFAVWLR		2445	0.999882	3	VPHWYHFSCFWK
17543	0.979412	3	LASLGLKFNK		2105	0.99994	3	QTSSLVPPYLGMILSALVQGLAGR
4316	0.977638	5	FGIFTCFDILFHDPAVTLVTK		2018	0.927128	4	AIYEVLFR
3094	0.977232	5	DGSEASLEWSSDRQDIAVISDSYFPR		2012	0.999892	3	RSSFICVILSHGEEGVIFGTDGPVDLK
16978	0.977006	3	LELLLQKR		2004	0.99987	3	YFDLSYLFRK
7991	0.976921	3	QITQVYGFYDECLRK		1899	0.916136	6	EAGVFVPR
6536	0.976854	3	SSDQALSYEYITLSETK		1872	0.927915	3	TLQVETLVQPPESCTESAAFGDTLHIHYTGSLVDGR
6220	0.975986	4	ASFNHFDKDHVGALGPEEFK		1851	0.999907	4	LRASILNTWISLK
527	0.975887	3	LEGSEVQLVEYEASAAGLIR		1796	0.999959	5	AVNAHVDLGYSTLVGVFHTEYGALNR
5792	0.975832	4	RIAEFAFEYAR		1696	0.999905	3	KDLLSWLLSHGVPR
6674	0.975726	3	YGISLCQAILDETKGDYEK		638	0.920516	6	NICQFLIEIGLAK
13734	0.975603	3	AVVGEEALTSDDLLYLEFLQK		10	0.921128	4	GAGSYTIMVLFADQATPTSPIR
7661	0.972956	3	ALDLDSSCK		Seed 123			
6944	0.972778	4	SLHTLFGDK		acc_code	mean_pred	n	
8628	0.972628	3	TVSLGAGAKDELHIVEAEAMNYEGSPIK		1196	0.999989	3	G3H8G0
6586	0.972279	3	VRDTYLDTQVVGQTGVIR		2649	0.999973	3	G3HVG5
3395	0.969533	5	GRVFEVSLADLQNDEVAFRK		3876	0.999757	9	G3HXF7
15910	0.969194	3	EYGGLNVLVNNAGIAFR		4892	0.999719	5	G3H215
4442	0.968051	6	TNDQMVVVYLASLIRSVVALHNLINNK		3752	0.99967	5	G3HG79

4617	0.966445	3	LLLPWLEAR		408	0.999663	12	G3I180
8779	0.966086	3	TVVFTQGRDDTVVATENEVMAFAVLDQNQK		3612	0.999574	5	G3HN43
429	0.963431	8	SLQDIIAILGMDELSEEDK		19	0.99951	7	G3HZE5
3336	0.96338	3	SITVIGGGFLGSELACALGRK		3843	0.999243	3	G3IFR0
6767	0.962323	5	GHNKESNPSNDIFLGHTNIEEIEK		2884	0.998954	3	G3H565
2949	0.962082	4	VFSVAITPDHLEPR		2538	0.99857	48	G3I4A0
2740	0.961965	5	SVKPQILSVFGDIALAIGGEFKK		4893	0.998186	7	G3I3X4
15880	0.961289	3	SQPLQNVSAVLQKPNPLYNQNADMVQK		4364	0.997897	5	G3HHK3
5689	0.961262	3	LFFIQACR		3667	0.99457	4	G3IG24
5737	0.961225	3	LLYGFLVR		2185	0.990849	9	G3HP82
7555	0.960738	17	VTVLFAGQHIAK		2621	0.977125	12	G3H8Z6
9505	0.959836	3	MAKPEEVLVVENDQGEVVR		3429	0.962774	5	G3H3F8
4020	0.959283	3	VLGIAFSSAR		3626	0.949267	5	G3I1G2
5559	0.959191	8	RLAYAIIQFLHGQLR		3200	0.944355	16	G3H0Q2
4099	0.958118	5	IGFEEKDIAANEENRK		2747	0.92016	3	G3HKF3
9301	0.957469	7	GTAFIIR		4265	0.918537	3	A5A422
16080	0.955424	4	TDEFQLHTNVNDGTEFGGSIYQK		1919	0.903875	6	G3HYC4
4528	0.953748	3	SRRPLTGCTIFLGYTSNLISSGIR		2082	0.902691	3	G3H7X9
6331	0.953532	3	FSLETEVDLR		pep_code	mean_pred	n	
15881	0.951604	6	SAFPFSDKLGELSDK		2613	0.999989	3	GLVVLGFPCNQFGHQENAK
2795	0.951071	6	EPVYLSGYGVELAIK		15719	0.999967	5	AEELKTELLK
6440	0.948531	4	MTQIMFEAFNTPAMYVAIQAVLSLYASGR		6247	0.999963	3	ATFAGLYSLDMGEEGDQAIAEALAAPNHFVLKPQR
7061	0.948302	3	YHGQHVTGSPFQFTVGPLGEGGAHK		15840	0.99994	3	FANSLVGVQQQLQAFNTYR
7307	0.948193	3	AVQGVHTGVNQAQKETEK		15020	0.999939	8	QQLLEDLK
15280	0.944492	4	LDNAKLSSDQNDK		227	0.999926	4	SADTLWGIQKELQF
5165	0.943813	3	TYIISILFK		13467	0.999925	3	GEYNVYSTFQSHEPEFDYLKSLEIEEK
5382	0.943342	3	ALKDEEKMELQEIQLK		12600	0.999925	3	VVLDVGSGTGILCMFAAK
2673	0.943276	4	INEKPQVIADYESGR		3763	0.999922	3	DLFPYEESKEK
1359	0.941926	3	TLVLIGASGVGR		16214	0.999921	4	KMAAGHELQPLAIVDQRPSSR
6680	0.94103	3	QHDHLEAQKQEYQQAVQQLEHK		15827	0.999916	4	GHYTEGAELVDSVLDVIR
4863	0.940178	4	SASYKPVFVTEITDDLHFYVQDVETGTQLEK		6586	0.999915	3	VRDTYLDTQVVGQTGVIR
6169	0.940042	5	VDHFGFYNSR		3832	0.999912	4	AQLVVIAHDVDPIELVVFLPALCRK
5474	0.938524	3	LPITVLNGAPGFINLCDALNAWQLVK		13128	0.999909	3	GVAALCIGGGMGIAMCVQR
9628	0.934831	3	IQFNAHIWTK		7225	0.999903	4	ECVNFASFNFLGLLANPR

15986	0.93169	13	KQLADETLLK		12438	0.999902	3	NPPVNSLSLEFLTEFVISLEK
12329	0.93162	5	GAMGIILVYDITDEK		7180	0.999898	3	EIEDFDGLEALRLEGNTVGVEAAK
3790	0.931377	3	NNQGTVNWSVEDIVK		9101	0.999891	3	LASDLLEWIR
863	0.930937	3	SNLGSVVLQLK		14910	0.999884	15	TFSFAIPLIEK
6947	0.930485	3	SQLLQYVYNLVPR		15489	0.999882	4	ADFVLAANSYDLAIK
4654	0.929444	6	AYIVQLQIEDLTRK		5942	0.999867	3	HNATEMIWAVLAAFNK
9821	0.929096	11	LGLVKYGEK		18952	0.999854	7	IANVSILK
12623	0.928012	3	DTTVIVWQVDPDTHLLK		13217	0.999849	6	GHYTIGKEIIDLVLDR
7066	0.927486	8	YMVADKFTELQK		4974	0.999842	3	YVATLGVEVHPLVFHTNRGPIK
13542	0.926587	3	QDLGFFQETKTGELNSR		3329	0.999827	3	FNGGGHINHSIFWTNLSPK
9334	0.925332	5	EVVHTVSLHEIDVINSR		15278	0.999815	4	IKIIAPPER
11460	0.924759	12	DYQDNKADVILK		5970	0.999809	5	QLGSGLLLVTGPLALNRVPLRR
8218	0.923045	9	DLAGSIIGKGGQR		6153	0.999783	4	KFDQLLAEEK
12837	0.92254	3	HLSNGASALVYFSR		8240	0.999767	11	EQNPYEKF
8526	0.921516	5	FCLDNGAK		6220	0.999762	3	ASFNHFDKDHVGALGPEEFK
16218	0.92114	3	TLAEHQQLIPLVEKAK		3729	0.999751	4	EAAWAISNLTISGRK
8906	0.920679	4	LTQTLQTLGDAQSK		7806	0.98008	14	DIVVAETLEDLDKNK
12838	0.920379	3	GNFNYIEFTR		15095	0.96742	8	LDQLIYIPLPDEK
12523	0.920241	3	VTEILGFEDDVVIEFIFNQLEVK		5559	0.964986	7	RLAYAIIQFLHGQLR
5058	0.919231	4	LQQEATEHITESEER		11649	0.962774	5	TVIVAALDGTFQR
12713	0.916206	4	EADIDGDGQVNYEEFVQMMTAK		3964	0.961597	4	RPWILTSLHNGVIQLWDYR
4818	0.916057	5	EGVTTYFSGNCTMEDAK		12413	0.955785	3	NNECCMAIPLSQIVFIEEQAAGIGKSAK
7775	0.915722	4	TLVLLDNLNVR		9821	0.954428	9	LGLVKYGEK
8909	0.915722	4	LHISPSNMTNQNTNEYLEK		12430	0.949267	5	NFAVWLR
5384	0.914203	6	RIVGCICEEDNCTVIWFWLHK		10733	0.944355	16	GGASDALLYR
2185	0.914169	7	RALAVHLWYLLPPTASISR		6048	0.940831	4	TQVCGILREELYQGDAFHQADTHIFIIMGASGDLAK
10733	0.913962	15	GGASDALLYR		14901	0.940669	8	LSLVSLLVR
5220	0.913536	6	KPLLESGTLGTKGNVQVVIPFLTESYSSSQDPPEK		12700	0.940588	9	TFVQEDVYDEFVER
7674	0.91344	4	ITVASLDASNLLASVPASQR		7555	0.937477	16	VTVLFAGQHIAK
11649	0.910847	3	TVIVAALDGTFQR		447	0.936158	7	NVFDEAILAALEPPEPKK
12710	0.910709	4	DLLLKCLTFNPAK		15683	0.93435	3	TVSFKLGDLEEAPER
15701	0.908702	4	EAVELPLTHFELYK		3094	0.929507	3	DGSEASLEWSSDRQDIAVISDSYFPR
1621	0.90839	3	LNTLLQR		8756	0.928163	4	LSQTSNVDKEEEAVTIEMNEPVQLTFALR
14901	0.907248	12	LSLVSLLVR		15910	0.918271	5	EYGGLNVLVNNAGIAFR

15095	0.906906	7	LDQLIYIPLPDEK		5474	0.91502	4	LPITVLNGAPGFINLCDALNAWQLVK
15963	0.906496	6	GKLEAIITPPPAK		11196	0.914822	14	SNLKIAELK
15575	0.906259	6	LEIGVQSVYEDVAR		3874	0.90949	4	YVISLNEDSTVHGFIVQLPLDSENPINTEAVINAITPDK
11196	0.906099	14	SNLKIAELK		8218	0.90728	19	DLAGSIIGKGGQR
1920	0.905557	3	QLHELAPSIFFYLVDAEHGR		5737	0.902691	3	LLYGFLVR
15227	0.904851	4	TEWLDGKHVVFGK		2729	0.900218	3	YFLNLVLVDEEDRR
7034	0.90301	3	DVFFYQADDEHYIPR					
9343	0.901164	5	GTVEPQLEAR					
4609	0.900741	5	MEGNAEESTLFCFAVR					
9537	0.90068	3	GLALDTQHHLATFHLDSEASIR					