
Proteomics of lipid accumulation and DGAT inhibition in
HepG2 liver carcinoma cells.

By

Bhumika Bhatt-Wessel

A thesis

submitted to Victoria University of Wellington
in fulfilment of the requirement for the degree of

Doctor of Philosophy

In

Cell and Molecular Biology.

Victoria University of Wellington

2017.

ABSTRACT

Non-alcoholic fatty liver disease (NAFLD) is a manifestation of the metabolic syndrome in the liver. It is marked by hepatocyte accumulation of triacylglycerol (TAG) rich lipid droplets. In some patients, the disease progresses to non-alcoholic steatohepatitis (NASH), characterized by cellular damage, inflammation and fibrosis. In some cases, cirrhosis and liver failure may occur. However, the pathogenesis of NAFLD is still unclear. The present project is based on the hypothesis that hepatocytes are equipped with mechanisms that allow them to manage lipid accumulation to a certain extent. Continued or increased lipid accumulation beyond this triggers molecular mechanisms such as oxidative stress, lipid peroxidation and cell death that aggravate the condition and cause disease progression. The aim of this project is to study the effects of lipid accumulation on the cells using proteomics approach to identify proteins involved in the disease progression.

A cell culture model was used in the study. HepG2 cells, a human liver carcinoma cell line, were treated with a mixture of fatty acids (FA) to induce lipid accumulation. The lipid accumulation in HepG2 cells was measured with Oil red O assay and the effect of lipid accumulation on the proliferation of the cells was measured using an MTT cell proliferation assay. HepG2 cells treated with 1 mM FA mixture for 6 hours induced lipid accumulation 1.4 times of control with 90% of cell proliferation capacity of the control cells.

The final and the only committed step in TAG biosynthesis is catalysed by acyl-CoA diacylglycerol acyltransferase (DGAT) enzymes. To investigate if limiting lipid accumulation in HepG2 cells would prevent molecular mechanisms of pathogenesis, inhibition of DGAT by small molecule inhibitors was performed. Among the three DGAT inhibitors (A922500, PF06424439 and PF04620110) tested, PF04620110 reduced the lipid accumulation to 1.2 fold of the control cells when they were treated with 100 μ M of the inhibitor in the presence of 1 mM FA mixture for 6 h.

Proteomic analyses were carried out for the control, FA-treated and inhibitor-treated cell groups to identify protein changes in the abundance. Functional analyses of the changed proteins identified suggest that lipid accumulation tends to adversely affect the functioning of the ER and the mitochondria. A complex interplay between the two organelles, possibly mediated by Ca^{2+} signalling may be vital in ensuring cell survival. PF04620110 was able to counter the FA induced changes in the abundance of some proteins involved in the

metabolic processes but it had limited effect on the ER chaperones whose abundance in the inhibitor-treated sample was comparable to that of the FA-treated sample. These data provided important information for future discoveries of biomarkers and molecular mechanisms involved in the progression of NAFLD.

ACKNOWLEDGEMENTS

The time spent working on this thesis has been a profound lesson in life summed up in three traits I had to use extensively - patience, resilience and fortitude. Patience to wait and not be tired of waiting, resilience in accepting unexpected outcomes and the fortitude to keep on going.

I have been supported by the best of people in this journey and tested thoroughly by some. They have all helped me put the best of me in this work. Thank you to my supervisors Dr. Bill Jordan and Prof. John Miller for the encouragement, support and advice – this project would not have been possible without your help. Dr Lifeng Peng, my primary supervisor, has introduced me to proteomics and my work is a reflection of her training; thank you for your critiques and support with the LC-MS/MS that is at the heart of this work. The introduction in this thesis has excerpts from the manuscript of a review Dr. Peng and I have worked on. Her feedback has improved the quality of the work and I hope the review will be published soon.

I must acknowledge Dr. Ian Vorster for the help with the MS analysis of A922500 and Dr. Christine Stockum and Sarah Corinder who initiated me into the lab and helped build my skills in Western blotting and tissue culture respectively. Thank you Leticia Castro for keeping a constant vigil against mycoplasma and, along with Remi Schneider, trying your luck at growing HepG2 cells when everything seemed to fail. A special thanks to Sushila Pillai - the confocal microscopy guru who has helped me with the imaging of immunolabelled cells. Thank you Dr. Lisa Woods for your insight into biostatistics.

Over the course of this PhD, this beautiful city has become my home and the amazing friends I have made are the best surrogate family I could hope for. Nishad, Karan, Jasmine, Christina, Rose, Rosie, Remi, Dinido and the whole Chiasma crew – you guys are all right!

There are perhaps not enough words to express the deep gratitude and love for my partner Ronald Wessel who has been as much part of this PhD as I have. He's the only bridge engineer I know who can talk about lipid accumulation in the liver like he knows it. At the end of the day, I hope I can make my parents and mota mummy proud and fulfil the aspirations they have for me.

TABLE OF CONTENTS

| | |
|--|-----|
| ACKNOWLEDGEMENTS | iii |
| LIST OF FIGURES | ix |
| LIST OF TABLES | xii |
| CHAPTER 1: INTRODUCTION | 1 |
| 1.1. Conception of this project..... | 1 |
| 1.2. Prevalence of non-alcoholic fatty liver disease | 1 |
| 1.3. Progression from NAFLD to NASH | 1 |
| 1.4. Factors affecting disease progression..... | 3 |
| 1.5. Molecular mechanism of disease progression | 4 |
| 1.5.1. First hit: Accumulation of triglycerides in lipid droplets..... | 4 |
| 1.5.2. ER stress | 5 |
| 1.5.3. Mitochondrial dysfunction..... | 6 |
| 1.6. Role of DGATs in triglyceride synthesis | 7 |
| 1.6.1. DGAT genes..... | 7 |
| 1.6.2. Membrane topology | 10 |
| 1.6.3. Source of substrate | 11 |
| 1.6.4. DGAT Inhibitors..... | 12 |
| 1.7. Models used to study NAFLD | 13 |
| 1.7.1. Clinical studies in humans..... | 13 |
| 1.7.2. Animal models of NAFLD | 14 |
| 1.7.3. Cell-culture based models..... | 15 |
| 1.8. Proteomics approach to lipid accumulation | 15 |
| 1.8.1. Separation of peptides and mass spectrometry | 16 |
| 1.8.2. Peptide sequencing and identification of proteins..... | 18 |
| 1.8.3. Quantitative proteomics | 19 |
| 1.9. Aims of the thesis..... | 21 |
| CHAPTER 2: MATERIALS & METHODS..... | 22 |
| 2.1 Materials | 22 |
| 2.2 Cell culture | 23 |

| | |
|---|----|
| 2.3 Lipid accumulation in HepG2 cells..... | 24 |
| 2.3.1. Preparation of fatty-acid stock solutions | 24 |
| 2.3.2. Optimization of treatment conditions to induce lipid accumulation..... | 24 |
| 2.3.3. Quantification of lipid accumulation using Oil Red O assay..... | 25 |
| 2.3.3.1. Preparation of Oil Red O staining solution | 25 |
| 2.3.3.2. Staining of lipid droplets..... | 25 |
| 2.3.3.3. Quantification of extracted dye | 25 |
| 2.3.4. MTT assay | 25 |
| 2.4. Effect of DGAT inhibitors on lipid accumulation | 26 |
| 2.5. Fluorescent DGAT assay | 27 |
| 2.5.1. Microsome preparation..... | 27 |
| 2.5.2. Reaction mixture for DGAT assay | 27 |
| 2.5.3. Separation and quantification of products | 28 |
| 2.6. Proteomic analysis by LC-MS/MS | 28 |
| 2.6.1. Type of replicates | 28 |
| 2.6.2. Sample preparation | 29 |
| 2.6.3. LC-MS/MS analysis of protein digests | 30 |
| 2.6.3.1. Instrumentation..... | 30 |
| 2.6.3.2. Separation of peptides using reverse-phase liquid chromatography | 30 |
| 2.6.3.3. Mass spectrometry analysis | 31 |
| 2.6.4. Protein identification..... | 31 |
| 2.6.6. Statistical analysis..... | 34 |
| 2.6.7. Functional analysis: Gene Ontology and KEGG pathways..... | 35 |
| 2.7 Validation of proteomics results | 35 |
| 2.7.1 Western Blots | 35 |
| 2.7.1.1. Separation of protein | 35 |
| 2.7.1.2. Detection of bands | 36 |
| 2.7.1.3. Quantification of bands | 37 |
| 2.7.2 Immunofluorescent labelling of protein | 37 |
| CHAPTER 3: DETERMINING TREATMENT CONDITIONS..... | 38 |
| 3.1. Introduction..... | 38 |
| 3.1.1. HepG2 cells as cell culture model for lipid accumulation | 38 |

| | |
|--|----|
| 3.1.2. DGAT inhibitors | 38 |
| 3.1.2. Measuring lipid accumulation..... | 39 |
| 3.2 Results | 39 |
| 3.2.1. Lipid accumulation in HepG2 cells | 39 |
| 3.2.1.1. Linear range of absorbance for Oil Red O..... | 40 |
| 3.2.1.2. Inducing lipid accumulation | 41 |
| 3.2.2. Effect of DGAT inhibitors on lipid accumulation..... | 46 |
| 3.2.2.1. DGAT1 inhibitor A922500 | 46 |
| 3.2.2.1.1. Effect of experimental conditions on the effect of A92 on cultured cells..... | 47 |
| 3.2.2.1.2. Stability of A922500 | 48 |
| 3.2.2.1.3. Inhibition of in vitro DGAT activity..... | 49 |
| 3.2.2.1.4. Expression of DGAT enzymes..... | 51 |
| 3.2.2.2. Pfizer DGAT inhibitors – PF04 and PF06. | 52 |
| 3.2.2.2.1. Effect of PF04 and PF06 on in vitro microsomal DGAT activity | 52 |
| 3.2.2.2.2. Effect of PF04 and PF06 in cell culture- based assays | 52 |
| 3.3. Summary | 56 |
| CHAPTER 4: PROTEOMICS ANALYSIS – PART I | 57 |
| 4.1. Introduction | 57 |
| 4.2. Results | 58 |
| 4.2.1. Selection of a lysis buffer for sample preparation..... | 58 |
| 4.2.2. Dataset LC1 | 60 |
| 4.2.2.1. LC-MS/MS parameters..... | 60 |
| 4.2.2.2. Protein identification | 61 |
| 4.2.2.3. Label-free quantification..... | 62 |
| 4.2.2.4. Validation of proteins detected | 66 |
| 4.2.3. Improving protein identification..... | 68 |
| 4.2.3.1. Comparison of LC-MS/MS parameters | 68 |
| 4.2.3.1.1. Optimization of LC gradient | 68 |
| 4.2.3.1.2. Parameters for data-dependent MS/MS fragmentation..... | 71 |
| 4.2.3.2. Parameters for protein identification | 73 |
| 4.2.3.2.1. Comparison of Proteome Discoverer workflows..... | 74 |
| 4.2.3.2.2. Comparison of protein databases..... | 75 |

| | |
|--|-----|
| 4.2.4. Dataset LC2 | 76 |
| 4.2.4.1. Protein identification and quantification | 76 |
| 4.2.4.2. Comparison of datasets..... | 77 |
| 4.3. Summary..... | 79 |
| CHAPTER 5: PROTEOMICS ANALYSIS – PART II..... | 84 |
| 5.1. Introduction | 84 |
| 5.2. Results..... | 85 |
| 5.2.1. Dataset LC3 | 85 |
| 5.2.1.1. Protein identification..... | 85 |
| 5.2.1.2. Label-free quantification using spectral counts | 86 |
| 5.2.1.3. Variance associated with data generated from unpooled replicates | 87 |
| 5.2.2. Dataset LC4..... | 92 |
| 5.2.2.1. Protein identification..... | 92 |
| 5.2.2.2. Label-free quantification using spectral counts | 93 |
| 5.2.2.3. Variance associated with data generated from the pooled group of LC4 | 96 |
| 5.2.3. Functional analysis of proteins identified in dataset LC4..... | 97 |
| 5.2.3.1. Enrichment of identified proteins in GO categories | 105 |
| 5.2.3.1.1. Comparison of FA-treated group to control group | 105 |
| 5.2.3.1.2. Comparison of inhibitor-treated group to FA-treated and control group | 107 |
| 5.2.3.2. Enrichment of identified proteins in KEGG pathways..... | 109 |
| 5.2.3.2.1. Metabolic processes | 109 |
| 5.2.3.2.1.1 Glucose metabolism | 109 |
| 5.2.3.2.1.2. Fatty acid metabolism | 112 |
| 5.2.3.2.1.3. Purine synthesis and folate metabolism | 114 |
| 5.2.3.2.2. Protein processing and trafficking in the ER | 115 |
| 5.2.3.3. Mitochondrial proteins..... | 119 |
| 5.2.4. Validation of identified proteins | 120 |
| 5.2.4.1. Optimization of Western blot protocol | 121 |
| 5.2.4.2. Target proteins | 123 |
| 5.2.5. Further validation of perilipin 2 by 2D-electrophoresis | 127 |
| 5.2.5.1. Confirmation of antibody specificity for perilipin 2 | 127 |
| 5.2.5.2. Immunolabeling of perilipin 2 in the cell..... | 128 |

| | |
|--|-----|
| 5.6. Summary | 129 |
| CHAPTER 6: DISCUSSION | 130 |
| 6.1. Background overview..... | 130 |
| 6.2. HepG2 cells as a cell-culture model for lipid accumulation..... | 130 |
| 6.3. Effect of DGAT inhibitors on lipid accumulation..... | 131 |
| 6.4. Proteomic changes in HepG2 cells..... | 134 |
| 6.4.1. Lipid metabolism in FA-treated HepG2 cells | 135 |
| 6.4.2. Increased mitochondrial metabolism | 136 |
| 6.4.3. Calcium signalling between mitochondria and the ER | 137 |
| 6.4.4. Protein processing in the ER | 138 |
| 6.5. Improving proteome coverage | 139 |
| 6.5.1. Sample preparation and separation of peptides | 140 |
| 6.5.2. Effect of instrument sensitivity and peptide-spectra matching | 140 |
| 6.5.3. Statistical analysis | 143 |
| 6.6. Summary of the proteomics changes observed | 144 |
| 6.7. Conclusions | 146 |
| 6.8. Future work..... | 146 |
| REFERENCES | 149 |
| APPENDIX | 177 |

LIST OF FIGURES

| | |
|---|----|
| Figure 1: Schematic representation of progression in non-alcoholic fatty liver disease..... | 2 |
| Figure 2: Schematic representation of LD formation..... | 5 |
| Figure 3: Schematic representation of TAG synthesis | 9 |
| Figure 4: Phylogenetic tree of the <i>DGAT</i> families | 9 |
| Figure 5: Predicted membrane topology of the <i>DGAT</i> enzymes..... | 11 |
| Figure 6: Molecular structures of <i>DGAT1</i> inhibitors..... | 13 |
| Figure 7: Schematic representation of LTQ-Orbitrap™ XL mass spectrometer | 17 |
| Figure 8: Fragmentation of peptides using CID [194, 195] | 18 |
| Figure 9: Schematic representation of enzyme reaction catalysed by <i>DGAT</i> enzymes..... | 27 |
| Figure 10: Protein identification workflows setup in Proteome Discoverer (v. 1.4.1.14). | 33 |
| Figure 11: <i>DGAT</i> inhibitors used in this project..... | 39 |
| Figure 12: Absorbance of ORO at different wavelengths | 41 |
| Figure 13: Visualization of lipid droplets in HepG2 cells. | 42 |
| Figure 14: Lipid accumulation measured using ORO assay..... | 43 |
| Figure 15: Lipid accumulation in HepG2 cells treated with 1 mM FA mixture for with different time lengths..... | 44 |
| Figure 16: Effect of FA on metabolic activity assessed using MTT assay..... | 45 |
| Figure 17: HepG2 cells treated with high concentration of FA mixture | 46 |
| Figure 18: Effect of 100 µM A92 on HepG2 cells..... | 47 |
| Figure 19: Lipid accumulation in HepG2 cells treated with 100 µM A92..... | 47 |
| Figure 20: Lipid accumulation in presence of <i>DGAT</i> inhibitor A92. | 48 |
| Figure 21: Spectral analysis of A922500 stock solution. | 49 |
| Figure 22: Comparison of mobile phases used for the separation of <i>DGAT</i> assay reaction mixture by TLC..... | 50 |
| Figure 23: Inhibition of microsomal <i>DGAT</i> activity by A922500.. | 51 |
| Figure 24: Western blots of <i>DGAT</i> enzymes in HepG2 cells treated with 100 µM A92..... | 51 |
| Figure 25: Effect of PF04 and PF06 on <i>in vitro</i> microsomal <i>DGAT</i> activity..... | 53 |
| Figure 26: Effect of Mg ²⁺ ions on <i>DGAT</i> activity | 53 |
| Figure 28: ORO staining of HepG2 cells treated with different concentrations of PF04 | 54 |
| Figure 27: Effect of PF04 and PF06 on lipid accumulation in HepG2 cells..... | 54 |
| Figure 29: Effect of <i>DGAT</i> inhibitors on metabolic activity of HepG2 cells treated with 1 mM FA mixture for 6 h measured by MTT assay | 55 |
| Figure 30: Western blot for <i>DGAT1</i> in HepG2 cells treated with 100 µM PF04..... | 55 |
| Figure 31: Schematic representation of proteomics workflow..... | 57 |

| | |
|---|---------------|
| Figure 32: Comparison of lysis buffer used for sample preparation | 59 |
| Figure 33: LC gradient used for separation of peptides analysed in dataset LC1 | 60 |
| Figure 34: The pipeline tree used for identification of proteins listed in LC1 | 61 |
| Figure 35: Proteins identification for dataset LC1 in Proteome Discoverer | 62 |
| Figure 36: Identifications of proteins, peptides and MS/MS spectra in different treatment groups in Scaffold..... | 63 |
| Figure 37: Western blots for the validation of dataset LC1..... | 67 |
| Figure 38: Representative chromatogram for Top 6 method using standard LC gradient | 69 |
| Figure 39: Representative chromatogram for the modified Top 6 method using gradient 1. | 69 |
| Figure 40: Optimization of LC gradient using GOAT | 70 |
| Figure 41: Comparison of gradients used for Top 6 methods | 71 |
| Figure 42: Comparison of data-dependent MS/MS settings. | 72 |
| Figure 43: Comparison of Top N methods | 73 |
| Figure 44: Comparison of number of proteins identified by different PD workflows..... | 74 |
| Figure 45: Comparison of proteins identified by the different PD workflows | 75 |
| Figure 46: Comparison of protein databases used for protein identification | 76 |
| Figure 47: Number of proteins identified for each biological replicate of dataset LC2 in PD. | 77 |
| Figure 48: Comparison of datasets LC1 and LC2..... | 78 |
| Figure 49: Number of proteins identified for dataset LC3 in Proteome Discoverer. | 86 |
| Figure 50: Identifications of proteins, peptides and MS/MS spectra in different treatment groups in Scaffold..... | 87 |
| Figure 52: Distribution of proteins with respect to the co-efficient of variance for LC3. | 89 |
| Figure 51: Comparison of proteins identified in the biological replicates of dataset LC3 .. | Error! |
| Bookmark not defined. | |
| Figure 53: Protein identification for dataset LC4 in Proteome Discoverer. | 92 |
| Figure 54: Identifications of proteins, peptides and MS/MS spectra in different treatment groups in Scaffold..... | 93 |
| Figure 55: Comparison of proteins identified in each LC-MS/MS run within a group analysed for dataset LC4 | 96 |
| Figure 56: Distribution of proteins with respect to the co-efficient of variance for LC4. | 97 |
| Figure 57: Schematic representation of the role of C1-THF synthase (MTHFD1), Formimidoyltransferase-cyclodeaminase (FTCD) and 5'-phosphoribosylglycinamide transformylase (GART) domain of the trifunctional purine biosynthetic protein adenosine-3 in folate metabolism. | 114 |
| Figure 58: Protein enriched in KEGG pathway for protein processing in the ER..... | 118 |
| Figure 59: Protein quantification for dataset LC1..... | 121 |
| Figure 60: Calibration curves for Western blots | 122 |
| Figure 61: Comparison of proteins identified in datasets LC3 and LC4..... | 123 |

| | |
|--|-----|
| Figure 62: Western blots for validation of target proteins. | 126 |
| Figure 63: Quantification of Western blot bands..... | 126 |
| Figure 64: Detection of perilipin 2 after 2D separation.. | 128 |
| Figure 65: Detection of perilipin 2 in HepG2 cells..... | 128 |
| Figure 66: Lipid accumulation over 24 h in control cells..... | 178 |
| Figure 67: Lipid accumulation over 24 h in HepG2 cells treated with 0.25 mM FA mixture. | 179 |
| Figure 68: Lipid accumulation over 24 h in HepG2 cells treated with 0.5 mM FA mixture. ... | 180 |
| Figure 69: Lipid accumulation in HepG2 cells treated with 2 mM FA mixture. | 181 |
| Figure 70: Lipid accumulation in HepG2 cells treated with 3 mM FA mixture | 182 |
| Figure 71: Uncropped images of Western blots for target proteins of dataset LC1..... | 183 |
| Figure 72: Uncropped Western blots images of target proteins from dataset LC3 and LC4. | 184 |

LIST OF TABLES

| | |
|---|-----|
| Table 1: Description of primary antibodies used for Western blots | 23 |
| Table 2 : Composition of lysis buffers used for sample preparation..... | 29 |
| Table 3: Description of databases searched for protein identification. | 32 |
| Table 4: Proteins with significant changes in abundance from dataset LC1 | 64 |
| Table 5: Target proteins from dataset LC1 for Western blot validation..... | 66 |
| Table 6: Comparison of LC-MS/MS gradients..... | 69 |
| Table 7: List of proteins identified in dataset LC2 with $p < 0.05$ | 80 |
| Table 8: Proteins with co-efficient of variance $\leq 10\%$ in dataset LC3..... | 90 |
| Table 9: QV and exclusive peptide counts for proteins with co-efficient of variance $\leq 10\%$ in dataset LC3..... | 91 |
| Table 10: Proteins with significant changes in abundance in dataset LC4 | 94 |
| Table 11: Proteins of dataset LC4 used for functional analysis | 98 |
| Table 12: Enrichment of protein identified in LC4 in GO categories for the comparison of FA-treated group to the control group. | 106 |
| Table 13: Enrichment of protein identified in LC4 in GO categories for the comparison of inhibitor-treated group to FA-treated group..... | 108 |
| Table 14: Enrichment of protein identified in LC4 in GO categories for the comparison of inhibitor-treated group to control group..... | 108 |
| Table 15: Enrichment in KEGG pathways..... | 109 |
| Table 16: List of proteins enriched in KEGG pathway hsa01100 for metabolic processes ... | 110 |
| Table 17: Proteins enriched in KEGG pathway hsa04141 for protein processing in the endoplasmic reticulum | 116 |
| Table 18: Proteins from dataset LC3 and LC4 selected for validation | 120 |
| Table 19: Comparison of studies describing lipid accumulation reported in HepG2 cells after 24 h treatment time..... | 131 |
| Table 20: File format used by software for proteomics analysis..... | 177 |
| Table 21: Organization and description of supplementary content in the electronic form.. | 177 |

ABBREVIATION

| | |
|--------------------------|---|
| A92 | DGAT1 inhibitor A922500 |
| ANOVA | Analysis of variance |
| BSA | Bovine serum albumin |
| CHAPS | 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate |
| CID | Collision-induced dissociation |
| CV | Co-efficient of variance |
| DGAT | Diacylglycerol acyltransferase |
| DMEM | Dulbecco's modified Eagle's medium |
| DMSO | Dimethyl sulfoxide |
| DTT | 1, 4-dithiothreitol |
| ER | Endoplasmic reticulum |
| FA | Fatty acids |
| FBS | Foetal bovine serum |
| FDR | False discovery rate |
| GO | Gene Ontology |
| GOAT [®] | Gradient Optimization and Analysis Tool |
| HBSS | Hank's balanced salt solution |
| ICAT | Isotope-coded affinity tags |
| iTRAQ | Isobaric tag for relative and absolute quantitation |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| LC-MS/MS | Liquid chromatography coupled with tandem mass spectrometry |
| LD | Lipid droplet |
| MS | Mass spectrometry |
| MTT | 3-(4, 5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide |
| NAFLD | Non-alcoholic fatty liver disease |
| NASH | Non-alcoholic steatohepatitis |
| NBD-palmitoyl coenzyme A | (N-((7-Nitro-2-1, 3-benzoxadiazol-4-yl)-methyl) amino) palmitoyl coenzyme A |
| ORO | Oil red O |
| PBS | Phosphate-buffered saline |
| PD | Thermo Scientific™ Proteome Discoverer™ software (version 1.4.1.14) |
| PF04 | DGAT1 inhibitor PF04620110 |
| PF06 | DGAT2 inhibitor PF06424439 |
| PLIN2 | Perilipin 2 |
| QV | Quantitative value |
| R _f | Retention factor |
| ROS | Reactive oxygen species |
| RT | Room temperature |
| SILAC | Stable-isotope labelling by amino acids in cell culture |
| SDS | Sodium dodecyl sulphate |
| SNP | Single nucleotide polymorphism |
| TAGs | Triglycerides |
| TLC | Thin layer chromatography |
| Tris | Tris(hydroxymethyl)aminomethane |
| UniProtKB | UniProt Knowledgebase |
| UPR | Unfolded protein response |
| WebGestalt | WEB-based GENE SeT Analysis Toolkit |

CHAPTER 1: INTRODUCTION

1.1. Conception of this project

Non-alcoholic fatty liver disease (NAFLD) is a disease state that starts with steatosis characterised by lipid accumulation in hepatocytes and progresses sometimes irreversibly to non-alcoholic steatohepatitis (NASH) to damage as a result of hepatic steatosis with inflammation and subsequent fibrosis and cirrhosis [1]. While all patients with NAFLD show simple steatosis, only some patients develop chronic hepatic inflammation which leads to increased mortality [2, 3]. This project aims to characterize the proteomic changes that accompany lipid accumulation in the cell.

1.2. Prevalence of non-alcoholic fatty liver disease

As of 2014, the World Health Organization has estimated the number of overweight persons to be 1.4 billion adults and about 600 million of these were categorized as obese [4]. Modern lifestyle is marked by increased caloric intake, decreased mobility, (and a paradoxical demand for slimmer body silhouettes) and the consequence of this is a widening waistline which, at the cellular level, is a result of lipid accumulation in adipose tissue [5]. The accumulated lipids are largely triglycerides (TAGs) sequestered into membrane-bound cytosolic lipid droplets (LDs)[6]. Morbidities more commonly referred to as the metabolic syndrome are known to accompany ectopic lipid accumulation in the liver, skeletal muscle, and pancreas [7]. In this context, NAFLD is widely considered as the manifestation of the metabolic syndrome in the liver and is increasingly being recognized as a risk-factor for end-stage liver diseases [8, 9]. Systematic review and meta-data reports on the epidemiology and natural history of NAFLD pegs the prevalence of this problem to be between 10 - 45 % depending on the population and the method used to establish the diagnosis [10, 11]. The prevalence of NAFLD in New Zealand has not been examined yet but considering 31 % of the adult and one in nine children between the ages of 2 and 14 years have been reported obese in New Zealand Health Survey 2014/15 [12], the NAFLD burden could be an emerging health care issue [13]. The close association between obesity and NAFLD [14, 15] is likely to increase these numbers as the obesity pandemic reaches new global highs [4].

1.3. Progression from NAFLD to NASH

NAFLD has been defined as a complex, non-linear, metabolic liver disease with two distinct phases – simple steatosis and NASH [16]. Simple steatosis is considered to be the onset of

NAFLD and is marked by excessive lipid accumulation in the liver in the absence of causes such as drugs, toxins, excess alcohol intake (>30 g/day in men and >20 g/day in women) and hepatitis due to viral infection or autoimmune problem [17]. NASH is characterised by excessive fat accompanied by liver cell injury and inflammation [18]. Simple steatosis has often been described as benign [19] with most patients not progressing to NASH [20]. Of those who do develop NASH, not everyone incurs progressive liver damage to develop fibrosis and necrosis [21]. About 25 - 40 % of patients with NASH show disease progression with increased risk of hepatocellular carcinoma even without cirrhosis (Figure 1, [22-24]). There has also been a suggestion to consider NASH and simple steatosis as independent but coinciding events since patients with NASH have been known to have inflammation without any or much steatosis [16, 25].

Improved diagnosis for distinguishing NASH from simple steatosis would be the key to distinguishing between the two. Non-invasive techniques like imaging and detection of biomarkers have been studied as tools to predict steatosis, NASH and fibrosis but each has its limitations in being able to diagnose the problem [26]. Liver biopsy is the gold standard of diagnosis and criteria used for histological scoring of the biopsy tissue are still being fine-tuned [27-29]. The intrusiveness of the biopsy-led diagnosis procedure and the asymptomatic nature of simple steatosis frequently make detection of NAFLD incidental [30]. Even when the diagnosis is made, the biopsy-led approach relies heavily on histology which provides very limited information about disease progression [31]. Although risk

Non-alcoholic fatty liver disease

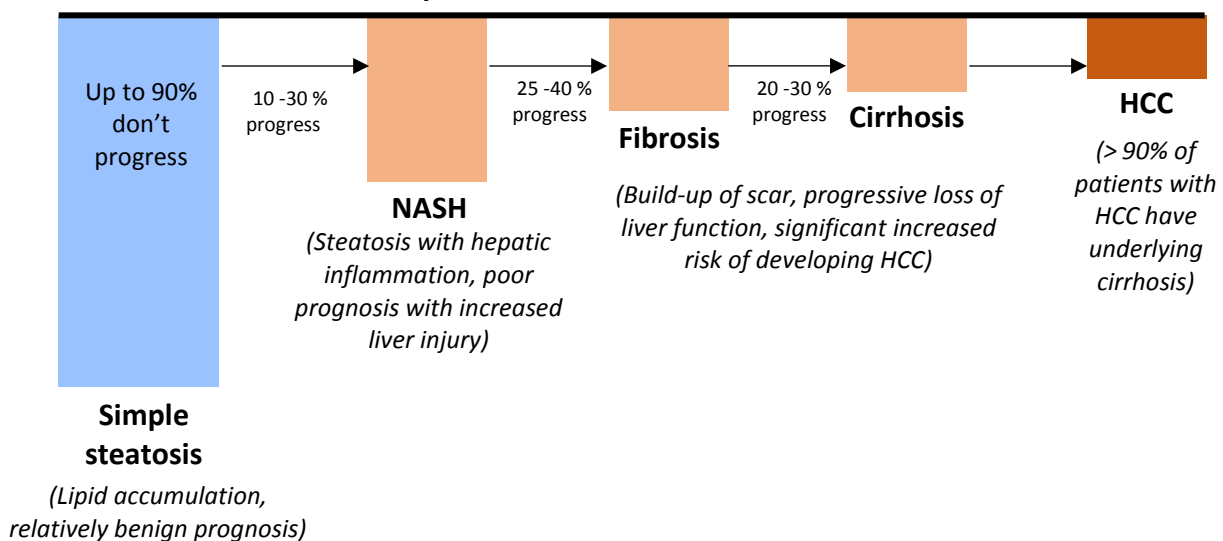


Figure 1: Schematic representation of disease progression in non-alcoholic fatty liver disease. Most patients with simple steatosis do not progress to steatohepatitis (NASH) and of the patients with NASH, less than half progress to irreversible liver damage in the form of fibrosis and cirrhosis [22-25]. However, cirrhosis significantly increases the risk of hepatocellular carcinoma (HCC).

factors like lifestyle and genetic predisposition are well-defined disease determinants [32, 33], the role and interplay of complex factors in disease progression at the cellular level is still being explored [34, 35].

1.4. Factors affecting disease progression

Drawing on the commonalities between the NAFLD and its alcohol-induced variant, C. P. Day proposed the 'two-hit' theory in which the excess lipid accumulation in the liver is the first hit which when left untreated, escalates to NASH and fibrosis as a result of the second hit [36]. The second hit is described as a combination of oxidative stress, endotoxin-mediated cytokine release and immunologically mediated mechanisms [36]. Further investigations into the molecular mechanisms of NAFLD have led to the notion of 'multiple, parallel hits' [37]. These hits, whether sequential or parallel, are brought about by complex interactions between intra-cellular factors affecting the hepatocytes and the environmental factors, host genetics and signals from other organs that affect the liver.

The characteristic feature of NAFLD is the increased accumulation of lipids in the hepatocytes and non-esterified fatty acids (FA) are the biggest contributors to this accumulation [38]. Almost two-thirds of the FA incorporated in the accumulating TAG are considered to be delivered to the liver on account of excessive lipolysis in the adipose tissue [39]. The environmental factor that contributes the most to this process is the dietary intake and the composition of the diet. This is evident in the high correlation between obesity and NAFLD [40]. While excessive consumption of high calorie diet is the most common underlying risk factor, inclusion of high amount of certain types of FA and fructose in the diet also increase the risk of developing NAFLD. The lipotoxicity of FA depends on whether it is saturated and unsaturated [41] and diets rich in saturated FA tend to be more harmful than those with monounsaturated or polyunsaturated FA [42]. Most of the naturally occurring unsaturated FA are *cis* isomers, but in partially hydrogenated vegetable oils, they may undergo *cis-trans* isomerization to form *trans*-FA that tend to have physical properties similar to saturated FA and therefore have similar deleterious effects when consumed [43]. The position of the double bond on the FA chain also seems to play a role in the toxicity they exhibit – the patients with NAFLD tend to have more *n-6* FA (the sixth carbon-carbon bond from the methyl end is a double bond) than *n-3* FA in the liver [44, 45]. Besides FA, diets with high fructose have also been shown to increase lipid accumulation in non-adipose tissues and increased the intrahepatic lipid content [46, 47].

The extent to which the high-calorie diet affects the individual is also likely to be affected by genetics. While no individual genetic locus has been identified as the cause of the problem, a genome-wide association study has shown some polymorphisms in the patatin-like phospholipase domain containing 3 gene increases the susceptibility of obese children and adolescents to diet-induced hepatic steatosis [48]. Another study has shown that ethnicity may as well be a part of the contributing factors – African-American children with obesity were reported to have the same extent of lipid accumulation in the liver as children of Caucasian or Hispanic roots but the extent of hepatocyte cell death was lower in the first cohort than the other two [49]. Given that disease progression varies in speed and severity among the NAFLD patients, it has been suggested that progression from simple steatosis to NASH may depend on how long and how strongly the adaptive mechanisms induced by the lipid accumulation can last [50]. Hepatocyte cell death and activation of the innate immune system have been shown to play an important roles in the progression of simple steatosis to NASH in both animal models and humans (reviewed in [51]).

1.5. Molecular mechanism of disease progression

Accumulation of TAG-rich LDs in the cytoplasm is a characteristic feature of NAFLD at the cellular level. The metabolic processes that contribute to this accumulation include *de novo* synthesis of FA in the cytoplasm, carbohydrate metabolism and FA breakdown in the mitochondria. The enzymes of the Kennedy pathway mainly localized on the endoplasmic reticulum (ER) membrane carry out the sequential esterification of acyl-CoA to the glycerol-3-phosphate backbone to produce TAG. The mitochondria and the ER are therefore also the organelles most affected by the lipid accumulation. While the lipid accumulation is considered as the ‘first hit’, the ‘second hit’ often manifest as ER stress and/or oxidative stress originating in the mitochondria.

1.5.1. First hit: Accumulation of triglycerides in lipid droplets

Although excessive lipid accumulation marks the onset of the disease, the sequestering of the FA into neutral TAG compartmentalized in LD has been mooted as an adaptive response to manage the lipotoxicity of non-esterified FA [41, 52, 53]. Although the exact mechanism of how the TAG would be packaged into the LDs is not known, according to one model, the formation of LDs starts with the deposition of neutral lipids between the leaflets of the ER membrane and as the droplet grows, it buds towards the cytoplasm, pinching away with it a leaflet of the ER bilayer that now becomes the phospholipid monolayer of the LD (Figure 2,

[54]). These nascent LDs may grow into larger LDs either through coalescence of the nascent droplets or through synthesis of neutral lipid [55, 56]. The size of LD is managed by the concerted efforts of lipases, autophagy and mitochondrial fusion [57, 58]. Proteins like those of the Kennedy pathway are known to translocate from the ER membrane to the LD – these translocation are thought to be facilitated by ER ‘bridges’ that provide a link between the two organelles [59]. A protein complex formed by ADP-ribosylation factor 1 and Coat Protein Complex I (Arf1/COPI) has been proposed as the facilitator of such connections [60, 61]. The presence of such proteins also allows for localised TAG synthesis in LD and so a cell may have two distinct LD subpopulations depending on the localization of these protein [59]. Given the dynamic role of LDs in fat homeostasis (reviewed in [62]) and the proximity of LDs to the ER, it is becoming increasingly evident lipid accumulation would have major effects on cell function and increase risk of developing metabolic diseases [63].

The close connections between the LDs and ER have also been investigated for the better understanding of ER stress – numerous reports suggest ER stress leads to lipid accumulation and that the LDs may also act as buffers for FA and phospholipids that help maintain ER homeostasis [64-67].

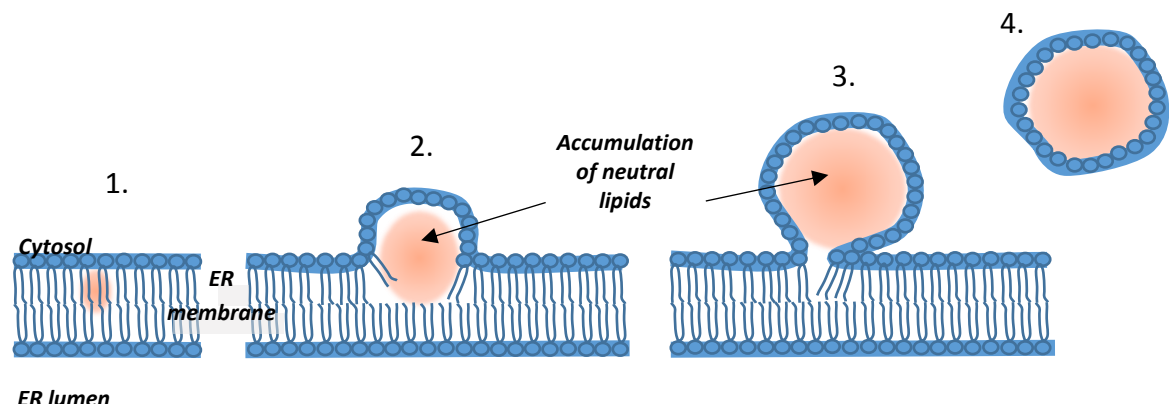


Figure 2: Schematic representation of LD formation. 1. Synthesis of TAG is catalysed by enzymes on the ER membrane. 2. Accumulation of TAG between the two leaflets of the ER bilayer leads to the distended of the cytoplasmic leaflet. 3. The cytoplasmic leaflet continues to bulge as more neutral lipids are deposited in the nascent LD before budding takes place. 4. The cytoplasmic leaflet of the ER membrane encloses the neutral lipids and forms a LD [55].

1.5.2. ER stress

In addition to being the site of TAG synthesis, the ER is also the protein processing hub of the cell. The protein processing includes post-translational modifications such as N-linked glycosylation, disulphide linkages and quality control to ensure the correct folding of the protein. Lipid accumulation observed in NAFLD has been reported to disrupt this critical ER function. This disruption originates in the capacity of saturated FA to induce various

components of unfolded protein response (UPR), an adaptive mechanism that allows the ER to cope with stress [68]. UPR is activated through three key proteins – inositol-requiring kinase 1 (IRE1), protein kinase-like ER kinase (PERK) and activating transcription factor 6 (ATF6). They are transmembrane proteins localized in the ER membrane and they serve as sensors of ER stress. In an unstressed cell, the three sensors are bound by 78kDa glucose regulatory protein (Grp78), an ER resident chaperone that also serves as a regulator of UPR. The binding of Grp78 to IRE1 and PERK prevents the oligomerization and auto-phosphorylation of these proteins and thereby prevents the activation of UPR [69]. Binding of Grp78 to ATF6 retains the latter in the ER lumen; dissociation of the two leads to the translocation of ATF6 to the Golgi. In the Golgi, ATF6 is modified into an active transcription factor that can activate genes associated with UPR [70]. Through the various downstream proteins and genes affected by the activation of IRE1, PERK and ATF6, the UPR functions to restore ER homeostasis.

Maintenance of high Ca^{2+} concentration in the ER lumen is also an important part of ER homeostasis. The ER is a major site for Ca^{2+} storage in the cell - while the concentration of Ca^{2+} in the cytosol has been estimated to be about 0.1 μM , the Ca^{2+} concentration in the ER lumen is in the millimolar range [71]. This pool of Ca^{2+} facilitates electrostatic interaction between proteins as is evident by the calcium-dependency of ER chaperones like calnexin and calreticulin [72]. Mobilization of this Ca^{2+} reservoir also plays a role in intracellular signalling like the pro-apoptotic trigger of Bak/Bax regulated release of Ca^{2+} from the ER [73]. In fact, the regulation of Ca^{2+} concentrations and signalling between the ER, mitochondria and cytoplasm has been proposed as an additional trigger of apoptosis [74]. In addition to the Ca^{2+} mediated signals, failure of UPR to cope with the stress also leads to pro-inflammation and pro-apoptotic signals like activation of c-Jun terminal kinase – the activation of which differentiates NASH patients from those of NAFLD with simple steatosis [75, 76].

1.5.3. Mitochondrial dysfunction

Metabolic adaptations including increased oxidation of FA in the mitochondria have been proposed as coping mechanism setup to limit lipid accumulation seen in NAFLD [77]. Patients with NAFLD tend to have mega mitochondria with inclusion bodies and loss of cristae [78]. This has been proposed as an adaptive response to accommodate the increased transport of FA to the liver also reported in patients with NAFLD [79]. Such mitochondrial

dysfunction can be a major source of reactive oxygen species (ROS) which can propagate the deleterious cascade of harm through lipid peroxidation [80]. ROS are highly reactive and have short half-lives [81]. They are capable of initiating lipid peroxidation that leads to the formation of aldehyde by-products like trans-4-hydroxy-2-nonenal and malondialdehyde that have much longer half-lives and can amplify the effects of oxidative stress [82, 83].

The aggravating factors described in this section – triggering UPR, imbalance of pro- and anti-apoptotic signals, signal transduction by Ca^{2+} and mitochondrial dysfunction are all in themselves complex biological events with a number of participating and affected proteins reviewed in [84] and [51]. Given the global impact of these disruptions and the number of proteins that could potentially contribute to disease progression, the current study has used a proteomics approach to study the changes that accompany lipid accumulation.

1.6. Role of DGATs in triglyceride synthesis

Although accumulation of lipid is a feature of NAFLD and as such poses risks to well-being, the intended evolutionary purpose of neutral lipids like TAG is to function as energy-dense storage molecules that allow a cell to survive during periods of reduced nutrient availability by providing membrane building blocks like FA or sterols [85]. Precursors like non-esterified FA, if not sequestered, have adverse effects on membrane integrity [86] and fatty alcohols such as sterols, retinols, and diacylglycerols may set off a plethora of reactions in their capacity as secondary messengers of transcriptional activators [87]. This makes the turnover of TAG a critical factor in maintaining health [88]. The sequestering of these molecules into neutral TAG compartmentalized in LD has been mooted as an adaptive response to manage the lipotoxicity [41, 52, 53, 89]. Given the role of TAG-rich LD in NAFLD, diacylglycerol acyltransferase 1 and 2 (DGAT1, DGAT 2) – enzymes that add an acyl moiety to diacylglycerol to form TAG – have been considered possible targets of intervention. TAGs are synthesized through step-wise esterification of acyl moieties to a glycerol-3-phosphate backbone via the Kennedy pathway [90], supplemented by the monoacylglycerol pathway that re-esterifies hydrolysed TAG (Figure 3, [91]). Both pathways are used by hepatocytes [92].

1.6.1. *DGAT* genes

DGAT1 and *DGAT2* are genes expressed ubiquitously in tissues such as intestine, adipose tissue and liver. Despite the proteins catalysing the same biochemical reaction, the DGAT genes share no significant sequence homology and belong to different gene families [93]. The two genes have evolved separately and lend their names to their respective gene

families [94]. Figure 4 generated using ClustalW and PHYLIP shows the phylogenetic tree for these genes [95, 96]. Human *DGAT1* has been mapped to human chromosome 8 by fluorescence *in situ* hybridization [97] and is now known to be located at position 8q24.3 [98]. The *DGAT1* family includes the gene for DGAT1 and the two genes *ACAT1* and *ACAT2* that encode their respective enzymes with acyl-CoA: cholesterol acyltransferase activity [99]. In fact, the similarity between the *DGAT1* and *ACAT* genes was used to identify DGAT1 through screening of expressed sequence tag clones homologous to the conserved C-terminus of ACAT [97]. Synthesis of TAG in *Dgat* ^{-/-} mice suggested the presence of another gene product with DGAT activity and this led to the identification of gene encoding a protein then named DGAT2. This gene had no sequence homology with any known genes of the *DGAT1* family and the newly identified gene was designated as DGAT2 [100]. The human *DGAT2* gene located on chromosome 11q13.5 [98] and it is part of the *DGAT2* gene family that includes three monoacylglycerol O-acyltransferases encoded by *MOGAT1*, *MOGAT2* and *MOGAT3* [101, 102], and two acyl-CoA: wax alcohol acyltransferases encoded by *AWAT1* and *AWAT2* [103]. The *DGAT2* family is suggested to have arisen by duplications of ancient gene tandem repeats via inter-chromosomal or genome duplications as well as further local (single) gene duplication and loss events [104]. The homology in nucleotide sequences is also seen in the amino acid sequences - human MGAT3 possesses significant DGAT activity and has an amino acid sequence which is 49, 44, and 46 % identical to that of DGAT2, MGAT1 and MGAT2 respectively [105-107]. Given the role of the DGAT enzymes in lipid metabolism, the mutations and polymorphisms in the DGAT genes have been studied for their role in predisposition to obesity and metabolic syndrome.

A German study assessing the benefits of a lifestyle intervention programme aimed at reducing adiposity and preventing Type 2 diabetes has reported that single nucleotide polymorphism (SNP) in *DGAT2* dampened the effects of the intervention. Although all participants showed a similar increase in insulin sensitivity, individuals with SNP rs1944438 in *DGAT2* showed a smaller decrease in liver fat than those without the SNP [108]. Polymorphisms identified in human *DGAT1* have not been conclusively linked to beneficial effects like lower body mass index [109, 110] but a splice site mutation in the gene has been shown to be a rare cause of a congenital diarrheal disorder, an extremely rare but severe recessive disorder characterised by hyperlipidaemia and protein-losing enteropathy and very

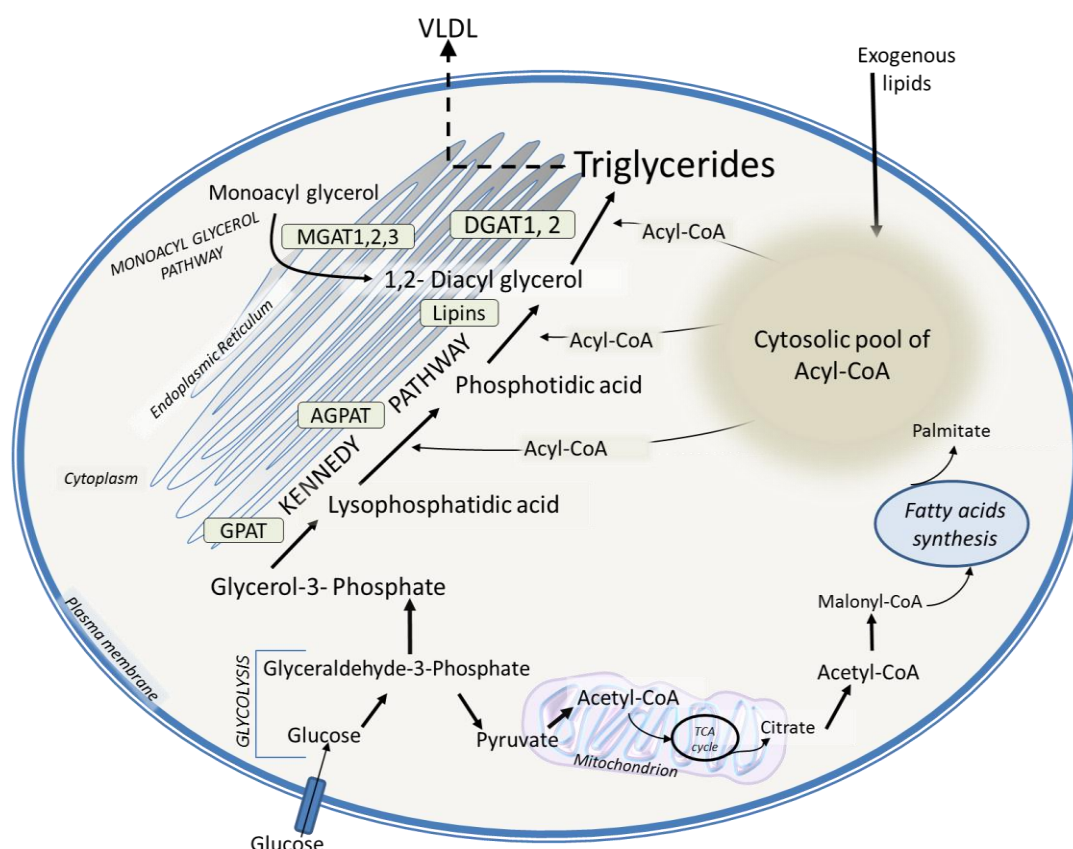


Figure 3: Schematic representation of TAG synthesis. Enzymes of the Kennedy pathway the monoacylglycerol pathway facilitate sequential esterification of the glycerol-3-phosphate derived from the glycolysis precursors. The acyl-CoA esterified into TAG may be originate in exogenous FA or be synthesized *de novo* in the cell from acetyl CoA. Key- GPAT: Glycerol-3-phosphate acyltransferase; AGPAT: Acyl-glycerol phosphate acyltransferase; DGAT: Diacylglycerol O-acyltransferase; MGAT: Monoacylglycerol O-acyltransferase.

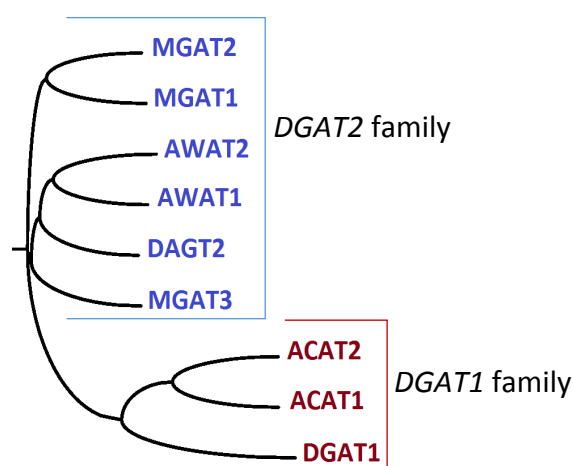


Figure 4: Phylogenetic tree of the DGAT families. The protein sequences of each protein was submitted to Biology Workbench (<http://workbench.sdsc.edu>) in fasta format for multiple sequence alignment using the ClustalW algorithm [82]. The aligned sequences were converted into the phylogenetic tree using Drawgram algorithm which is part of PHYLIP (Phylogeny Inference Package, version 3.5c, distributed by the author J. Felsenstein of University of Washington, Seattle, USA) [83]. Protein sequences were retrieved from UniProt KB using accession numbers P35610, O75908, Q58HT5, Q6E213, O75907, Q96PD7, Q96PD6, Q3SYC2 and Q86VF5. Key- MGAT: Monoacylglycerol O-acyltransferase, AWAT: Acyl-CoA wax alcohol acyltransferase, DGAT: Diacylglycerol O-acyltransferase, ACAT: Acyl-coenzyme A: cholesterol acyltransferase.

poor chances of survival [111]. Polymorphisms identified in human *DGAT1* have not been conclusively linked to beneficial effects like lower body mass index [109, 110] but a splice site mutation in the gene has been shown to be a rare cause of a congenital diarrheal disorder, an extremely rare but severe recessive disorder characterised by hyperlipidaemia and protein-losing enteropathy and very poor chances of survival [111]. Cholestyramine and intravenous albumin were found helpful in correcting the symptoms [111]. Stunted growth was also reported in calves homozygous for a naturally-occurring mutation that lead to a truncated protein without the region encoded by exon 16 of bovine *DGAT1*. This truncated enzyme is incapable of esterifying FA to the diacylglycerol and calves homozygous for this mutation require intravenous supplementation of essential and unsaturated lipids to restore the growth [112].

1.6.2. Membrane topology

DGAT1 and DGAT2 are integral membrane proteins and about 40 % of the total amino acid residues in these proteins are hydrophobic in nature [113]. The amino acid sequences for human DGAT1 and DGAT2 were retrieved from UniProt Knowledgebase (UniProtKB) and, using the hidden Markov model, the possible number and orientation of the transmembrane domains were predicted for these proteins (Figure 5) [114]. Human DGAT1 (O75907) has been described as a multi-pass membrane protein [115]; the hidden Markov model for membrane topology predicted nine transmembrane domains for this protein. Similar predictions have been made for murine DGAT1 but experimental verification of the topological model suggests that the protein may have three transmembrane domains instead of the eight predicted by the topology models generated in the McFie lab [116]. The McFie model suggests that the three transmembrane domains are oriented such that the N terminus is in the cytosol while the C terminal region and almost half of the amino acid residues of the proteins, including the conserved histidine residue, are in the ER lumen [116]. The experimentally verified membrane topology of DGAT2 on the other hand, comprises of two transmembrane domains connected by a short loop in the lumen or alternatively, a single hydrophobic domain that embeds itself in the membrane [117]. Such a topology would orient both termini of DGAT2 towards the cytosol and the C-terminal domain distal to the transmembrane domains would form the bulk of the protein [117].

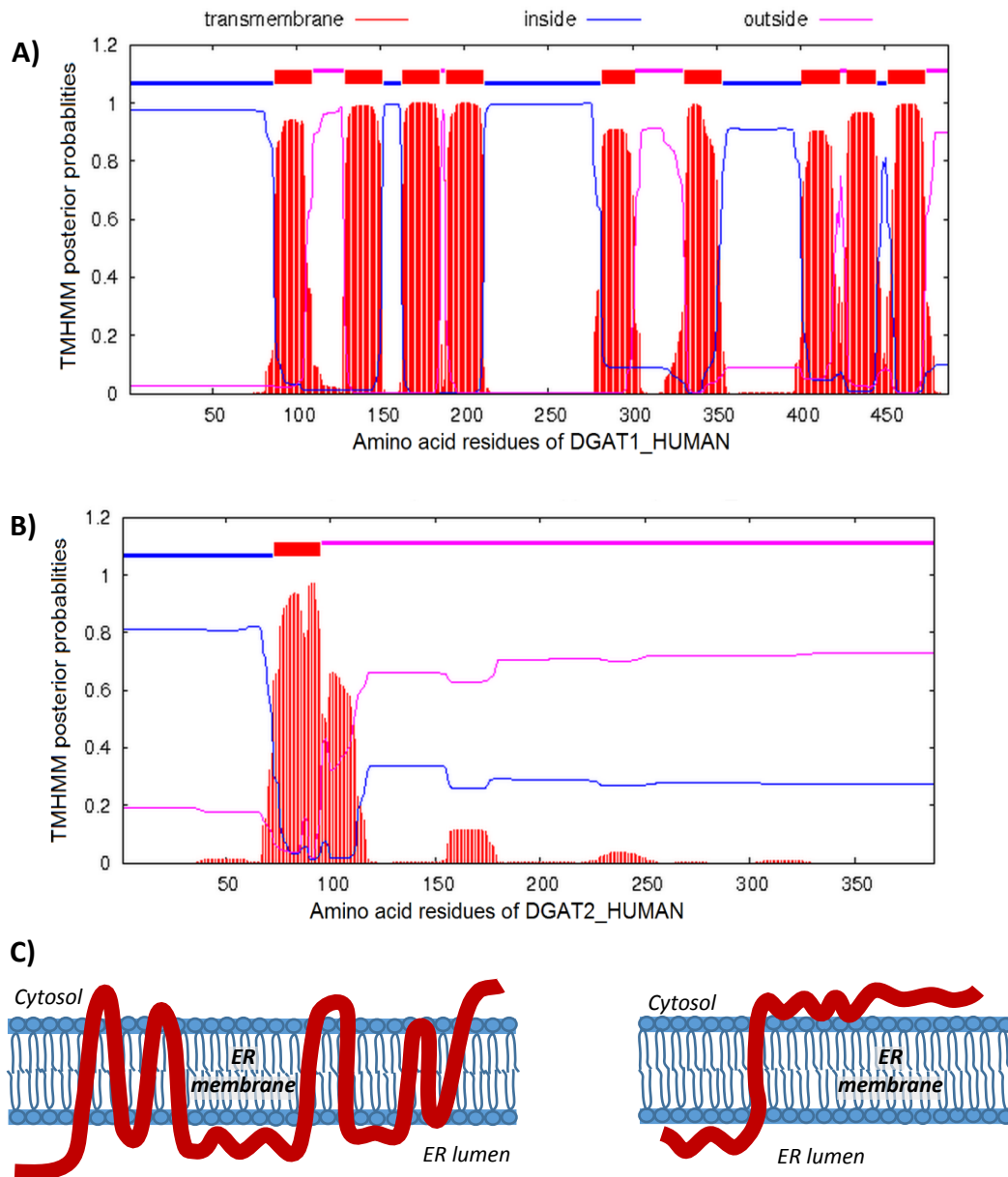


Figure 5: Predicted membrane topology of the DGAT enzymes. The protein sequence for human DGAT1 (O75907) and DGAT2 (Q96PD7) were retrieved from the SwissProt database in the fasta format and analysed using the transmembrane hidden Markov model tool available on www.cbs.dtu.dk/services/TMHMM [96]. The model plots probabilities for each residue residing in a trans-membrane domain and is not a prediction of the most probable overall protein structure. The model predicts nine domains for DGAT1 (**A**) and one major trans-membrane domain for the DGAT2 (**B**). **C**) Diagrammatic representation of the DGAT enzymes based on the orientation and trans-membrane domains predicted by the model.

1.6.3. Source of substrate

DGAT1 and DGAT2 are integral membrane proteins localized in the ER; however, DGAT2 is also known to translocate to the LD membrane [118]. Also unique to DGAT2 are the interactions of this protein with acyl-CoA synthetase FATP1 [56] and MGAT2 [119] to promote TAG synthesis and enlargement of the LD. Within the ER, DGAT2 is found in the

vicinity of stearoyl-coenzyme A desaturase 1 (SCD1). SCD1 facilitates the formation of a double bond by removing hydrogen from the saturated fatty acyl-CoA formed through *de novo* FA synthesis. The proximity of SCD1 and FATP1 to DGAT2 may facilitate channelling of *de novo* FA between the enzymes [120]. Also, disruption of the *Scd1* gene leads to reduced levels of hepatic TAG, a deficiency that cannot be corrected by dietary supplementation of mono-saturated FA [121]. This points to the possibility of different substrate pools of endogenous FA and exogenous FA that may be located in different microenvironments of the ER where they are utilized for different purposes [122]. Owing to its specific localization in the ER, DGAT2 may associate with the pool of *de novo* FA that is not available to DGAT1, which is localised in a different place in the ER [120, 121].

Complementing the localization studies, experiments with *Dgat* knockout mice and labelled substrates have provided further evidence of source-dependent preferential use of FA by the two DGAT enzymes. Liver-specific *Dgat1*^{-/-} mice are resistant to hepatic steatosis induced by the exogenous FA available in high-fat diets or released from the adipose tissue after prolonged fasting [123]. However, DGAT1 inhibition does not prevent hepatic steatosis caused by lipodystrophy and liver X receptor activation which increase *de novo* lipogenesis in the liver [123]. Further investigations using radiolabelled substrates have confirmed that DGAT1 preferentially esterifies exogenous FA, whereas DGAT2 uses endogenous FA derived from *de novo* lipogenesis [124]. Also, it has been shown that DGAT1 is more active at higher (> 200 μ M) oleoyl-CoA concentrations associated with an influx of exogenous FA in an *in vivo* system, whereas DGAT2 is more active at lower oleoyl-CoA concentrations (up to 50 μ M) [100].

1.6.4. DGAT Inhibitors

Knowledge about DGAT activities and the role of individual enzymes has spurred interest in developing molecules capable of inhibiting the enzymes and ameliorating conditions like obesity, diabetes and NAFLD that are characterized by the build-up of TAG. Effects of knocking out *Dgat1* in mice include resistance to obesity, increased insulin sensitivity and even resistance to diet-induced hepatic steatosis [125-127]. This has led interest in the development of small-molecule DGAT1 inhibitors of multiple chemical classes and many have been shown to ability to improve metabolic parameters in preclinical models [128]. Since most of these drugs are orally administered, the gastro-intestinal system is where the molecules affect the DGAT1 inhibition. Effects include delayed gastric emptying following an

oral lipid challenge and upset post-prandial balance of hormones plasma glucagon-like peptide-1 [129, 130]. Similar effects were reported in human clinical trials for Novartis molecule LCQ908 (pradigastat); Pfizer molecule PF04620110, AstraZeneca molecule AZD7687 (Figure 6, [131-134]).

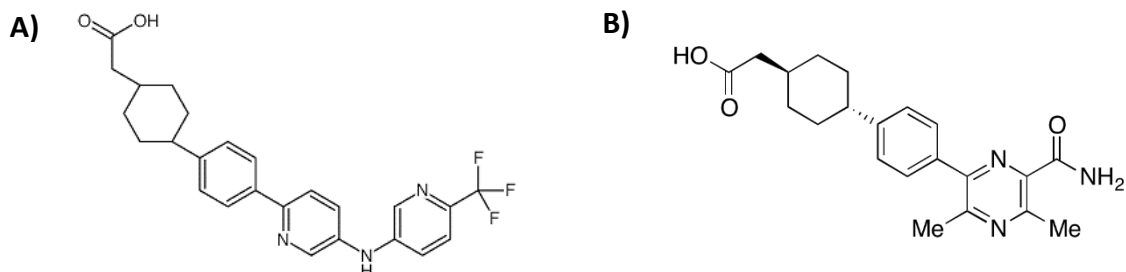


Figure 6: Molecular structures of DGAT1 inhibitors. A) Pradigastat and B) AZD7687. The images used are part of the suppliers' catalogue.

The interest in DGAT2 has been limited by the apparent indispensability of the gene. Knocking out *Dgat2* in mice leads to lethal lipopenia and decreased skin permeability [135]. Decreased *DGAT2* mRNA in the sebaceous glands of patients with psoriasis also highlights the vital role of this gene and its product [136]. This notwithstanding, efforts have been made to inhibit DGAT2 with anti-sense oligonucleotides and generate liver-specific TAG reduction in mice [137]. While these results have not attracted the same enthusiasm as the DGAT1 inhibitors, the distinct role played by the two isozymes in the cell merits the use of selective inhibitors that can help further differentiate between the two. Wurie *et al.* have successfully used isozyme specific small-molecule inhibitors to distinguish between the activities of the two enzymes and even suggest that “DGAT1 may act up-stream of DGAT2” [138]. Given that new molecules are being screened for DGAT2-specific inhibition [139, 140], it may be possible in the future to further differentiation of the two isozymes, especially in functions that could reduce adverse effects of DGAT1 inhibition. This project describes the use of DGAT inhibitors to limit lipid accumulation induced by excess FA and the proteomic changes that allow the cell to cope with these metabolic events.

1.7. Models used to study NAFLD

1.7.1. Clinical studies in humans

Several clinical trials have used samples from liver biopsies to extract information about gene expression in individuals with and without NAFLD – these include testing for known

lipogenic targets, gene profiling to identify new ones and comparisons of methylomes and transcriptomes [141-144]. However, an invasive liver biopsy is poorly suited as a diagnostic test because of its expense and risks of complications [145, 146]. Additionally, imaging techniques or biopsies give limited information about mechanisms involved in disease progression.

Early studies have used radiolabelled tracers to measure the flux of FA and metabolites like TAG and ketone bodies [147, 148]. More recent approaches like positron emission tomography [149] and stable-isotope-labelled FA have proved safer and have provided insight into the source of FA and the compartments they are assimilated into [39, 79]. While these studies have furthered the understanding in the field, all clinical studies in humans have an inherent limitation of variable compliance and poor availability of subjects. Diversity of genotypic and demographic factors may also limit the inferences in a cohort that can be made in these studies during the brief period the subjects are under observation.

1.7.2. Animal models of NAFLD

The use of the laboratory mouse, *Mus musculus* allows control over factors such as genetic heterogeneity, diet and environmental variation [150, 151]. Variables such as age and gender can also be controlled for. An ideal animal model of NAFLD would demonstrate hepatic histological and pathological features seen in the diseased state in humans. The mouse models however, have a limited capacity to reflect the metabolic changes of NAFLD in humans. In fact, no single animal model has been successful in reflecting all aspects of the disease. Models are usually chosen for a specific purpose – while some provide good models for liver steatosis in the context of the metabolic syndrome, others are better models of hepatic inflammation and fibrosis [152, 153]. Moreover, the aspects of the disease that are represented need to be put into the human context since hepatic TAG metabolism tends to vary between species [154] and the composition of dietary intakes may skew the interpretation of the final results [155].

The aim of this project is to study the proteomic changes that happen within the cell as a consequence of lipid accumulation like that observed in steatosis. A cell culture-based model allows the study of lipid metabolism with stricter control of substrate influxes and possible molecules that could affect homeostasis.

1.7.3. Cell-culture based models

Liver is central to lipid metabolism and is attributed with performing diverse functions like detoxification and processing drugs and non-drug xenobiotics [156-158], metabolizing macromolecules and vitamins [159-161]; and secreting of bile, cholesterol and lipoproteins [162]. At any given point in time, the liver tends to be under the influence of a variety of factors [163-165]. An effort to understand intracellular changes in the liver would be helped by a reduction in the complexity of such extra-hepatic influences. Although it has the limitation of not being a perfect representative of a hepatocyte in liver tissue, a cell-culture based model maintained in a medium of known composition allows for assessment of causal links of the effects seen in the cell. Hepatocytes are the functional units of the liver and constitute almost 80 % of the total liver volume [166]. Based on this, primary human hepatocytes isolated by the collagenase perfusion method described by Bhogal *et al.* have been proposed to be the ideal cell-based system [167]. Their use, however, is limited by the availability and variability of human liver tissue samples to start the cultures from. A limited lifespan of primary cells in culture, phenotypic instability over a period of time [167-169], and poor recovery after cryopreservation [170] further restrict their use.

Some of these limitations of primary cultures can be overcome by use of HepG2 cells, a human hepatoma cell line derived from a well-differentiated hepatocellular carcinoma of a 15-year-old Caucasian American male [171]. This adherent cell line has epithelial morphology and appears as flattened, grossly polygonal in shape. The cells are mainly arranged in monolayers [171]. The current study is based on this model in which HepG2 cells are treated with a mixture of FA to induce lipid accumulation with the least possible cytotoxic effects. This model is further described in section 3.1.1.

1.8. Proteomics approach to lipid accumulation

Although NAFLD has dyslipidaemia at its core, the molecular mechanisms that lead to the dysfunction are driven by proteins that bind, transform and carry the various metabolites to and from their compartments. Proteins, and the pathways they are a part of, represent these mechanism and so a comparison of the changes induced in them by lipid accumulation would represent the ways in which the cell copes with the influx of exogenous FA. The goal of this thesis was to describe the proteomic changes that occur in HepG2 following lipid accumulation. The term “proteomics” was coined from the merger of “protein” and “genomics” in the 1990s [172] and has evolved into an expanding field of biology that

includes identification and quantification of global changes in protein abundances and post-translational modifications [173, 174]. Early proteomics work relied heavily on gel-based technologies with post staining or fluorescent pre-labelling but it has since moved toward mass spectrometry (MS) -based analysis [175-177]. The current work has used the 'shotgun' or bottom-up proteomics approach. A typical bottom-up proteomics experiment can be divided into three parts – first, sample preparation in which the protein sample is reduced, alkylated and digested into peptides; second, separation of peptides in the mixture and third, identification of the separated peptides by MS [178]. Sample preparation is a critical step in a proteomics experiment and the strategies used for protein extraction depend largely on the origin of the sample and information required from the experiment; various procedures for sample preparations are described by Canas *et al.* in [179]. The extracted proteins are usually digested by trypsin, a serine protease that cleaves proteins to generate peptides with either arginine or lysine at the C termini. Although proteases such as chymotrypsin, LysC, LysN, AspN, GluC and ArgC have been described and used for digestion of proteins for proteomics analysis, sample preparation using trypsin remains the more popular approach [180]. The peptide mixture is analysed using reverse-phase liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS).

1.8.1. Separation of peptides and mass spectrometry

The peptides in the sample mixture are separated by LC usually performed with a reverse-phase column [181]. Most columns are packed with silica and the stationary phase is usually made of hydrophobic alkyl chains commonly known as C4, C8 or C18 depending on the carbon chain length [182]. The mobile phase formed by the mixing of two reverse phase solvents. By convention, the aqueous solvent and the organic solvent are installed on the HPLC channels A and B respectively and thereby referred to as solvent A and solvent B. A weak organic acid like formic acid or trifluoroacetic acid is included in the mobile phase to improve the peak shape and provide a source of protons [183]. The separated peptides are channelled into the mass spectrometer for gathering the spectral information required for peptide and protein identification.

The mass spectrometer consists of an ion source and optics to generate and focus the ions, a mass analyser like the LTQ™ and/or Orbitrap™ and the data processing electronics that record the information (Figure 7). Ionization of proteins and peptides requires soft ionization techniques that can transfer the polar, non-volatile, and thermally unstable fragments into

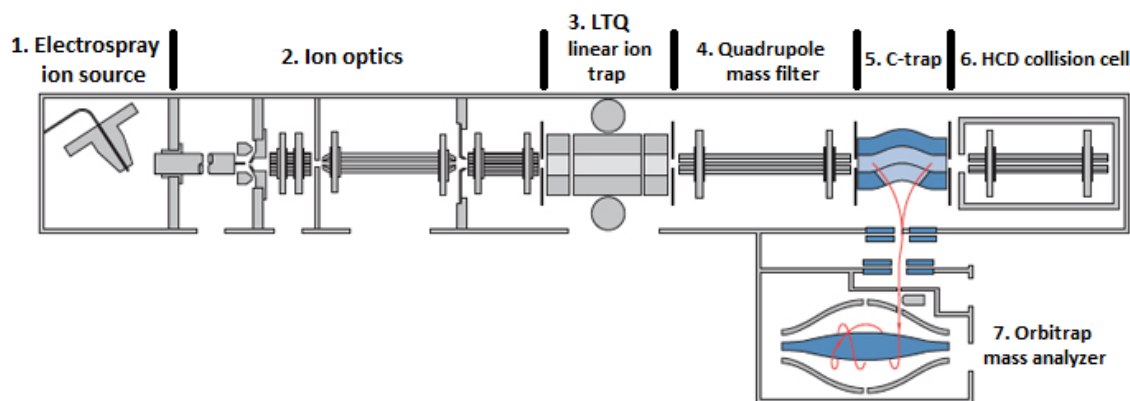


Figure 7: Schematic representation of LTQ-Orbitrap™ XL mass spectrometer. The information available in the user guide made available with the instrument and the schematic on its online version (<http://planetorbitrap.com/ltq-orbitrap-xl#.V49tUdKUdrE>) were used to make this figure.

the gas phase with minimal degradation [184, 185]. The LTQ-Orbitrap™ XL (Thermo Fisher Scientific Inc.) used for this project uses electrospray ionization [186, 187]. Although there are several models for ion formation through electrospray ionization, the general approach includes generating an electrically charged spray of the analyte such that the solvent from the charged droplets evaporates and the size of the droplet continues to reduce till it reaches the Rayleigh limit [188]. At this point, the droplet subdivides since the Coulomb repulsion forces exceed the surface tension [188]. The sub-division of the droplets continues till, in theory, each droplet has an average of one macromolecule whose spectra can now be analysed [188-190]. The LTQ-Orbitrap is a hybrid instrument that utilizes the LTQ (Thermo Scientific version of linear ion trap) for isolation and fragmentation of ions and the Orbitrap mass analyser measures the accurate mass of ions transferred to it [191]. The analytes undergo two fragmentations - the survey scan selects MS1 parent ions that are further fragmented by collision induced dissociation (CID) to form MS2 daughter ions used to obtain the peptide sequence by a database search [192].

As the name suggests, fragmentation based on CID involves the collision of kinetically excited peptide ions with an inert gas like helium or argon. The translational energy imparted to the peptide ion in each collision is converted into vibrational energy that is distributed in picoseconds through all covalent bonds; this surge in internal energy allows bond cleavage and formation of fragment ions [193]. Peptides ionized using electrospray ionization tend to fragment along the protonated amide bonds and form complementary b and y type product ions (Figure 8). The mass differences observed between homologous ions helps identify an amino acid and facilitates peptide sequence analysis [194].

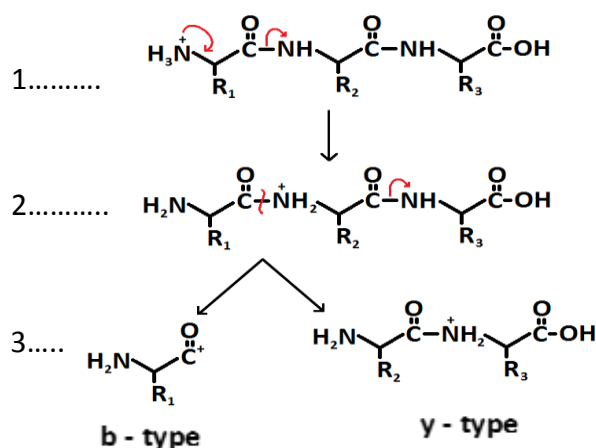


Figure 8: Fragmentation of peptides using CID [194, 195]. The amide nitrogen in the kinetically excited peptides ion is protonated (1) and the collision of these ions with those of an inert gas increases the internal energy of the ion. When the internal energy exceeds the activation barrier required for a particular bond cleavage, the bond breaks (2). The low-energy CID (collision energies less than 100 eV/charge) employed by the LTQ-Orbitrap collision energy used in CID produces the cleavage of peptide bond to form two positively charged ions of b and y type (3).

1.8.2. Peptide sequencing and identification of proteins

Identification of peptides on the basis of MS2 fragments generated by LC-MS/MS is done using search algorithms. The general principles that drive these algorithms include matching of the tandem mass spectra derived from peptide fragmentation with the spectra generated from *in silico* digestion of a protein sequence database [195, 196]. The quality of match is controlled by parameters such as mass tolerance, numbers of missed cleavages accepted and amino acid modifications. The peptide sequences are used for making inferences about the protein that would contain them [196]. SEQUEST, first described by Eng *et al.* in 1994 [197] has been developed over the last 20 years to add new capabilities to the original algorithm that was capable of producing automated, high-throughput matches for the tandem mass spectra against a protein sequence library [198]. Successor algorithms include SEQUEST-HT [198]. Cottrell *et al.* provided a probability-based protein identification through the use of the MASCOT search engine [199]. Since then, more search algorithms like the Open Mass Spectrometry Search Algorithm [200], Andromeda [201] and X! Tandem [202] have become available for use.

Although various algorithms work by the same general principles, they tend to have their own metrics to evaluate the quality of the match and the likelihood of false identifications. As a result, different algorithms may report very different peptide matches for the same set

of spectra [203]. Given the uncertainty in the arbitrary scores assigned by each algorithm and the inability to perform false discovery rate (FDR) calculations, the peptide matches generated by the search algorithms are evaluated by stand-alone algorithms like PeptideProphet [204, 205] and Percolator [206, 207] that estimate FDR using the target-decoy approach. This approach includes the search of a database of random or reversed sequences and of comparable size to the target database with the aim of identifying incorrect identification. These databases are known as decoy databases and the sequences within them are the decoys. Since the decoy database does not have valid protein sequences, presence of a decoy in the data helps estimate the possible incorrect identifications in the target database [208, 209]. Since a combination of bioinformatics tools is required to perform all analytical steps necessary to obtain a result, software suites and pipelines like MaxQuant, Proteome Discoverer and Scaffold have been developed to perform the required steps for streamlined data analysis [210-212]. This project has used the Proteome Discoverer platform for protein identification and Scaffold for quantification of the proteomics data.

1.8.3. Quantitative proteomics

Having identified the proteins, an absolute or relative quantification of proteins provides a comparison between two or more samples analysed. These samples could be of different origins or could represent different physiological states of the same biological system. MS-based quantification methods are broadly classified into labelled or label-free methods [213]. Labelled methods include the use of stable isotope or chemical mass tags for quantification. Metabolic labelling of proteins is a term used to describe labelling techniques that are based on the incorporation of a labelled metabolite into newly synthesized proteins [214]. Metabolic labelling using stable-isotope labelling by amino acids in cell culture (SILAC) as the name suggests, relies on the incorporation of amino acid labelled with heavy isotopes of hydrogen (^2H), carbon (^{13}C) or nitrogen (^{15}N) into newly synthesized proteins in a cell culture system. The experimental design includes growing cells in a modified medium that includes the labelled amino acids. At the end of the treatment, the labelled cells are mixed with a population of cells grown in an identical medium but with unlabelled amino acids. The proteomes of the two populations are analysed and differentiated on the basis of the predictable shift in mass of the labelled peptides of a protein in comparison to the peptides generated from a protein synthesized using 'light', unlabelled amino acids [215]. In the mass

spectra, the peptides appear as a pair separated by the known mass difference on account of the incorporation. Ratio of peak intensities of the pair are then compared and used as a quantitative parameter [133]. Krijgsveld *et al.* have shown that metabolic labelling of the multicellular organisms like *C. elegans* and *D. melanogaster* can be achieved by feeding them on ^{15}N -labeled *E. coli* and yeast, respectively [216].

Like the metabolic labelling methods, quantification using isotope-coded affinity tags (ICAT) also includes comparison of two proteomes based on the peak intensities for heavy and light peptides but the point of difference is that ICAT relies on the chemical modification of cysteine residues of a peptides by the ICAT reagent during the sample preparation stage. The ICAT reagent has three components – a biotin tag that allows enrichment of labelled peptides using avidin affinity chromatography; a linker that contains the stable isotope signatures that allow light and heavy peptides to be differentiated by MS and a thiol-specific reactive group to tag the cysteines [217]. The linker of the ‘heavy’ ICAT reagent includes nine ^{13}C to generate a mass difference of 9 Da between the light and heavy peptides [129]. The obvious limitation of ICAT labelling is the cysteine-specific tagging that can severely restrict the number of proteins identified [129]. Quantification using isobaric tag for relative and absolute quantitation (iTRAQ) could be used to overcome this limitation. It uses N-hydroxysuccinimide ester to tag primary amino groups in lysine residues which are generally present in high abundance [218]. The iTRAQ reagent is designed to include a ‘mass balance group’ that links the N-hydroxysuccinimide to a reporter group – this mass balance group allows for multiple sample comparisons in contrast to the binary comparisons offered by isotope labelling methods. The commonly used ‘4-plex’ reagent generates reporter ions with m/z values of 114.111, 115.108, 116.112, and 117.115 after CID fragmentation [219]. The advantages and disadvantages of these methods are discussed further in Section 6.5. The quantitative proteomics experiments described in this thesis are based on a label-free technique.

Label-free methods measure either the peptide ion intensity or count of the number of fragment-ion spectra acquired for peptides of a given protein [220, 221]. The first approach measures chromatographic peak areas of peptide precursor ions - the intensity of each peak is visualized in an extracted ion chromatogram which allows the determination of the area under the curve used as the measure of protein expression [222]. The second approach is called spectral counting and it provides a relative quantification of proteins based on the

empirical observation that the tandem mass spectra of a particular peptide increase with an increasing amount of the corresponding protein [221]. The advantages and disadvantages of each method have been reviewed in [223, 224]. For the current work, spectral counting was used to quantify and compare the proteomes of HepG2 cells after treatment.

1.9. Aims of the thesis

This thesis is based on the hypothesis that an influx of exogenous FA induces accumulation of TAG in HepG2 cells and the proteomic changes that accompany this lipid accumulation would represent mechanisms that allow the cell to cope with the cellular stress induced by the TAG. Inhibition of DGAT enzymes could then be a feasible approach to reduce the extent of TAG accumulation. Based on this the aims of the current work were:

- Characterize a treatment regime that would induce lipid accumulation in HepG2 cells with minimal lipotoxic effects.
- Determine the proteomic changes induced on account of lipid accumulation in HepG2 cells.
- Describe the effect of DGAT inhibitors on lipid accumulation in HepG2 cells and determine the proteomic changes effected by the DGAT inhibitors

CHAPTER 2: MATERIALS & METHODS

2.1 Materials

All chemicals, including bovine serum albumin (BSA), dimethyl sulfoxide (DMSO), sodium dodecyl sulphate (SDS), 1, 4-dithiothreitol (DTT), 3-[(3-cholamidopropyl) dimethylammonio]-1-propanesulfonate (CHAPS), glacial acetic acid and formic acid, were purchased from Sigma-Aldrich (New Zealand) and were of analytical grade unless specified otherwise. DGAT2 inhibitor PF06424439 [(3R)-1-[2-[1-(4-chloro-1H-pyrazol-1-yl)cyclopropyl]-3H-imidazo[4,5-b]pyridin-5-yl]-3-piperidinyl]-1-pyrrolidinyl-methanone methanesulfonate] and DGAT1 inhibitors A922500 [(1R, 2R)-2-[[4'-[[phenylamino) carbonyl] amino] [1,1'-biphenyl]-4-yl]carbonyl] cyclopentane carboxylic acid] and PF04620110 [(2-((1r,4r)-4-[4-(4-amino-5-oxo-7,8-dihydropyrimido [5,4-f][1,4] oxazepin-6(5H)-yl) phenyl] cyclohexyl) acetic acid)] were also purchased from Sigma-Aldrich (New Zealand).

The HepG2 cell line (ATCC® HB-8065™), a human hepatoma cell line, was obtained from the American Type Culture Collection (ATCC; Manassas, VA, USA). Foetal bovine serum (FBS), Dulbecco's modified Eagle's medium (DMEM) with high glucose (4.5 g/L), 4 mM glutamine and phenol red, Gibco® 2.5 % trypsin and Hank's balanced salt solution (HBSS) were purchased from Life Technologies™ (New Zealand). Trypsin was diluted to 0.25 % in HBSS before use.

Water and solvents used for chromatography including hexane, methanol and acetonitrile were from LiChrosolv®, Merck Millipore (Billerica, MA, USA). Diethyl ether was obtained from Ajax Finechem (now integrated into Thermo Fisher Scientific NZ Ltd). Thin layer chromatography (TLC) plates purchased from Merck Ltd (New Zealand) were 20 x 20 cm with a 200 µm thickness of Silica gel 60 on aluminium backing and a concentrating zone.

NBD-palmitoyl coenzyme A [(N-((7-nitro-2-1,3-benzoxadiazol-4-yl)-methyl) amino) palmitoyl coenzyme A (ammonium salt)] and 1,2-dioleoyl-*sn*-glycerol used for a fluorescence-based DGAT activity assay were obtained from Avanti Polar Lipids Inc. (Alabaster, AL, USA).

The secondary antibodies and most primary antibodies were from Abcam plc (Cambridge, UK); primary antibodies against actin, DGAT1 and DGAT2 were from EMB Millipore (Billerica, MA, USA), Santa Cruz Biotechnology, Inc (Dallas, TX, USA) and Atlas Antibodies (Stockholm, Sweden) respectively (Table 1). Goat anti-rabbit IgG Fc tagged with Alexa Fluor® 647 and goat anti-mouse IgG H&L tagged with Alexa Fluor® 555 were as secondary antibodies.

Table 1: Description of primary antibodies used for Western blots

| Target | Antibody | Host | Type | Description of antigen |
|--|-----------|--------|------------|--|
| Perilipin-2 | ab78920 | Rabbit | Polyclonal | Synthetic peptide (ab 90212) conjugated to Keyhole Limpet Haemocyanin (KLH) derived from within residues 1 - 100 of perilipin 2. |
| 75 kDa Heat shock protein, mitochondrial | ab151239 | | Polyclonal | Recombinant fragment, corresponding to a region within aa 392-704 of human TRAP1. |
| Glutathione reductase, mitochondrial | ab137513 | | Polyclonal | Recombinant fragment, corresponding to a region within aa 164-403 of human glutathione reductase. |
| Glutaredoxin-3 | ab170794 | | Polyclonal | Synthetic peptide within human TXNL2 conjugated to KLH*. |
| DGAT1 | sc-32861 | | Polyclonal | Antibody raised against aa 1-100 mapping at the N-terminus of human DGAT1. |
| DGAT2 | HPA013351 | | Polyclonal | Recombinant Protein Epitope Signature Tag (PrEST) antigen sequence. |
| Fatty acid synthase | ab22759 | | Polyclonal | Synthetic peptide ab25719 conjugated to KLH derived from within residues 2450 to the C-terminus of mouse FAS. |
| ERLIN1 | ab171372 | | Polyclonal | Synthetic peptide within human ERLIN1 aa 245-275 (internal sequence) conjugated to KLH*. |
| BAG family molecular chaperone regulator 2 | ab79406 | | Monoclonal | Synthetic peptide corresponding to human BAG2 aa 1-100 (N terminal)*. |
| Actin | MAB1501 | Mouse | Monoclonal | Purified chicken gizzard actin epitope corresponding to aa 50-70 [225]. |
| 60 kDa heat shock protein, mitochondrial | ab110312 | | Monoclonal | Hela cell lysate |
| Catalase | ab125688 | | Monoclonal | Recombinant full length human catalase produced in HEK293T cells (NP_111743). |
| *The exact sequence is proprietary. | | | | |

2.2 Cell culture

HepG2 cells placed in liquid nitrogen for long-term preservation were thawed and grown in DMEM with 10 % FBS at 37 °C in an incubator with a humidified atmosphere of 5 % CO₂ and 95 % air until the cells were confluent [226]. The medium was changed every three days or more frequently depending on the confluency of the cells and the number of cells required for subsequent work. Excess cells were returned to long-term storage as cell suspensions of 10⁶ cells in DMEM with 10 % FBS and 10 % DMSO. The cell suspensions were frozen overnight in CoolCell® (BioCision, San Rafael, CA, USA) freezing containers at -80 °C and then placed in liquid nitrogen cryovials [227]. CoolCell® is a styrofoam container that facilitates a slow and controlled drop of temperature of about -1 °C per min [228]. Cells were passaged once they reached confluency. Confluent cells were washed with warm HBSS and incubated with 0.25 % trypsin to detach the cells from the flask surface.

The detached cells were suspended in DMEM with 10 % FBS and pelleted by centrifugation at 180 g for 3 min. The cell pellet was re-suspended in the medium and transferred to flasks or plates. Tissue culture flasks (75 cm²; Corning Falcon®, Corning, NY, USA) were inoculated with 2 x 10⁶ cells for passage of the cell culture, and multi-well plates (Corning Falcon®, Corning, NY, USA) used for experiments were inoculated as required for the specific assays. Cells were allowed to attach and grow for at least 18 h before treatment.

2.3 Lipid accumulation in HepG2 cells

2.3.1. Preparation of fatty-acid stock solutions

BSA-bound FA stock solutions were prepared by dissolving a measured amount of FA in a aqueous solution and adding this to the BSA solution as described by Spector[229]. To prepare 30 mM stock solutions of sodium oleate, 91.32 mg of the sodium salt was solubilized in 5 mL of 150 mM sodium chloride at 70 °C in a water bath. This was added in aliquots to a warm solution of 5 mM BSA in 150 mM sodium chloride at 37 °C with occasional vortexing until the solution was clear. Once the solution was clear, the volume was made up to 10 mL. The BSA-bound oleate was passed through 0.2 µm Minisart® syringe filters (Sartorius, Göttingen, Germany) before being stored at –20 °C. A 30 mM stock solution of palmitate was prepared in the same way by solubilizing 83.52 mg of sodium palmitate in 10 mL of 5 mM BSA in 150 mM sodium chloride. The stock solution of FA mixture was prepared by mixing 2 volumes of the BSA-oleate stock with 1 volume of the BSA-palmitate stock. This 30 mM stock solution of FA mixture was diluted to give the FA concentrations used for the experiments [230].

2.3.2. Optimization of treatment conditions to induce lipid accumulation

HepG2 cells were inoculated in a 24-well plate (10⁵ cells per well) and grown overnight in DMEM with 10 % FBS. The attached cells were washed with warm HBSS and incubated with serum-free DMEM containing 1 % BSA for 1 h at 37 °C [231]. At the end of the incubation, the media was replaced with FA-containing media which was prepared by diluting the stock solution of FA mixture serum-free DMEM containing 1 % BSA to give final concentrations ranging from 0.25 to 3 mM. Using the Oil red O (ORO) assay, lipid accumulation in HepG2 cells was measured at 1, 3, 6, 12, 18 and 24 h exposure to FA mixture [230]. The metabolic activity of the cells was also measured at these time points (see Section 2.3.4).

2.3.3. Quantification of lipid accumulation using Oil Red O assay

2.3.3.1. Preparation of Oil Red O staining solution

ORO was dissolved in isopropanol at a concentration of 3.5 mg/mL. This stock solution was allowed to rest overnight and undissolved dye was removed by passing through a 0.2 µm syringe filter. The ORO staining solution was made by mixing 6 parts of 3.5 mg/mL ORO stock solution with 4 parts of distilled water. The working ORO solution was allowed to stand overnight and filtered before use to remove any precipitates [232].

2.3.3.2. Staining of lipid droplets

Treated HepG2 cells were washed with HBSS and fixed in 4 % paraformaldehyde at room temperature (RT) for 30 min. The fixed cells were washed with filtered phosphate-buffered saline (PBS) and, after a wash with 10 % isopropanol, cells were covered with the ORO staining solution for 15 min at RT. The excess dye was removed, and the cells were washed multiple times with distilled water till the cytosol appeared unstained under the microscope. The cells with red stained lipid droplets were examined under a light microscope [233]. Preparations intended for cell staining only were counter-stained with haematoxylin and eosin to visualize the nucleus and the cell body.

2.3.3.3. Quantification of extracted dye

Equal volumes of isopropanol were added to each well to extract the dye bound to the accumulated lipids. The ratio of absorbance of the dye extract from the treated cell to that of the extract from the untreated, control cell was expressed as 'fold of control'. This was used as a relative measure of lipid accumulation in the cells on account of the treatment [230]. The wavelength at which this absorbance would be measured was determined using a spectrum scan and preliminary experiments described in section 3.2.1.1. The absorbance and the spectrum scan were performed on the plate reader (EnSpire™ 2300 Multilabel Reader, Perkin Elmer, Waltham, MA, USA). Two-way analysis of variance (ANOVA) with Bonferroni's multiple testing correction was used to determine statistical significance of the extent of lipid accumulation seen at different combinations of treatment FA concentrations and time. Significance was accepted at $p < 0.01$ and the comparisons were performed using GraphPad Prism® (v.5) [GraphPad Software Inc., La Jolla, CA, USA].

2.3.4. MTT assay

The metabolic activity and the proliferation capacity of the cells was measured using

the 3-(4, 5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) assay. The assay is based on the protocol described first by Mosmann [234]. Briefly, at the end of the treatment time, the cells were incubated for 2 h with 0.8 mg/mL MTT in serum-free medium. After 2 h, the medium was discarded, cells were washed with 1 mL HBSS and 300 μ L of DMSO was added to each well. The dissolution of the formazan crystals in DMSO was facilitated by repeated pipetting. The absorbance of the resulting solutions was recorded at 560 nm using a plate reader (EnSpire™ 2300 Multilabel Reader, Perkin Elmer, Waltham, MA, USA). Results were presented as a percentage of the control values [235]. Two-way ANOVA with post-hoc Bonferroni's multiple comparison test was performed to compare the means between different treatments.

2.4. Effect of DGAT inhibitors on lipid accumulation

DGAT2 inhibitor PF06424439 (PF06) and DGAT1 inhibitors A922500 (A92) and PF04620110 (PF04) were added to the culture medium, and their effect on the lipid accumulation in the cell was quantified using the ORO assay described above. Stock solutions of 10 mM PF06 were prepared in sterile distilled water, and 50 mM A92 and 10 mM PF04 were dissolved in DMSO. The stock solutions were stored at -20 °C.

For the assay, the inhibitors were diluted to various concentrations in DMEM with 1 % BSA. HepG2 cells grown overnight in a 24-well plate were washed with warm HBSS and incubated for 1 h at 37 °C with medium containing the inhibitors. At the end of the incubation, an aliquot of the FA mixture stock was added to the wells to give a final concentration of 1 mM and the cells were further incubated in the medium with the FA mixture for 6 h at 37 °C [124]. Two controls were included on each plate- a positive control of cells treated with 1 mM FA mix in DMEM with 1 % BSA without the inhibitors and a negative control of untreated cells in DMEM with 1 % BSA. Since A92 and PF04 were dissolved in DMSO, an aliquot of DMSO equivalent to the volume of the highest concentration in the inhibitor-treated samples (100 μ M) was added to the control wells.

A combination of PF04 and PF06 inhibitors was also performed. The treated cells were then stained with ORO as described above, and the ORO-based quantification was used to determine the extent of inhibition of TAG synthesis by these inhibitors. A potential rescue of cells from lipotoxicity by the inhibitor was assessed using the MTT assay at the end of the treatment.

2.5. Fluorescent DGAT assay

2.5.1. Microsome preparation

Pellets of untreated HepG2 cells stored at -80 °C were thawed and suspended in sucrose buffer [0.25 mM sucrose in 50 mM Tris-HCl (pH 7.6)]. The suspended cells were lysed by trituration through a 27-gauge needle multiple times. The lysate was centrifuged at 10,000 g for 30 min at 4 °C in an Optima™ L-100 XP ultracentrifuge (Beckman Coulter Inc., Fullerton, CA, USA). The pellet was discarded and the supernatant further centrifuged at 100,000 g for 1 h at 4 °C to give the microsomal pellet. This pellet was suspended and dissolved in about 200 µL of sucrose buffer. The protein content of this suspension was quantified using a Bio-Rad Protein Assay reagent (Bio-Rad Laboratories Pty Ltd., NZ) based on Bradford's method of protein quantification which is a Coomassie dye-binding assay that uses the differential change in dye colour as a measure of protein concentration. Aliquot equivalent to 30 µg protein sample was diluted with the sucrose buffer to give a final volume of 50 µL. This was used as the enzyme source for the assay.

2.5.2. Reaction mixture for DGAT assay

A reaction mixture containing 50 µL of enzyme preparation along with 20 µL of 1 M Tris-HCl (pH 7.6), 4 µL of 1 M MgCl₂, 10 µL of 4 mM dioleoyl glycerol and 96 µL distilled water was warmed in a water bath at 37 °C for 5 min. Stock solutions of the inhibitor were diluted in the reaction mixture to give the desired final concentration in a total reaction volume of 200 µL. The reaction (Figure 9) was started by adding 10 µL of 500 µM NBD-palmitoyl CoA to the pre-warmed reaction mixture and was allowed to proceed for 30 min at 37 °C with occasional shaking [236]. Each experiment included a set of blanks that would serve as the negative controls for the reaction. The substrate blank did not have 1,2-dioleoylglycerol but included the fluorescent substrate NBD-palmitoyl CoA in the reaction mixture. The enzyme

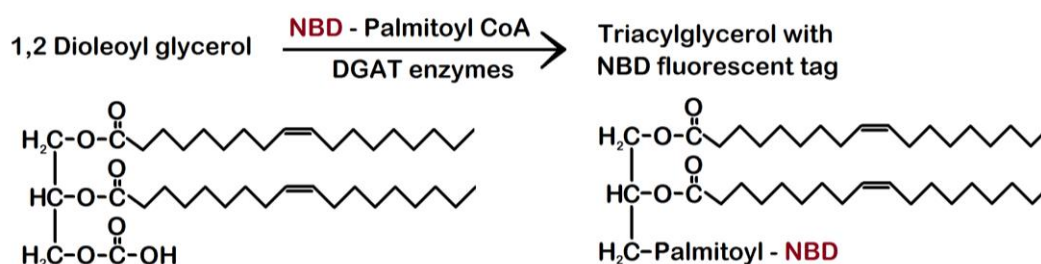


Figure 9: Schematic representation of enzyme reaction catalysed by DGAT enzymes. The reactants in the fluorescent assay include dioleoyl glycerol and a palmitoyl CoA tagged with fluorescent NBD. Esterification of the tagged acyl to the dioleoylglycerol by the DGAT enzymes produces a fluorescent NBD-tagged triacylglycerol which is detected after separation of the reaction mixture by TLC.

blank included all substrates and heat-inactivated aliquot of the microsomal preparation. In addition to these, a solvent blank containing equivalent volume of DMSO was also prepared for the inhibitors A92 and PF04.

2.5.3. Separation and quantification of products

The reaction was terminated by adding 750 μ L chloroform/methanol (2 volumes: 1 volume) and vortexing thoroughly. Water was added and, after another thorough vortexing, the phases were allowed to separate at RT for 1 h. The reaction solution was centrifuged for 5 min at 16,000 g to improve phase separation [236]. A 400 μ L aliquot of the lower organic phase was collected and dried in a CentriVap® Vacuum Concentration System (Labanco, Kansas City, MO, USA). The lipids were re-suspended in 20 μ L chloroform/methanol (2:1, v/v), and 10 μ L was spotted in the concentration zone of a TLC plate. The plate was placed in a solvent chamber saturated with the solvent vapour, and the mobile phase was allowed to run till the solvent front was about 3 cm from the top edge of the plate.

Preliminary experiments were performed using a mobile phase that consisted of *n*-hexane: diethyl ether: acetic acid (80: 20: 1; v/v). Based on the results of these experiments, an improvement of band separation was sought and so, a solvent mixture of diethyl ether: *n*-hexane: methanol: acetic acid (55: 45: 5: 1, v/v/v/v) was tested and used as the mobile phase to separate the lipids [119]. Once the lipids were separated on the plate, the plate was removed from the solvent chamber, and the solvents were allowed to evaporate in a fume hood. The newly synthesized labelled TAG was detected on a Fujifilm FLA-5100 (Fujifilm Life Science, Stamford, CT, USA) molecular imaging system.

2.6. Proteomic analysis by LC-MS/MS

2.6.1. Type of replicates

The experimental design for this study included comparison of three treatment groups – untreated HepG2 cells; cells treated with FA mixture and cells treated with FA mixture and DGAT inhibitor. Each treatment group was represented by four biological replicates. Cell samples collected from different flasks on different days were considered biological replicates. In one scheme of sample preparation, the four biological replicates were pooled into one sample. To prepare the pooled sample, aliquots of the cell lysate equivalent to 7.5 μ g protein from each biological replicate were mixed just after lysis to give 30 μ g total protein of the pooled sample. For the unpooled samples, 30 μ g protein aliquot was prepared for each biological replicate. One peptide digest was prepared from each sample as

described in section 2.6.2, pooled or otherwise. Multiple LC-MS/MS runs were performed for peptide digest. Each LC-MS/MS run was considered a technical replicate. All statistical analysis were performed biological replicates unless stated otherwise.

The experimental design had two variabilities in the data: the biological variability of the cell population and the technical variability between the LC-MS/MS runs. Different experimental designs were tested using individual biological replicates and pooled samples. They are described in detail in chapters 4 and 5.

2.6.2. Sample preparation

HepG2 cells representing the three groups were washed with cold HBSS and scraped off the flask surface. Protease inhibitor cocktail (Catalogue Number P8340, Sigma-Aldrich, St. Louis, MO, USA) was included in the final wash. The cells were pelleted by centrifugation at 180 g for 3 min and stored at -80 °C after discarding the supernatant.

Once all the replicates were gathered, the cell pellet of approximate 4×10^6 cells was re-suspended in 300 μ L of lysis buffer and vortexed for 30 min at 4 °C. Given that lysis buffer can affect the quality and quantity of proteins extracted, preliminary samples of untreated HepG2 cells were prepared using four different lysis buffers, including glycerol-Triton-X 100 buffer [237], RIPA buffer (Sigma-Aldrich) [237], SDS-DTT buffer [238] and urea- CHAPS buffer [239] (Table 2). Protease inhibitor P8340 was included in all the lysis buffers.

Table 2 : Composition of lysis buffers used for sample preparation.

| Glycerol- Triton-X 100 | RIPA | SDS-DTT | Urea- CHAPS |
|-------------------------|----------------------------|-----------------------------|----------------------------|
| 50 mM HEPES (pH 7.4) | 25 mM Tris-HCl (pH 7.5) | 100 mM Tris-HCl (pH 7.5) | 30 mM Tris-HCl (pH 8.8) |
| 150 mM KCl | 150 mM NaCl | 4 % SDS | 7 M Urea |
| 10 % Glycerol | 1 % SDS | 0.1 M DTT | 2 M Thiourea |
| 0.5 % Triton-X 100 | 1 % Triton X-100 | --- | 4 % CHAPS |

Cell debris was pelleted by centrifugation at 10,000 g for 30 min at 4 °C. The proteins from the supernatant were precipitated using a ProteoExtract® protein precipitation kit (Merck Ltd., NZ) according to the manufacturer's protocol and dissolved in digestion buffer (8 M urea in 100 mM Tris-HCl, pH 8). The protein concentration of the sample was determined using 2-D Quant Kit (GE Healthcare, NJ, USA) following the manufacturer's protocol. An aliquot equivalent of 30 μ g protein in 50 μ L of the digestion buffer was reduced by treatment with 5 mM DTT at 56 °C in a water bath for 35 min and alkylated with 10 mM iodoacetic acid for 40 min at RT in the dark [240]. The reduced and alkylated sample was

digested with modified sequencing grade trypsin (Roche Custom Biotech, Indianapolis, IN, USA) in the presence of 1 mM CaCl₂ after 3-fold dilution with 100 mM Tris-HCl, pH 8, to bring the total urea concentration to about 2 M. The digestion was stopped by the addition of 90 % formic acid, and the resulting digest was de-salted and cleaned using OMIX C18 tips (Varian Inc., Palo Alto, CA, USA).

The peptides bound to the OMIX C18 matrix were eluted in 20 µL of 0.1 % formic acid in 50 % acetonitrile followed by elution with 20 µL of 0.1 % formic acid in 70 % acetonitrile. The two eluents were combined and concentrated to about 10 µL in a CentriVap® Vacuum Concentration System. The final volume of the peptide preparation was adjusted to 150 µL with 0.1 % formic acid. This solution was used for analysis by LC-MS/MS.

2.6.3. LC-MS/MS analysis of protein digests

2.6.3.1. Instrumentation

LC-MS/MS was carried out using a Dionex UltiMate™ 3000 RSLCnano liquid chromatography system (Thermo Scientific, USA) coupled to a Thermo Scientific™ LTQ Orbitrap XL™ Hybrid mass spectrometer via a nanospray ion source. The LTQ Orbitrap is a Fourier Transform Mass Spectrometer based on Thermo Scientific™ LTQ XL™ linear ion trap and Orbitrap mass spectrometry technologies. Thermo Scientific™ Xcalibur™ software (v 2.1.0) was used for data acquisition, instrument control and method setup.

2.6.3.2. Separation of peptides using reverse-phase liquid chromatography

Samples of tryptic peptides were separated on an Acclaim® PepMap100 C18 analytical column [particle size: 3 µm, pore size: 100 Å, dimension: 75 µm inner diameter × 15 cm length] at a constant flow rate of 0.3 µL/min. Samples were loaded using an auto-sampler maintained at 10 °C. The peptides were separated on the principle of reverse-phase LC – the most hydrophilic and the least hydrophobic eluates are the first to exit the column, and transit time is also affected by peptide mass. The mobile phase consisted of a solvent mixture set up by mixing 0.1 % formic acid (solvent A) and 0.1 % formic acid in 80 % acetonitrile (solvent B) in a multi-step gradient. Initial experiments were performed over a 300-min retention period. This gradient was extended to 352 min and modified further based on the analysis of LC-MS/MS runs using Gradient Optimization and Analysis Tool (GOAT®, The University of Texas Southwestern Medical Center, Dallas, TX, USA)[241]. The three gradients are compared with respect to their duration, composition and performance in chapter 4.

2.6.3.3. Mass spectrometry analysis

The separated peptides were introduced into the tandem MS system by electrospray ionization using a PicoTip™ emitter (360 µm x 20 µm; New Objective, USA) at a voltage set to 1.8 kV through a transfer tubing of 25 µm inner diameter. The acquisition method used was an Nth order double play set to acquire data over two scan events. The N most intense ions ('Top N' ions) detected in scan event 1 were selected for data-dependent MS/MS scans performed in scan event 2. Scan event 2 was performed starting with the most intense ions of scan event 1 and proceeded in order of decreasing intensity of the top N ions selected. The method was tested in 3 settings : 'Top 6', 'Top 8' and 'Top 10' ions – these would be the 6th, 8th and 10th most intense ions detected in their respective scan event 1.

Scan event 1 was a full scan for all ions eluted at a particular retention time with m/z range of 200 – 1800 at a resolution of 30,000 at m/z 400 at a scan rate of 0.6 to 0.8 Hz. The scan was performed in the positive ion mode with the heated capillary temperature set at 200 °C. Scan event 2 was a data-dependent scan that used a preview scan and dynamic exclusion settings to pick the top N precursor ions for CID fragmentation and detection by the ion trap analyzer. CID was operated at an isolation width of 1.0 m/z , 35 % normalized collision energy, activation Q of 0.25 and activation time of 30 ms. Maximum injection time for the MS scan was 500 ms and 150 ms for the MS/MS scans. The dynamic exclusion settings included repetition duration of 30 s, exclusion list size of 500 and exclusion duration of 90 s.

2.6.4. Protein identification

Tandem mass spectra were searched against the human protein database obtained from UniProtKB using Thermo Scientific™ Proteome Discoverer™ (PD) software (version 1.4.1.14). UniProtKB database can be broadly divided into two categories – SwissProt entries that have been manually reviewed and annotated and TrEMBL entries that have been annotated from literature but await manual review [115]. Four subsets of the UniProt database were generated for evaluation in this project (Table 3). These include proteins identified in humans were retrieved from the UniProtKB using the keywords 'organism: "Homo sapiens (Human) [9606]"' and this list was further stream-lined using the keywords 'Protein existence: evidence at protein level' that selected entries supported by experimental evidence of existence. The proteins retrieved using keywords 'organism: "Homo sapiens (Human) [9606]"' included all human proteins in the database. Filtering this list further using the keywords 'Protein existence: evidence at protein level' formed a database of human

proteins the existence of has been experimentally verified. The databases were curated into a .fasta format and parsed by the search algorithms to identify the proteins from the LC-MS/MS spectra.

Different workflows as shown in Figure 10 were tested to identify the workflow that identified the maximum number of proteins. All workflows require 'Spectrum selector' node for input of MS/MS data in the form of a .raw file. All MS/MS data were analysed using Mascot (Matrix Science, UK; version 1.4.1.14) and/or SEQUEST HT (Thermo Fisher Scientific, USA; version 1.4.1.14). Both Mascot and SEQUEST HT were set up to search the databases described in Table 3, assuming the digestion enzyme used was trypsin with allowance for a maximum of 2 missed cleavage sites. Mascot and SEQUEST HT were searched with a fragment ion mass tolerance of 0.60 Da and a parent ion tolerance of 10.0 ppm.

Table 3: Description of databases searched for protein identification.

| Database | Keyword | Number of sequences | | |
|-----------|---|---------------------|----------|------------|
| | | Total | Reviewed | Unreviewed |
| UniProt | organism: "Homo sapiens (Human) [9606]" | 148,986 | 20,196 | 128,790 |
| | organism: "Homo sapiens (Human) [9606]" AND existence: "evidence at protein level" | 51,543 | 14,685 | 36,858 |
| SwissProt | organism: "Homo sapiens (Human) [9606]" | 20,196 | 20,196 | 0 |
| | organism: "Homo sapiens (Human) [9606]" AND existence: "evidence at protein level" | 14,685 | 14,685 | 0 |

Carbamidomethylation of cysteine was specified as a fixed modification, and oxidation of methionine was specified as a variable modification for both search algorithms. All workflows included the 'Percolator' node. Percolator is an algorithm that uses semi-supervised machine learning to improve the discrimination between correct and incorrect spectrum identifications [206]. The algorithm searches a decoy database with negative examples for the classifier and positive examples constituted by a subset of the high-scoring matches from the target database to generate q values and posterior error probabilities for each match [242]. The q values determine the FDR, and filters were set to accept identifications with $q < 0.01$. This equates to a FDR of 1 % or less. The maximum ΔC_n in the scoring options of the 'SEQUEST HT' node was set to 1 to allow the maximum amount of data to be assessed by the 'Percolator' node [205]. In addition to the above, workflow 'Combined 2' includes the 'MS2 – Spectrum Processor' node that performs charge state deconvolution and deisotoping of the data.

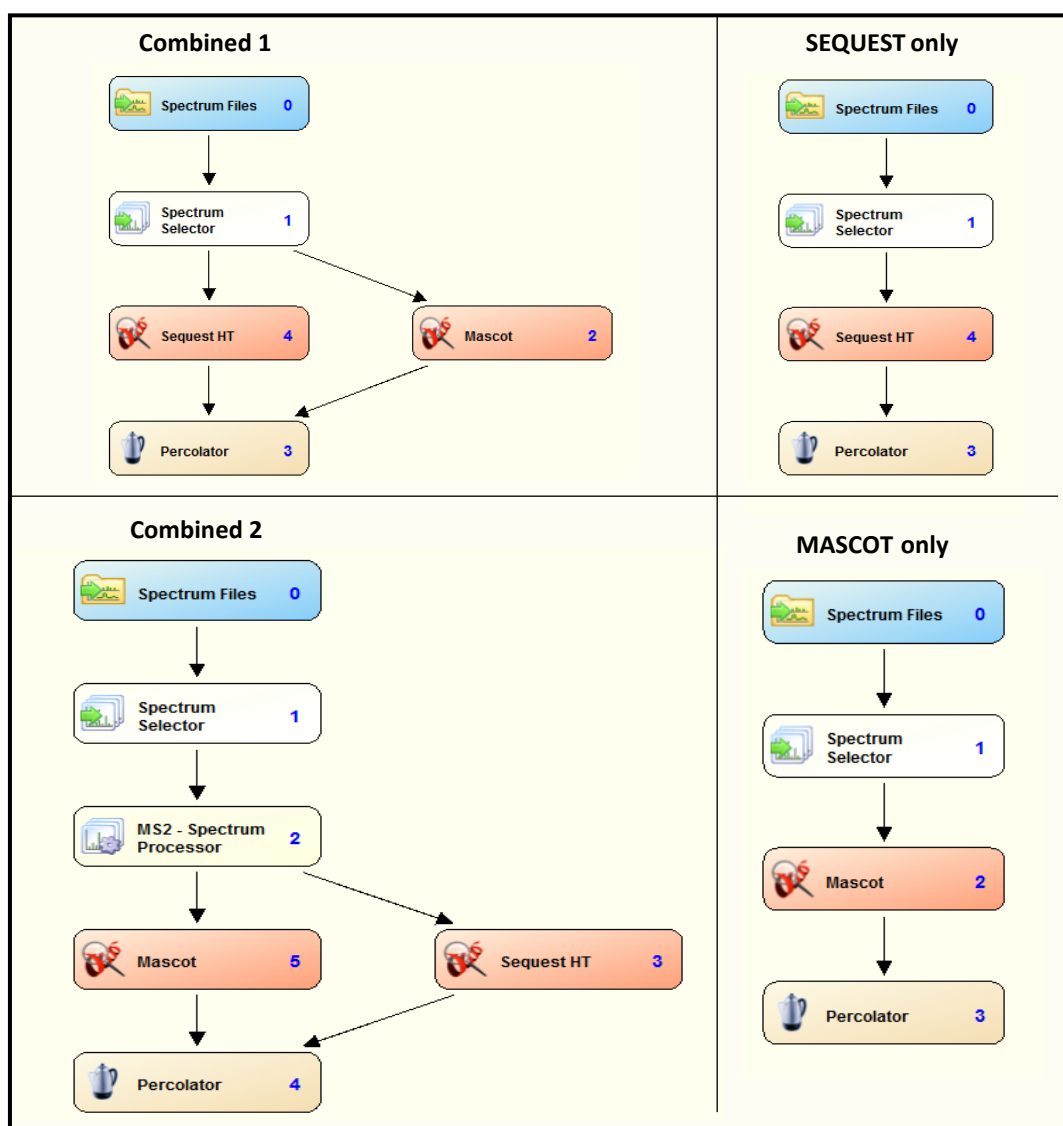


Figure 10: Protein identification workflows setup in Proteome Discoverer (v. 1.4.1.14).

2.6.5. Label-free quantification

Label-free quantification based on spectral counts was performed using Scaffold (version 4.4.6, Proteome Software Inc., USA). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. The samples were organized into three categories – ‘Control’, ‘FA-treated’ and ‘Inhibitor-treated’. Each category had four ‘biosamples’ which represented biological replicates and each biological replicate was represented by four LC-MS/MS runs. The LC-MS/MS runs representing the pooled samples were generated using one digest. For experiments with pooled sample, each category has one biosample and each biosample was represented by five LC-MS/MS runs. The information from each technical replicate was condensed to give cumulative spectral data representative of the sample.

This cumulative spectral data was referred to as the 'total spectral count'. It is the sum of all the spectra associated with a specific protein within a sample. For proteins that showed a significant change in expression, manual verification was carried out to exclude peptides that were shared with other proteins. The total spectral count was normalized to allow comparisons between samples. The normalisation scheme in Scaffold adjusts the sum of the selected quantitative values for all proteins in the list within each MS sample to an average of the sums of all MS samples present in the experiment. This was achieved by applying a scaling factor for each sample to each protein or protein group. This normalized total spectral count was referred to as the 'quantitative value (QV)'. The ratio of the quantitative value of two groups was expressed as the 'fold change' and represented the differential expression of the protein for that sample.

2.6.6. Statistical analysis

Depending on the experimental design, one of the three tests was used to determine the statistical significance of the difference in the quantitative value reported for a protein in the three treatment groups tested. These tests were Fisher's exact test, *t*-test and simple one-way ANOVA. Fisher's exact test and *t*-test were used to compare the relative abundance between two groups at a time while ANOVA was used to compare all three groups. The results of Fisher's exact test and *t*-test were reported as the probability (p-value) that the difference between the means could occur by chance. The result of the one-way ANOVA test was also a probability (p-value) but this value represented the probability of variation among the different groups considered for the test – a low p-value indicated a high probability of the variation between categories. Multiple testing corrections were performed for all three tests using the Hochberg-Benjamini method with a FDR threshold of 5 %.

Comparison of all three groups by one-way ANOVA requires three or more replicates per treatment group and so this analysis was possible only for experiments that used datasets with unpooled samples. The one-way ANOVA test performed in Scaffold tests the equality of mean QVs across three or more treatment groups and low p-values indicate a difference among the categories. The test however, does not indicate what categories are different from each other. And so, the categories were compared two at a time using either Fisher's exact test or *t*-test. Fisher's exact test can be applied to fewer than three replicates and so it was used for the statistical analysis of datasets with pooled samples in which one sample represented a treatment group. The *t*-test was used to determine the statistical significance

in analysis datasets with unpooled samples. These datasets included four biological replicates for each of three treatment groups.

2.6.7. Functional analysis: Gene Ontology and KEGG pathways

WEB-based GENE Set Analysis Toolkit (WebGestalt) was used to interpret the protein lists in a biological context [243]. WebGestalt is a suite of tools for functional enrichment analysis in biological contexts; it compares the uploaded list with entries in predefined functional categories to identify those categories with enriched numbers of user-uploaded gene. The tools hosted by WebGestalt were used to identify possible functional enrichment of proteins in the Gene Ontology (GO) database and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. SwissProt identifiers (for example: P31946) of the proteins were provided as input “user IDs”. The human genome was used as a reference set for the enrichment analysis. A hypergeometric test was used for enrichment evaluation analysis with the significance level set at $p < 0.01$ using a Benjamini and Hochberg multiple test adjustment. A requirement of at least two genes was set for a category or pathway to be considered.

2.7 Validation of proteomics results

2.7.1 Western Blots

2.7.1.1. Separation of protein

HepG2 cells were harvested, washed with PBS and pelleted by centrifugation at 180 g for 3 min. The pellet was stored at -80°C . The sample for Western blots was a cell lysate obtained after re-suspending the cell pellet in lysis buffer. Protease inhibitor cocktail P8340 was added to the sample, and the sample was stored at -80°C . Since the lysis buffer did not contain urea, the protein content of the sample was assessed using the manufacturer’s protocol for the Bio-Rad Protein Assay [244].

The cell sample was separated using two-dimensional electrophoresis to validate the anti-perilipin 2 antibody. This included precipitation of 100 μg proteins using Calbiochem precipitation kit and re-suspending the proteins in 125 μL rehydration buffer [2 M thiourea, 7 M urea, 2 % pH 3-11 IPG buffer (GE healthcare, USA), 2 % DTT, 4 % CHAPS and trace bromophenol blue]. The sample was pipetted into the Immobiline DryStrip Reswelling tray (GE healthcare, USA) and the IPG strip was placed on the sample and covered with mineral oil to be left overnight for passive rehydration and loading of the sample onto the strip. The isoelectric focusing for the separation of protein in the first dimension was performed using

Ettan Multiphor II Electrophoresis system (GE Healthcare, USA) attached to EPS3501 XL power pack (GE Healthcare, USA) to setup the voltage gradient for separation and a temperature controlled water bath (Julabo, Seelbach, Germany) to maintain the temperature at 20 °C.

At the end of the focussing, the proteins on the strip was prepared for separation by SDS-PAGE on 4–12 % gradient NuPAGE gels (Invitrogen, USA; now Thermo Fisher Scientific) by reducing them in 1 % DTT in equilibrating buffer (6 M urea, 30 % glycerol, 2 % SDS and traces bromophenol blues in 50 mM Tris-HCl, pH 8.8) and alkylating them in 1 % iodoacetamide in equilibrating buffer. For Western blots of proteins other than perilipin 2, the separation was performed only in one-dimension using SDS-PAGE on 4–12 % gradient NuPAGE gels. Samples for these gels were prepared by reducing and alkylating aliquots with 10 µg proteins using NuPAGE® LDS Sample Buffer (Life Technologies, New Zealand) and NuPAGE® Sample Reducing Agent (Life Technologies, New Zealand).

A pre-stained fluorescent marker (Precision Plus Protein™ Dual Colour standards, Bio-Rad) was run along with the protein samples to indicate molecular weights. The separated proteins were stained with Coomassie blue G250 (Sigma-Aldrich) to confirm uniformity and accuracy of the protein quantification across different samples. Gels for Western blots were not stained. The separated proteins were transferred onto Hybond-LFP membrane (GE Healthcare Life Sciences) for Western blotting by wet electro-transfer in NuPAGE® transfer buffer for 1 h at 30 V. The transfer system was cooled by ice-cold water. The extent of transfer was assessed by post-transfer Coomassie blue staining of the gel.

2.7.1.2. Detection of bands

Membranes were placed in blocking solution [8 % (w/v) skimmed milk in Tris-buffered saline with 0.1 % Tween-20 (TBS-T)] for 3 h to prevent non-specific protein interactions between the membrane and the antibodies. The blocked membrane was incubated overnight at 4 °C with primary antibodies against the target proteins. Anti-rabbit Alexa Fluor 647 (Invitrogen) and anti-mouse Alexa Fluor 555 (Invitrogen) secondary antibodies were used at 1 in 5000 dilution. The protein concentration and optimum titre for each primary antibody was determined individually. Detected bands not corresponding to molecular weight of the target protein was considered a non-specific band. These were taken into consideration before selecting the target proteins that could be detected using multiplexed blots. Testing the antibody at different protein concentrations was done to confirm there was no

saturation of signal at the given antibody concentration [245]. Actin was used as a loading control [246]. The membrane was washed with four 7-min incubations in TBS-T before incubating with secondary antibody for 1 h at RT in the dark. Excess antibodies were removed by washing the membrane with four 7-min incubations in TBS-T, and the membrane was scanned on a Fujifilm FLA-5100 (Fuji Photo Film, Japan).

2.7.1.3. Quantification of bands

Bands were quantitated using ImageJ [247, 248]. Anti-rabbit Alexa Fluor 647 and anti-mouse Alexa Fluor 555 generated images in the red and green channel respectively. The channels were merged to give a RGB composite image which was then converted into 8-bit grey scale image to allow for quantification of the bands. ImageJ converts pixel intensities into optical density using the function $OD = \log_{10} (255 / \text{pixel value})$.

2.7.2 Immunofluorescent labelling of protein

HepG2 cells used for immunolabeling of perilipin 2 [also known as adipose differentiation-related protein or adipophilin] were grown on 13 mm sterile coverslips placed in wells of a 24-well plate and treated as described above. At the end of the treatment, the cells were fixed with 4 % paraformaldehyde for 20 min at RT [249]. The fixed cells were incubated in blocking buffer (1 % BSA in PBS with 0.05 % Tween 20) for 1 h at RT to minimize non-specific binding of antibodies [250]. At the end of the incubation, the cells were rinsed once with PBS and incubated overnight at 4 °C with polyclonal anti-perilipin 2 antibody (1 in 500 dilution) raised in rabbit. The primary antibody was diluted in 1 % BSA in PBS. The excess and unbound primary antibody was washed away by six 5-min washes of PBS, and the bound antibodies were detected using a 1 in 500 dilution of goat anti-rabbit secondary antibody (Abcam®) conjugated with Alexa Fluor® 488 (Life Technologies Corporation, USA). The excess and unbound secondary antibodies were removed with 6 washes of 5 min each. The washed coverslips were then lifted off the plate and mounted in 5 µL of Vectashield® anti-fade mounting medium with 4', 6-diamidino-2-phenylindole (DAPI) as nuclear counterstain. Fluorescent staining was examined with an Olympus FluoView FV1000 confocal laser scanning microscope using a 60× oil-immersion objective. Images were acquired in sequential imaging mode with dichrome filters allowing a bandwidth of 425 to 460 nm for DAPI and 485 to 545 nm for Alexa Fluor 488.

CHAPTER 3: DETERMINING TREATMENT CONDITIONS

3.1. Introduction

3.1.1. HepG2 cells as cell culture model for lipid accumulation

As described in section 1.4.1, lipid accumulation, largely on account of excessive inflow of FA from the adipose tissue into the liver, constitutes the first hit in disease progression. Treatment of hepatocyte cell lines and primary hepatocytes with mono-unsaturated and/or saturated FA has been shown to reproduce this key feature of NAFLD in humans [230, 251]. This project has used HepG2, a human hepatocarcinoma cell line, to study this feature and describe the proteomic changes that accompany it. Unlike primary hepatocytes, HepG2 cells are not limited in availability and can be maintained in cell culture media with more ease than hepatocytes. Moreover, these cells are capable of performing a wide variety of liver-specific functions related to cholesterol and TAG metabolism [252]. The use of HepG2 in assessing the effects of exogenous FA and ensuing lipid accumulation has also been well-documented [231, 253]. This chapter describes the effects of FA on HepG2 cells and outlines a treatment regime that would provide a HepG2-based cell-culture model that represents the proteomic changes associated with lipid accumulation. DGAT inhibitors have also been studied for their effect on lipid accumulation induced by such a treatment.

3.1.2. DGAT inhibitors

Three DGAT inhibitors have been tested in this project. They are DGAT1 inhibitors A92 and PF04 and DGAT2 inhibitor PF06 (Figure 11). A92 was developed by Abbott Laboratories (Abbott Park, IL, USA) and is a urea analogue of biaryl keto acids class of compounds that has been reported to cause weight loss and significant reductions in hepatic and serum TAG in mouse models [254, 255]. PF04 (Pfizer Global Research and Development, Groton, CT, USA) has a pyrimidooxazepinone structure and has been reported to have an IC_{50} (half maximal inhibitory concentration) of 19 nM for human DGAT1 in human intestinal epithelial cells (HT-29) cells [131]. It is an orally-active compound capable of reducing serum TAG concentration in rat models [256] and reducing blood glucose levels in patients with Type II diabetes [257]. The DGAT2 inhibitor PF06, also developed by Pfizer Global Research and Development, has been reported to be a highly selective, imidazopyridine-based inhibitor capable of reducing TAG synthesis in cultured human hepatocytes by as much as 50 % in conjunction with PF04 [258]. It was also capable of reducing serum TAG concentration in a rat model [258].

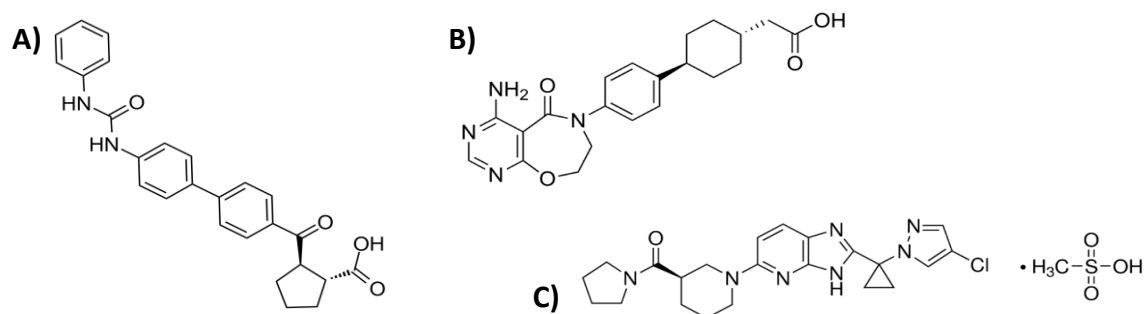


Figure 11: DGAT inhibitors used in this project. A) A922500; B) PF04620110 and C) PF06424439. The images have been sourced from the suppliers' catalogue.

3.1.2. Measuring lipid accumulation

Assays used to measure the TAG content and lipid accumulation are either enzyme-based commercial kits [259, 260] or use of lipophilic dyes like ORO or Nile red [261, 262]. While each commercial kit-based assay has its proprietary components, the steps involved include enzymatic hydrolysis of TAG by a lipase and measurement of the glycerol or FA produced [259, 260]. Nile red is a fluorescent lipophilic stain that exhibits solvatochromism [263]. The colour of solvatochromatic dyes changes with the change in solvent polarity, and so Nile red fluoresces yellow-gold (excitation, 450 - 500 nm; emission, > 528 nm) in the presence of neutral molecules like TAG but in the presence of polar molecules like phospholipids in cellular membranes it emits a red fluorescence (excitation, 515 - 560 nm; emission, > 590 nm) [264, 265]. The yellow-gold fluorescence has been used to measure accumulation of neutral lipids in the cell. ORO on the other hand is a non-fluorescent, fat-soluble diazole dye which stains neutral lipids but, unlike Nile red, does not stain biological membranes [266]. LDs stained with ORO appear red under a light microscope. The principle for staining is that ORO is more soluble in TAG-rich LDs than it is in the solvent (60 % isopropanol) used to prepare the staining solution [267]. This property has been used to stain TAG-rich LDs in the HepG2 cells used for the present study [171]. Since it exclusively stains the LDs in the cell, the amount of dye retained has been used as a measure of the extent of intracellular lipid accumulation. This chapter describes the use of ORO to characterise the extent of lipid accumulation induced by exogenous FA in HepG2 cells and identify the treatment conditions to be used for the proteomics experiment.

3.2 Results

3.2.1. Lipid accumulation in HepG2 cells

The aim of this project was to identify the proteomic changes that accompany lipid

accumulation in HepG2 cells. And, since the broader purpose of this objective was to identify molecular mechanisms involved in a diseased state like simple steatosis, a treatment condition that would induce lipid accumulation with minimal cytotoxicity was required. HepG2 cells were incubated with different concentrations of FA mixture for different treatment durations and two parameters – lipid accumulation and cell viability – were measured at each time point. The lipid accumulation was measured using the ORO assay described in section 2.3.3. and cell viability was measured using the MTT assay described in section 2.3.4. This section describes the results of these assays and based on these results, treating HepG2 cells with 1 mM FA mixture for a duration of 6 h was found suitable for the proteomics analysis. The exact duration of the treatment was confirmed after taking into consideration the effects of the DGAT inhibitors discussed in section 3.2.2.

3.2.1.1. Linear range of absorbance for Oil Red O

The ORO assay used to measure lipid accumulation in HepG2 cells was performed using the ORO staining solution prepared in our laboratory and so, a spectrum scan ranging from 300 to 600 nm was performed to determine a wavelength at which the absorbance of the working ORO solution could be measured. Although many studies have used ORO to measure lipid accumulation in HepG2 cells, there is no consensus on the measurement units or wavelength at which ORO absorbance is measured [68, 231, 253]. The spectrum scan performed identified a two ranges – 340 to 350 nm and 500 to 540 nm – within which absorbance of the dye was at its peak (Figure 12A). This was consistent with the UV–visible spectra reported for ORO [268]. Two wavelengths, 510 and 405 nm, have been cited in literature for quantification of ORO [231, 233, 262]. Absorbance of a series of two-fold diluted ORO working solutions was measured at both these wavelengths and, consistent with the spectrum scan, the absorbance of the ORO working solution at 510 nm was higher than that observed at 405 nm at all concentrations tested. And so, although the absorbance had a linear correlation ($R^2 \approx 0.98$) with the dye concentration for up to 0.525 mg/mL for both wavelengths (Figure 12B), all subsequent measurements were made at 510 nm. The linear correlation between the dye concentrations up to 0.525 mg/mL corresponded to an absorbance ranging from 0.1 to 0.93. Therefore, the volume of isopropanol used to extract the dye from the cells was adjusted to yield an extract that would have an absorbance between 0.15 to 0.93 and, since the dye from each cell sample was extracted using the same volume of isopropanol, it was assumed that any difference in the absorbance would be due to the difference in the amount of dye that had accumulated in the cell. The absorbance of

the ORO extracted from the treated cells was compared to that from the untreated ‘control’ cells and expressed as ‘fold of control’ to provide a relative measure of lipid accumulation in the cell as described in section 2.3.3.3.

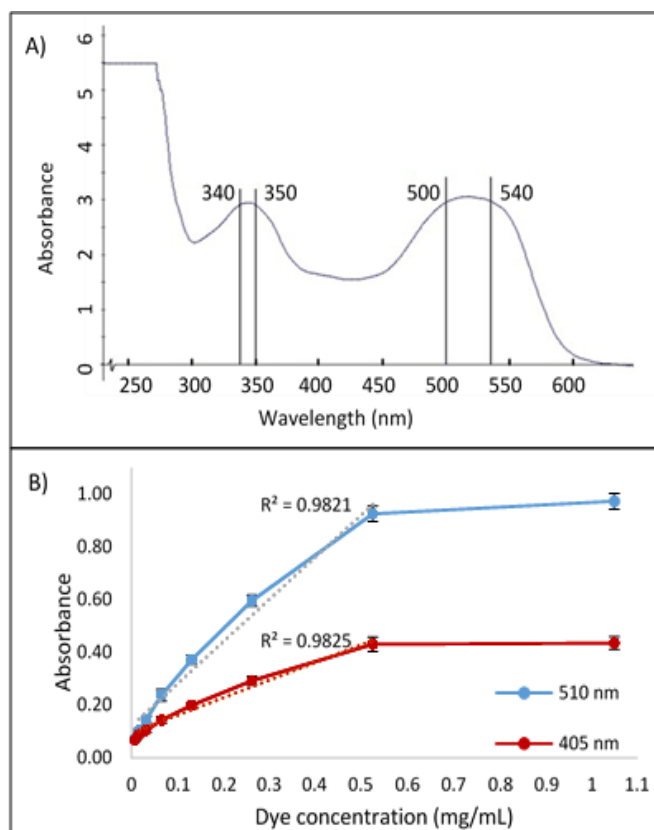


Figure 12: Absorbance of ORO at different wavelengths. A) Spectrum scan of undiluted ORO working solution. **B)** Absorbance of ORO working solution were measured over a range of 0.008 to 1.05 mg/mL at two wavelengths – 510 nm and 405 nm. This concentration range corresponds to two-fold serial dilution of the ORO working solution.

3.2.1.2. Inducing lipid accumulation

The concentrations of FA mixture tested ranged from 0.25 to 3 mM and each concentration was evaluated for lipid accumulation at treatment times ranging from 1 to 24 h. Exposure to the FA mixture led to a dose-dependent and time-dependent increase in lipid accumulation. The accumulated lipid droplets were seen using a phase-contrast microscope as refractile particles in the cytosol and were stained red by the lipophilic dye ORO (Figure 13)[171]. The bound dye was extracted and the absorbance of the extract was measured at 510 nm. To account for the loss of cells, absorbance was normalized using the MTT data described in section 3.2.1.3. The ratio of absorbance of the FA-treated sample to that of the control sample was expressed as the fold of control and provided a relative measure of lipid accumulation. The average value for each combination of treatment time and concentration

of FA mixture was compared for statistically significant differences using two-way ANOVA. This test also indicated a statistically significant interaction between the two variable factors that had affected the outcome i.e. treatment time and the concentration of FA mixture were likely to have synergistic effect on the extent of lipid accumulation.

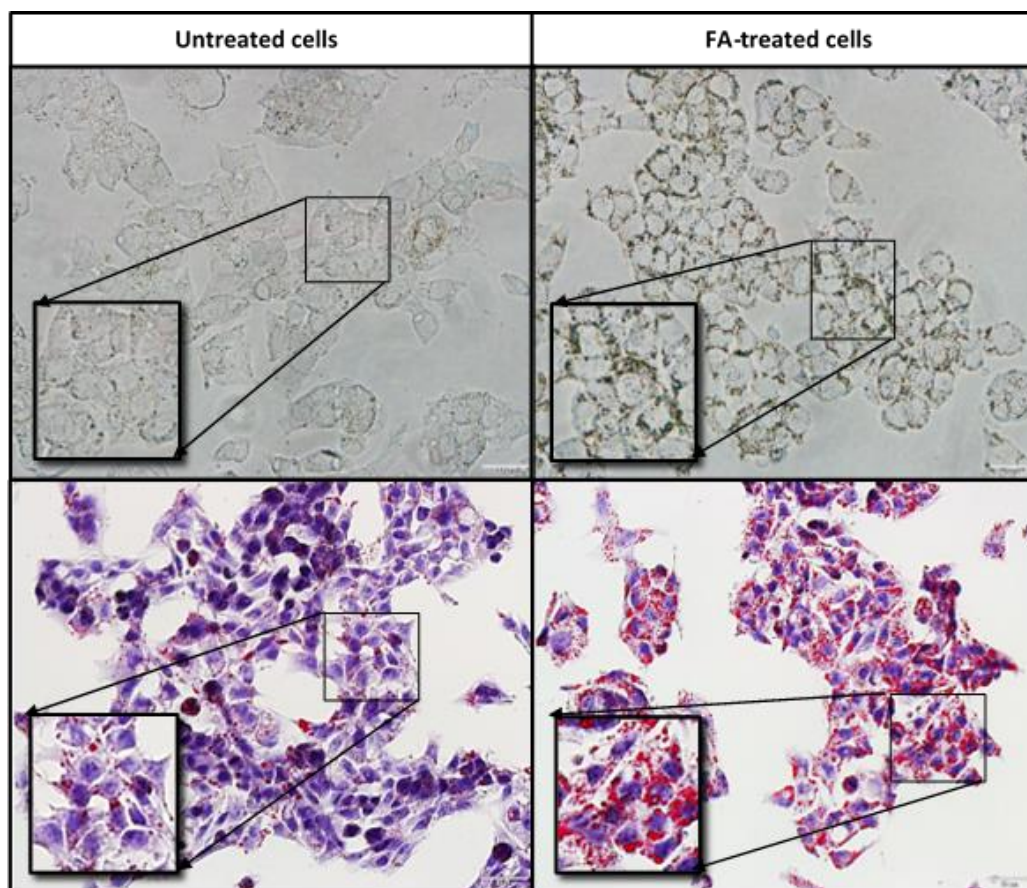


Figure 13: Visualization of lipid droplets in HepG2 cells. The lipid droplets appear as birefringent particles under a phase-contrast microscope in the upper panels. The lower panels shows lipid droplets stained red with ORO and the nucleus stained violet with haematoxylin against an eosin-stained cytoplasm. The insets have a digital magnification of 1.5x. The cells shown in this figure were treated with 1 mM FA for 12 h.

Consistent with this, 3 mM FA mixture had the most pronounced effect on lipid accumulation - treatment with this concentration of FA mixture produced significant effects within an hour of exposure while it took up to 18 h to induce a significant level of lipid accumulation using 0.25 mM FA (Figure 14). At the end of the 24 h treatment time, the extent of lipid accumulation induced by 3 mM FA mixture was about 6 times of control and the highest for all concentrations tested. The highest extent of lipid accumulation observed after a 24 h treatment with 0.25 and 0.5 mM FA mixture was 1.5 and 1.9 times of control respectively. Through ORO staining, Figure 15 shows the lipid accumulation induced by 1

mM FA mixture over a period of 24 h. This corresponds to lipid accumulations 2.5 times of control over a period of 24 h. The extent of lipid accumulation induced by 2 mM FA mixture for the same duration is about 4 times of control (Appendix Figures 66-70 show lipid accumulation in HepG2 cells treated with various concentrations of FA mixture and stained with ORO for observation and quantification.)

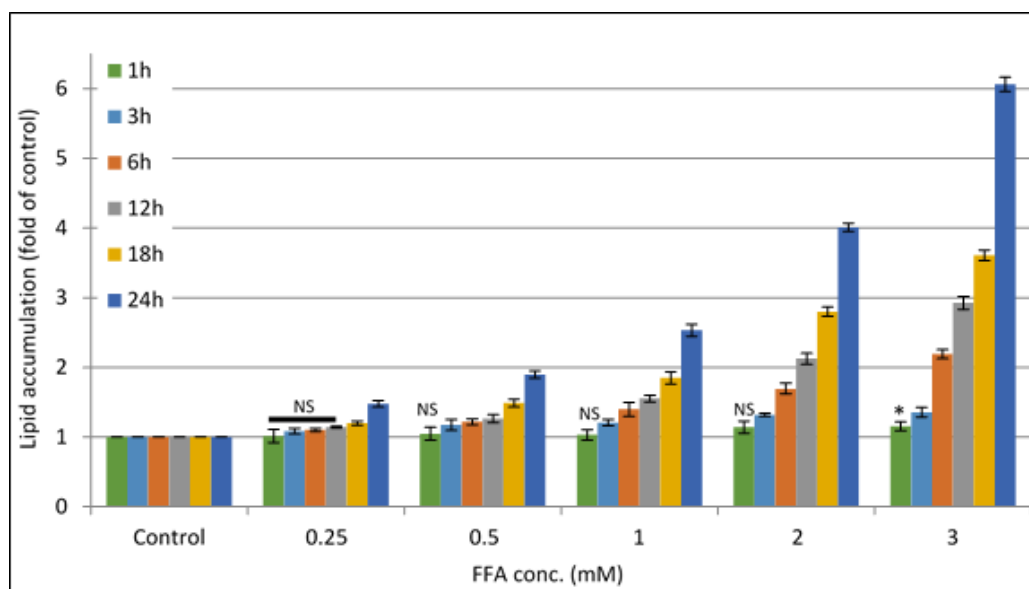


Figure 14: Lipid accumulation measured using ORO assay. In HepG2 cells, lipid accumulation post exposure to FA was found to increase with increase in treatment time as well as FA concentration. The ORO assay provides a relative quantification of lipid accumulation in terms of 'fold of control'. The values in the graph represent mean \pm SEM of four experiments with three technical replicates for each experiment. Values were significant at $p < 0.01$ unless otherwise specified (NS: No significant change with respect to the control at the same time point; * $p < 0.05$).

3.2.1.3. Cytotoxic effects of lipid accumulation

Cytotoxic effects of FA, especially saturated FAs, have been well documented in the literature [269-271]. This cytotoxicity and the consequent loss of cells was measured using the MTT assay. Living cells reduce MTT, a yellow tetrazole, to a purple formazan. The amount of formazan produced was estimated colourimetrically after solubilisation and a comparison of the optical densities of treated cells to that of the untreated, control cells provided a relative measure of how lipid accumulation may have affected the metabolic activity of the cell. Since non-viable HepG2 cells tend to detach from the plate surface, they were removed from the plate and thereby from the assay when the treatment media was replaced with 0.8 mg/mL MTT in DMEM at the end of the treatment time. As a result, this test also provided an indirect measure of cell viability defined here as the number of cells attached and metabolically active.

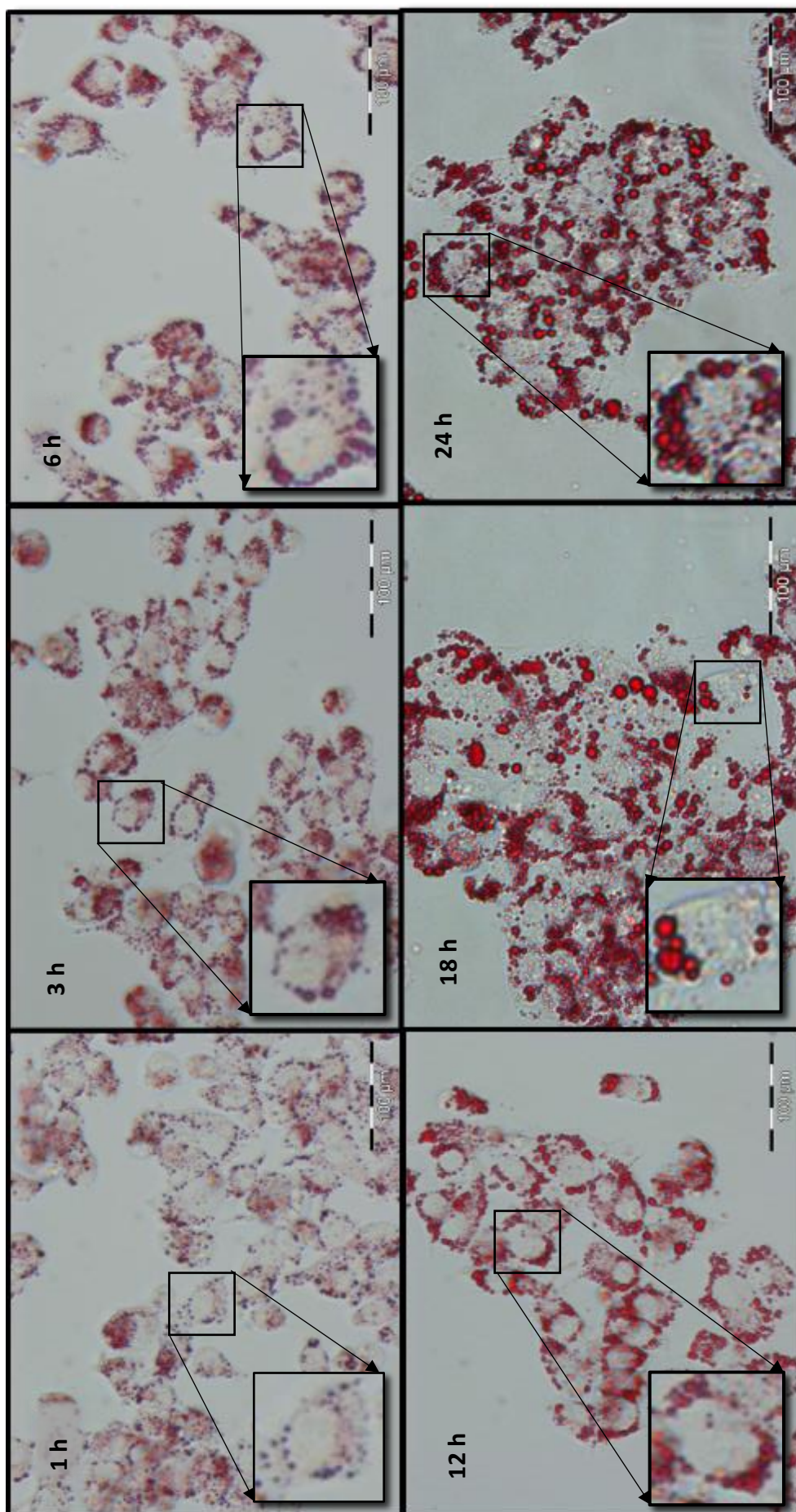


Figure 15: Lipid accumulation in HepG2 cells treated with 1 mM FA mixture for with different time lengths. Cells stained by ORO to show red lipid droplets accumulated within the cells treated with 1 mM FA mixture. Prolonged exposure produces larger lipid droplets that occupy almost all of the free space in the cell. At this point the usually flat and polygonal HepG2 cells seem to be larger and rounder. The round unstained regions in the centre of the cells are the nuclei which are unstained. The insets are at 2x digital magnification.

The cell viability thus estimated was used to normalize the ORO absorbance values and account for the loss of cells observed at the higher concentrations and/or after prolonged treatment times. The extent of cytotoxicity depended on the treatment time and the concentration of FA mixture tested with the effect being most pronounced at higher concentrations. Cells treated with 0.25 mM FA showed the least reduction in absorbance while those treated with 3 mM showed significant decrease within an hour of exposure. A 24 h exposure to 2 mM FA reduced the metabolic activity to almost half that of the control while a 3 mM FA mixture led to an almost 70 % decrease (Figure 16).

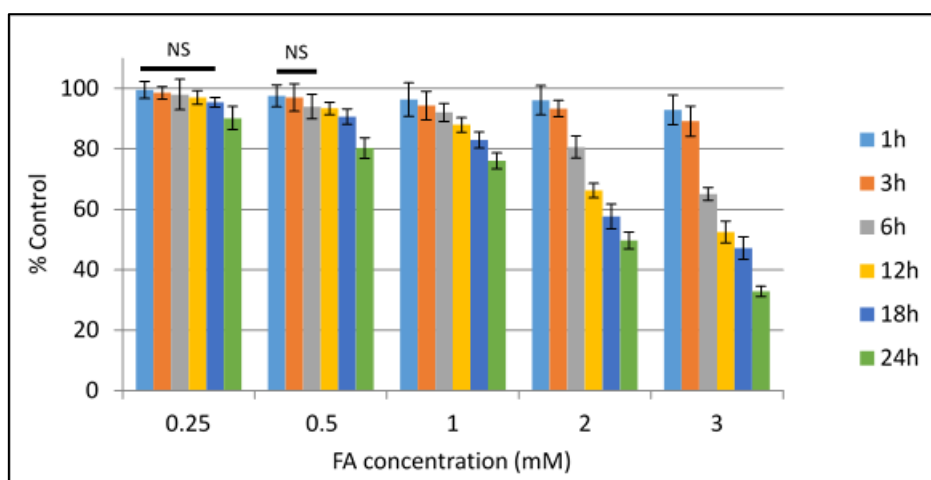


Figure 16: Effect of FA on metabolic activity assessed using MTT assay. Lipotoxic effects are most pronounced at FA concentrations of 1 mM and higher. The effects are exacerbated over period of time. The values represent mean % control \pm SD of four independent experiments with three biological replicates for each experiment. Values were considered significant at $p < 0.01$ unless otherwise specified (NS: No significant change with respect to the control at the same time point).

Lysed cells surrounded by a large number of ORO-stained LDs were observed at the end of 24 h treatment with 2 and 3 mM FA (Figure 17). The significant cytotoxic effects ruled out the use of 2 and 3 mM FA in further experiments. The cells treated with 1 mM FA mix had a cell viability of up to 92 % of control for a 6 h treatment time; the cell viability reduced to more steadily thereafter. At 18 h treatment duration, the cell viability was found to be about 80 % of control cells and this reduced to about 76 % of control cells after 24 h treatment duration. The increase in lipid accumulation observed within 6 to 18 h of treatment with 1 mM FA mixture appeared to be sufficient to allow detection of a DGAT inhibitor-induced decrease in lipid accumulation. Since the study design requires a system with low cytotoxicity but with maximum possible lipid accumulation, 1 mM FA mixture was selected as a suitable treatment condition to study the changes brought about by lipid accumulation.

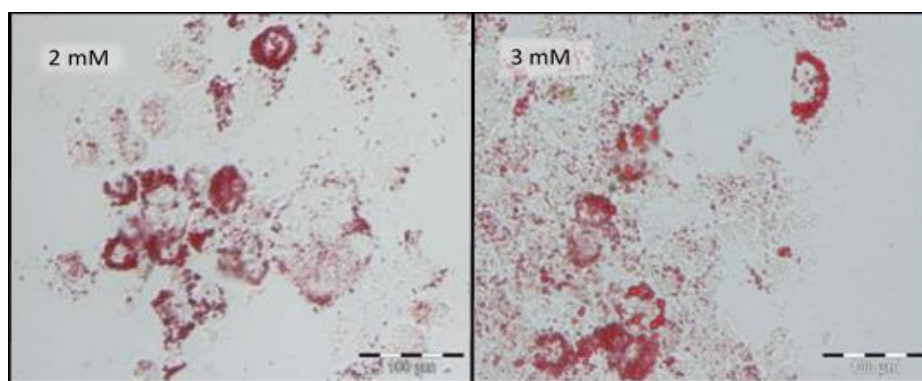


Figure 17: HepG2 cells treated with high concentration of FA mixture. Prolonged treatment with high concentration (2 and 3 mM) of FA mixture leads to significant reduction in cell viability and the remnants of lysed cells stained with ORO are shown in this figure.

3.2.2. Effect of DGAT inhibitors on lipid accumulation

3.2.2.1. DGAT1 inhibitor A922500

The DGAT1 inhibitor A92 is reported to be a potent, orally-active inhibitor with an IC_{50} value of 7 nM for recombinant human DGAT1 protein [254]. A92 is also reported to be very specific for DGAT1 with IC_{50} values for DGAT2 and ACAT1/ACAT2 of 53 μ M and 296 μ M, respectively [254, 272]. In HepG2 cells, A92 has been reported to inhibit incorporation of radio-labelled oleic acid into TAG at concentrations as low as 0.1 μ M [273]. In my work, concentrations ranging from 1 μ M to 100 μ M A92 were tested in the presence of 1 mM FA mixture. The extent of lipid accumulation was measured using the colorimetric ORO assay. Various combinations of suitable treatment time and inhibitor concentration were tried. Initial experiments of the treatment of HepG2 cells with 1 mM FA for 6 h in the presence of 100 μ M A92 showed reduction of lipid accumulation by 22.8 % with respect to the cells treated with the FA mixtures without the inhibitor. The treatment was tested in three independent experiments using cells from the same batch but different passage numbers, however, the effects of the inhibitor could not be reproduced. Each batch of cells originated from a single vial of stock cells, stored under liquid nitrogen, originally purchased from ATCC. The effect of A92 was also tested on cells of different batches originating in different vials frozen down at different times but the results remained irreproducible (Figure 18 and 19). Each batch of cells was tested for mycoplasma and, having ruled out such contamination, different culture conditions such as longer incubation periods and lower concentrations of FA mix were also tested. The inhibitor did not have any effect on lipid accumulation in these

altered conditions either. Additionally, the DGAT1 inhibitor was also tested for stability and effectiveness in a cell-free system. This section describes the results of these assays.

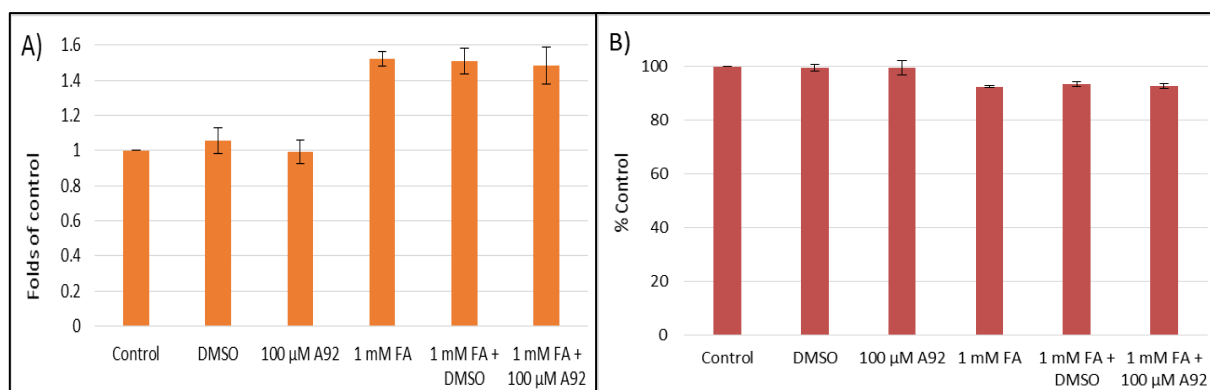


Figure 18: Effect of 100 μ M A92 on HepG2 cells. A) Lipid accumulation measured using ORO assay after 6 h treatment with 1 mM FA mixture; **B)** Metabolic activity of cells as measured by MTT assay after 6 h treatment with 1 mM FA mixture.

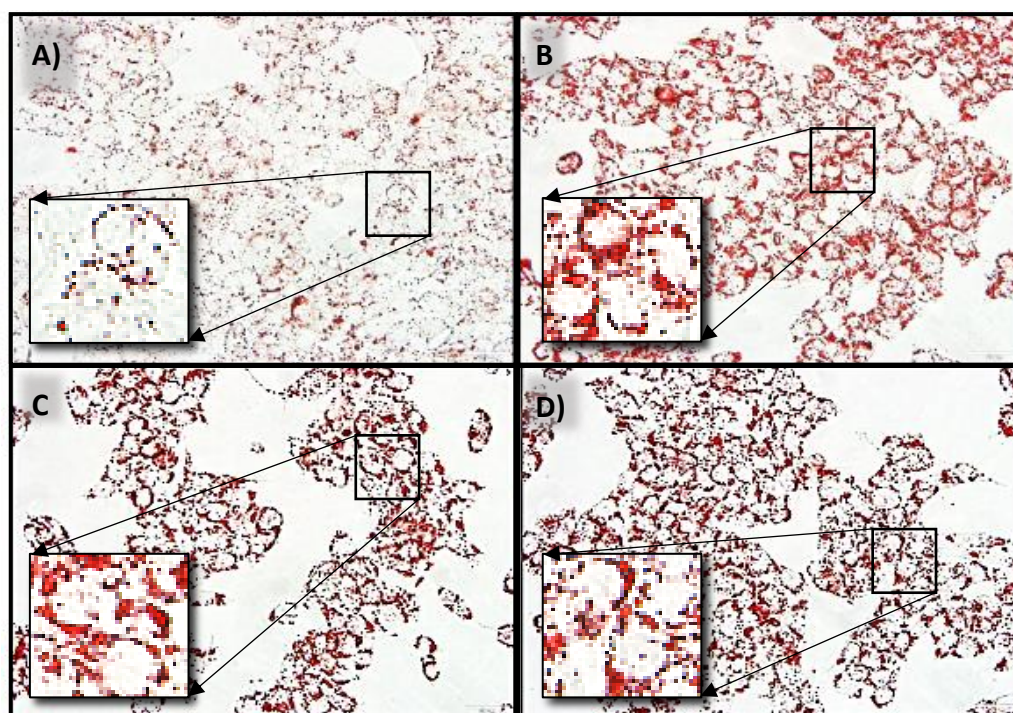


Figure 19: Lipid accumulation in HepG2 cells treated with 100 μ M A92. A) Control cells; **B)** Cells treated 1 mM FA mixture for 6 h only; **C)** Cells were treated with 1 mM FA mixture with 0.2 % DMSO to represent solvent control for the effect of A92; **D)** Cells treated with 1 mM FA mixture in presence of 100 μ M A92.

3.2.2.1.1. Effect of experimental conditions on the effect of A92 on cultured cells

The initial experimental conditions tested included incubation of cells with 100 μ M A92 in DMEM with 1 % BSA for 1 h prior to the addition of FA mixture to the medium. After the addition of FA mixture, the cells were incubated further for 6 h. Three changes to this protocol were tested: 1) omitting the pre-incubation, 2) increasing the incubation time after

the addition of FA mixture to 18 h and 3) using medium without BSA. The prior incubation with the DGAT inhibitor was omitted to minimize any possible break-down or metabolism of the inhibitor molecule before its effect could take place. Increased incubation time was tested to allow maximum possible time for the inhibitor to have an effect on the cell. The removal of BSA from the medium was considered to rule out any possible interference of the BSA with the uptake of inhibitor molecule. None of these changes in the treatment conditions led to inhibition of lipid accumulation as seen in the preliminary experiments. Incubation with lower concentrations of FA mixture (0.25 mM and 0.5 mM) for 6 h was also tested, but again, no inhibition was seen (Figure 20).

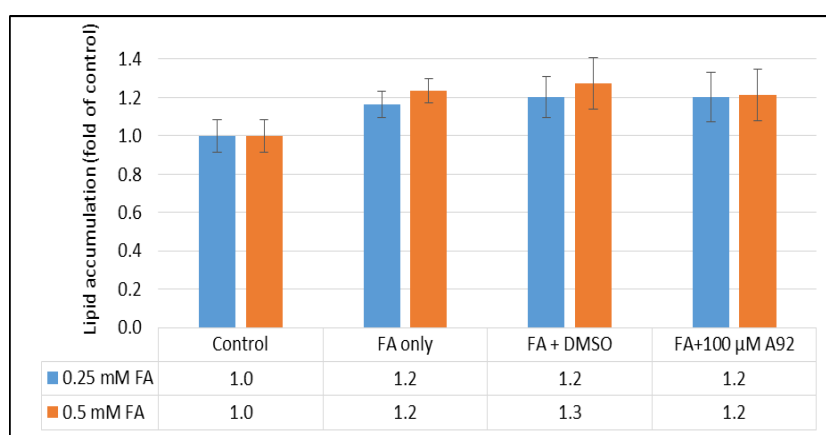


Figure 20: Lipid accumulation in presence of DGAT inhibitor A92. HepG2 cells were treated 100 μ M A92 in presence of 0.25 mM or 0.5 mM FA with a control for DMSO. The lipid accumulation observed in HepG2 cells treated with 100 μ M A92 was not significantly different from that observed in cells treated with FA alone.

The reasons for inconsistent inhibition were not clear although we assumed it might be related to the cell culture protocols especially trypsinisation and preparation of cell stocks for cryopreservation. It appeared at that time that the cells would attach but stop proliferating after 8 to 10 passages and any further passage would result in no cells adhering to the flask surface. Since the problem persisted in all cultures generated from stock vials available at our facility, we sourced a frozen vial of HepG2 cells from the Malaghan Institute of Medical Research (New Zealand) and used the cells from this vial for the rest of the study.

3.2.2.1.2. Stability of A922500

A 50 mM A92 stock solution was prepared in DMSO and stored at -20 $^{\circ}$ C for no more than three months before use. Although this conforms to the manufacturer-recommended storage conditions, the stability of A92 in the stock solution was confirmed by mass spectrometry using electrospray ionization in a Q-TOF LC/MS mass spectrometer (Agilent

Technologies 6530). The results showed that ions corresponding to $[M+H]^+$ (calculated m/z : 429.1809; measured m/z : 429.1827) and $[2M+H]^+$ (calculated m/z : 857.3618; measured m/z : 857.3583) for A92 (formula: $C_{26}H_{24}N_2O_4$) were detected (Figure 21). The absence of any other detected ions ruled out the possibility of contamination or possible degradation of A92 in the stock solution as the cause of its lack of inhibitory effect in the cell culture assay.

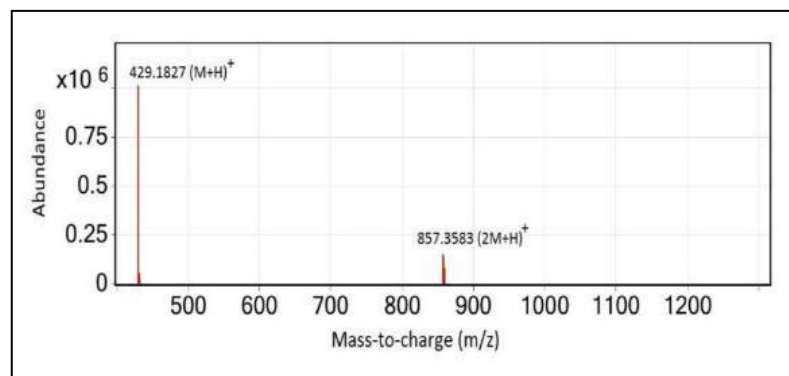


Figure 21: Spectral analysis of A922500 stock solution. The stock solution was analysed using the Agilent Accurate-Mass™ Q-TOF system in the positive ion mode. The sample was considered uncontaminated since only peaks corresponding to A92 were observed.

3.2.2.1.3. Inhibition of *in vitro* DGAT activity

Having confirmed the stability of the stock solution, an *in vitro* assay was used to test the activity of A92 in the absence of cells. The reaction mixture containing DGAT substrates dioleoyl glycerol and NBD-palmitoyl CoA was prepared as described in section 2.5.2, and, after incubating the mixture for 30 min at 37 °C, the components of the mixture were separated using TLC as described in section 2.5.3. Only molecules that would have incorporated the fluorophore would be detected by the fluorescence scanner. A preliminary experiment was performed using the mobile phase described by McFie *et al.* in the original article describing this assay [236]. This mobile phase consisted of *n*-hexane: diethyl ether: acetic acid (80: 20: 1; v/v). Separation of bands in these experiments was poor and the possibility of unresolved bands could not be ruled out. The band that had moved furthest from the loading spot had a retention factor (R_f) of 0.12 when the solvent was allowed to run to 18 cm from the base of a 20 cm plate (Figure 22A). A second band was seen very close to the loading spot (R_f = 0.05). The presence of a band superimposed on the loading spot indicated suggested a need for a mobile phase that gave better resolution. A mobile phase consisting of diethyl ether: *n*-hexane: methanol: acetic acid (55: 45: 5: 1, v/v) was tried on the recommendation of the authors of the original article [119, 236]. Although the band superimposed on the loading spot persisted, the sample did resolve into more bands and the

separation of the bands was also improved separation as seen by the higher Rf values (Figure 22B). The TAG product band was identified by its absence in the negative control, and had moved the farthest from the loading spot. Based on this result, the mobile phase of diethylether: *n*-hexane: methanol: acetic acid (55: 45: 5: 1, v/v) was used to separate the reaction products of the *in vitro* assay for DGAT activity. The ability of A92 to inhibit microsomal DGAT activity was tested using the modified mobile phase. In this assay, A92 inhibited all microsomal DGAT activity at 50 μ M and 100 μ M (Figure 23). This indicated that, in spite of the lack of inhibition in the cell-culture based assay, A92 was capable of inhibiting DGAT activity without the need for *in vivo* processing or metabolism that might occur in cells or intact animal models.

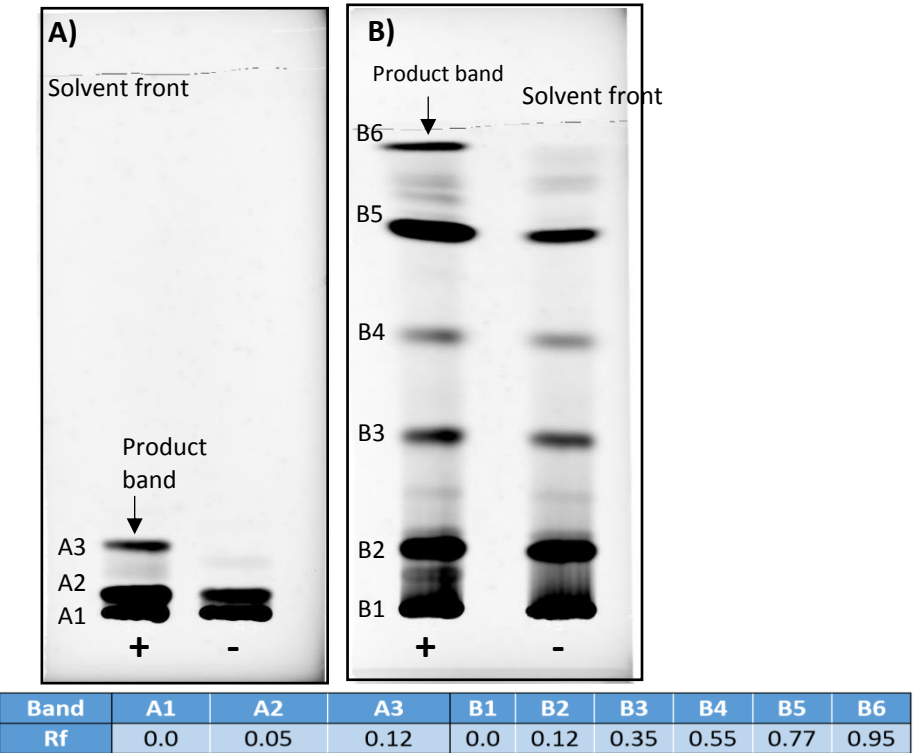


Figure 22: Comparison of mobile phases used for the separation of DGAT assay reaction mixture by TLC. The lane marked '+' represents a positive control which had all components of the reaction mixture. The '-' lane is a negative control which had heat-inactivated enzymes. The reaction mixtures were separated using different solvent systems as mobile phases and TAG reaction product band was inferred by its absence in the negative control. It corresponds to the bands A3 and B6 on their respective chromatograms. **A)** Products separated using a mobile phase of *n*-hexane: diethyl ether: acetic acid (80: 20: 1; v/v). **B)** Mobile phase of diethyl ether: *n*-hexane: methanol: acetic acid (55: 45: 5: 1, v/v) was used to separate reaction products.

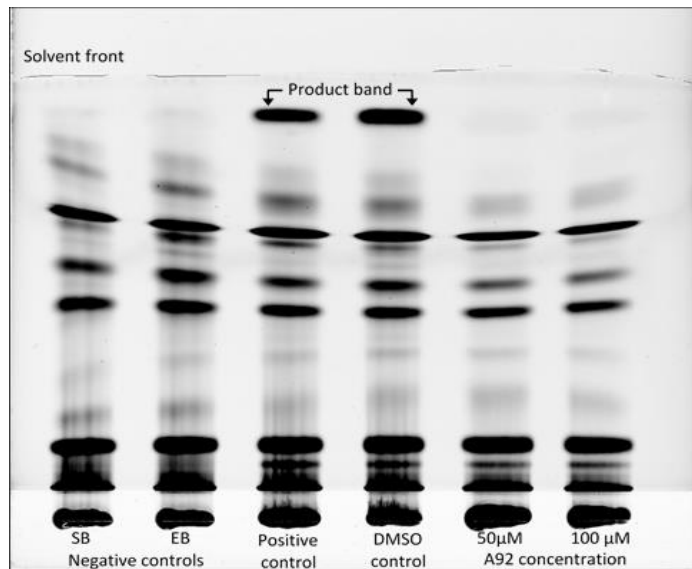


Figure 23: Inhibition of microsomal DGAT activity by A922500. The negative control 'SB' was a substrate blank which included NBD-palmitoyl CoA but not dioleoylglycerol and 'EB' was an enzyme blank that included both substrates but the enzyme was heat-inactivated. Positive control had all the reaction components. The DMSO control was included to check the effect of inhibitor solvent on the reaction. The reaction systems in both of these controls was able to form the TAG product band which is absent in the negative controls. DGAT1 inhibitor A92 was tested at 50 and 100 μM and was able to inhibit microsomal DGAT activity at both concentrations.

3.2.2.1.4. Expression of DGAT enzymes

It was possible that up-regulation of DGAT2 may have compensated for inhibition of DGAT1 to maintain enough DGAT activity to continue producing TAG in the presence of FA [274]. Expression of DGAT1 and DGAT2 were therefore verified using western blots. There was no significant change in the abundance of either protein (Figure 24).

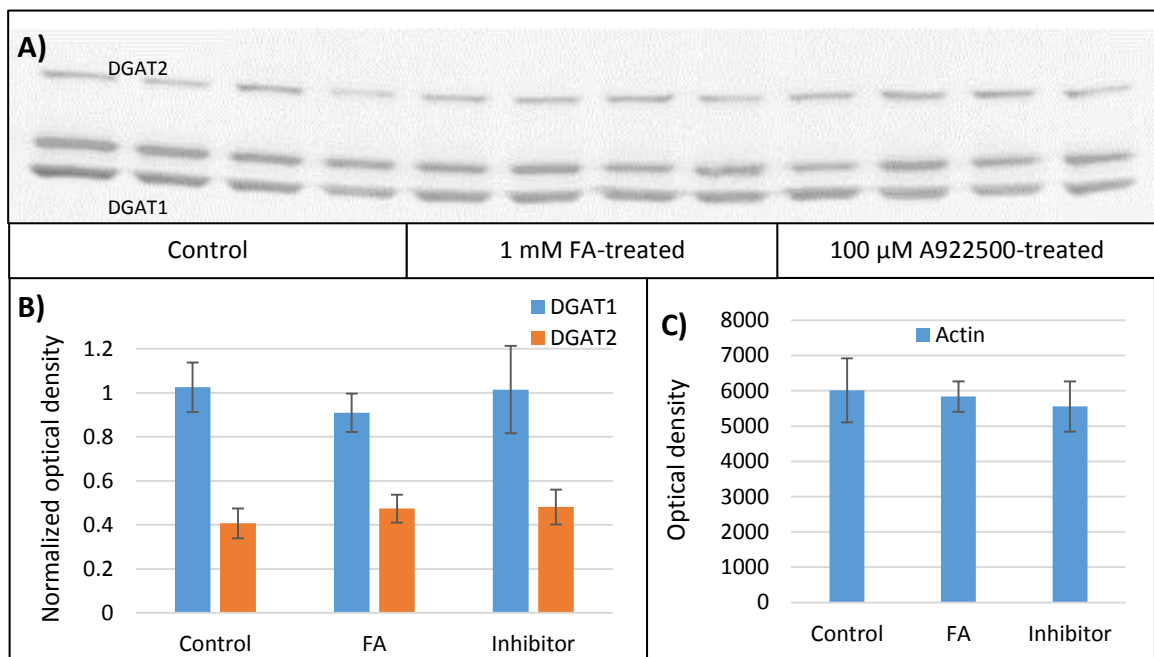


Figure 24: Western blots of DGAT enzymes in HepG2 cells treated with 100 μM A92. **A)** Each set of four lanes in the blot represents a set of four biological replicates treated as indicated. **B)** Normalized optical density was calculated as the ratio of optical density of the DGAT band to the optical density of actin in the same lane. No changes in protein expression were observed for either DGAT enzyme. **C)** Optical density of actin used as the loading control.

3.2.2.2. Pfizer DGAT inhibitors – PF04 and PF06.

The use of inhibitors to block DGAT activity was limited by the commercial availability of molecules other than A92. However, during the current study, the DGAT inhibitors PF04 and PF06 became available for testing [258, 275]. PF04 is a DGAT1 inhibitor, and PF06 is a DGAT2 inhibitor. The efficacy of these molecules was first tested in the *in vitro* assay for DGAT activity at different concentrations of Mg^{2+} (section 3.2.2.2.1). Section 3.2.2.2.2 describes the results of the ORO assay and the MTT assay performed to test the effects of DGAT1 inhibitor PF04 on lipid accumulation in HepG2 cells treated with 1 mM FA mixture for 6 h.

3.2.2.2.1. Effect of PF04 and PF06 on *in vitro* microsomal DGAT activity

DGAT inhibitors PF04 and PF06 were tested in the *in vitro* DGAT assay at final concentrations of 5, 25, 50, 75 and 100 μ M. The reaction mixture used dioleoylglycerol and NBD-palmitoyl CoA as the substrates and included 20 mM $MgCl_2$ in the reaction mixture. While PF04, the DGAT1 inhibitor, was able to inhibit all DGAT activity even at the lowest concentration tested, DGAT2 inhibitor PF06 was not able to produce the same effect (Figure 25). Although concentration of $MgCl_2$ in the *in vitro* assay is much lower than the $[Mg^{2+}]$ reported to adversely affect DGAT2 activity [100], the assay was repeated at lower concentrations of $MgCl_2$ to confirm the lack of effect described above for PF06. The reduction in the $MgCl_2$ did not help; the product formed in the presence of 100 μ M PF06 was comparable to that seen in the positive control while the presence of 100 μ M PF04 was able to block TAG formation completely at all concentrations of $MgCl_2$ in the same assay (Figure 26).

3.2.2.2.2. Effect of PF04 and PF06 in cell culture- based assays

Different concentrations of PF06, PF04 and an equivalent volume of DMSO were added to the media to test their effect on lipid accumulation in the cell culture-based ORO assay. An MTT assay was performed parallel to the ORO assay. The lipid accumulation seen in cells treated with 5 and 25 μ M DGAT1 inhibitor PF04 was up to 15.5 % less than that in cells treated with 1 mM FA mixture alone. However, the effects of the inhibitor were statistically significant only at concentrations of 50 μ M and higher. DGAT2 inhibitor PF06 had no effect at 5 μ M. At concentrations of 25 to 100 μ M, PF06 caused an average decrease of 6.3 % in lipid accumulation but the decrease was not statistically significant. The two inhibitors were also tested in combination at 75 μ M and 100 μ M. No additive effect was seen, and the extent of inhibition observed was similar to that seen for PF04 alone (Figure 27). The dose-dependent effect of PF04 on lipid accumulation in cells using ORO is presented in Figure 28.

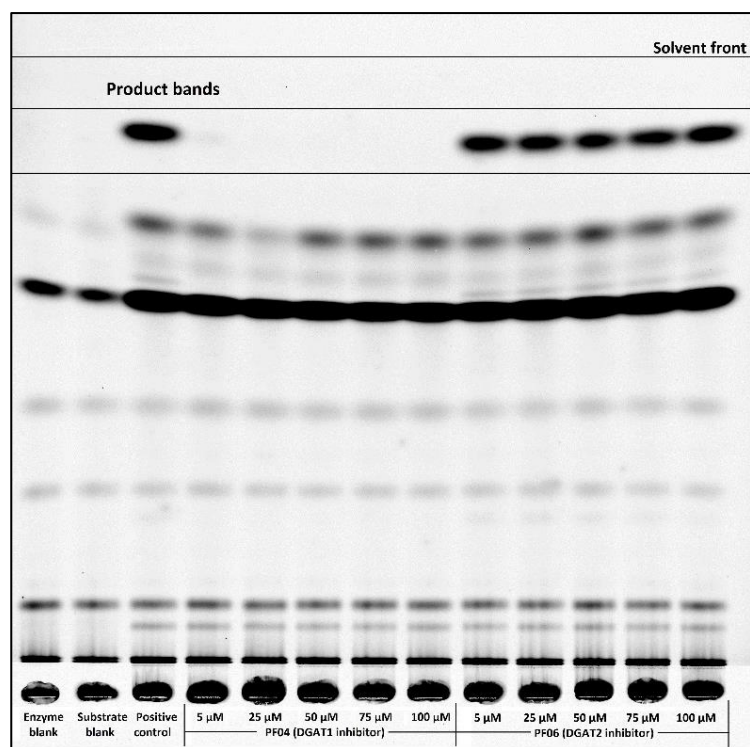


Figure 25: Effect of PF04 and PF06 on *in vitro* microsomal DGAT activity. Different concentrations of DGAT inhibitors PF04 and PF06 were tested as indicated. The negative controls included were enzyme blank (EB) and substrate blank (SB). Inhibition of DGAT activity was inferred by comparing the tests to the positive control (+ve control). The corresponding product band was absent for all concentrations of PF04 tested.

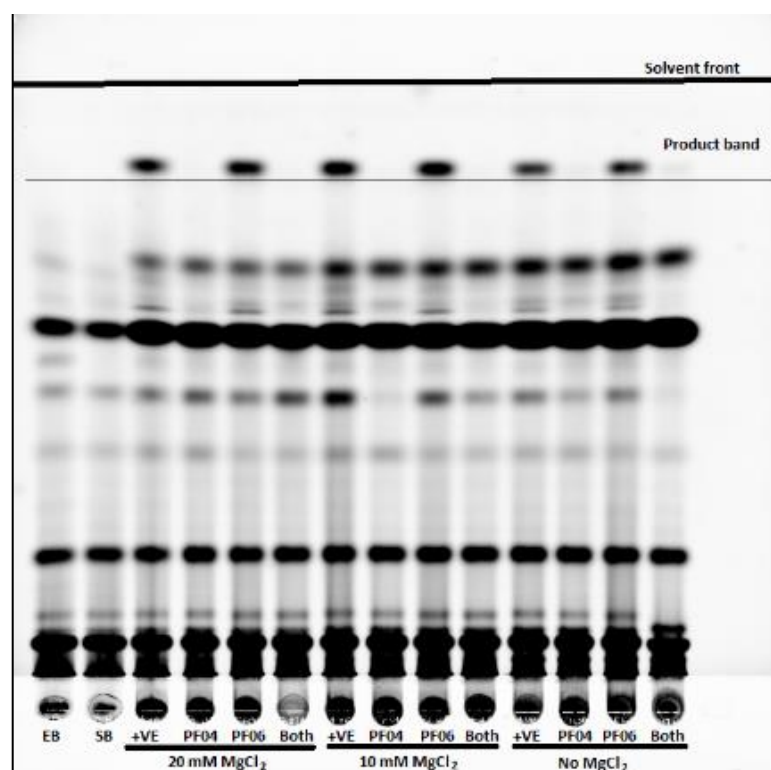


Figure 26: Effect of Mg^{2+} ions on DGAT activity. Activities of DGAT1 and DGAT2 can be distinguished based on presence of Mg^{2+} ions in the reaction buffer – the latter is inhibited at higher concentrations. PF04 was able to block microsomal DGAT activity irrespective of the $[Mg^{2+}]$ while PF06 had no effect despite the reduced $[Mg^{2+}]$.

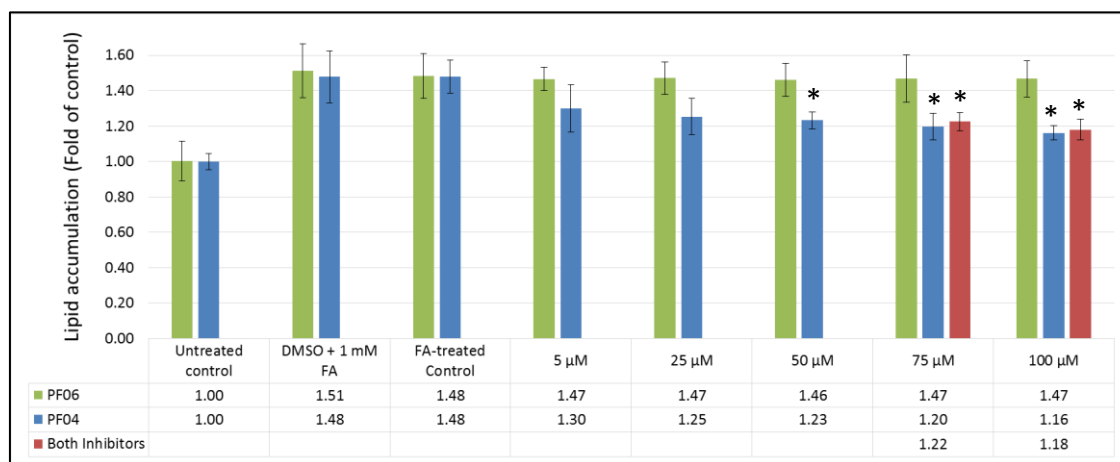


Figure 28: Effect of PF04 and PF06 on lipid accumulation in HepG2 cells. HepG2 cells were treated with different concentrations of DGAT inhibitors in the presence of 1 mM FA mixture for 6 h. Reduction in lipid accumulation was considered significant for $p < 0.05$ and the values in the chart represent mean \pm SD of four experiments with three measurements per experiment.

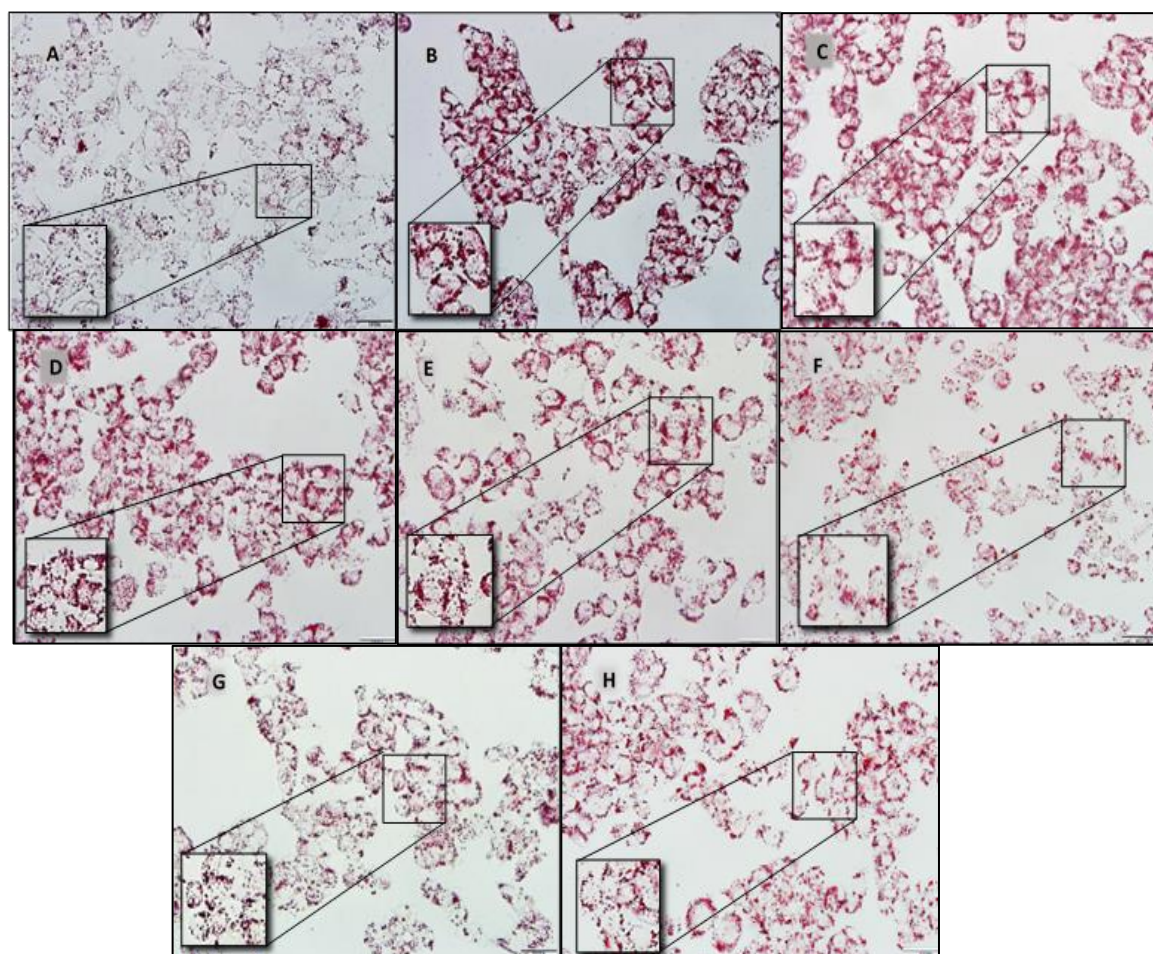


Figure 27: ORO staining of HepG2 cells treated with different concentrations of PF04. A) Untreated control; B) Cells treated with 1 mM FA mixture; C) Cells treated with 1 mM FA mixture with 0.2 % DMSO; D – H) Cells treated with 1 mM FA mixture in the presence of 5, 25, 50, 75 and 100 μ M PF04, respectively.

The metabolic activity of cells treated with 1 mM FA mixture for 6 h was about 92 % of the untreated HepG2 cells. The presence of DGAT inhibitors did not rescue the cells from this decrease in MTT absorbance. There was no statistically significant difference between the MTT values observed for FA-treated cells and those treated with 1 mM FA mixture in the presence of the DGAT inhibitors (Figure 29). Having confirmed that PF04 does not affect the expression of the protein (Figure 30), treatment of HepG2 cells with 100 μ M PF04 was selected for the proteomic analysis of the changes that accompany lipid accumulation in the presence of DGAT inhibitors.

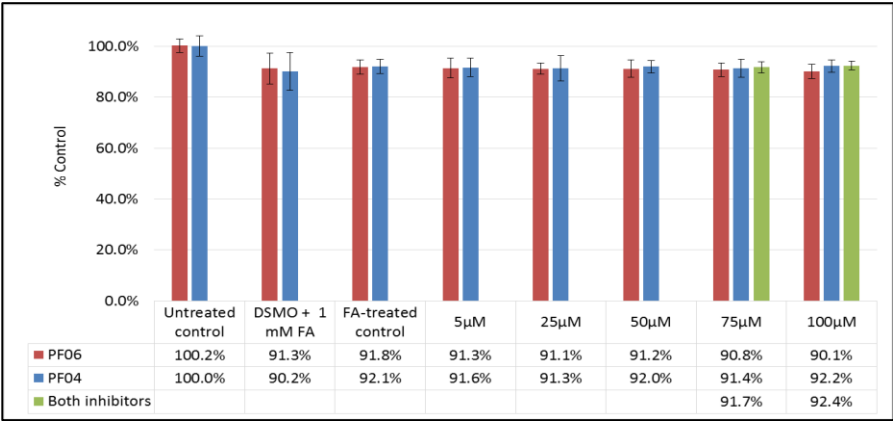


Figure 29: Effect of DGAT inhibitors on metabolic activity of HepG2 cells treated with 1 mM FA mixture for 6 h measured by MTT assay. The values represent mean % control \pm SD of four independent experiments with three biological replicates for each experiment. The treated cells were compared to the untreated control cells and no significant changes were observed. The difference in the mean values would have been significant for values of $p < 0.05$.

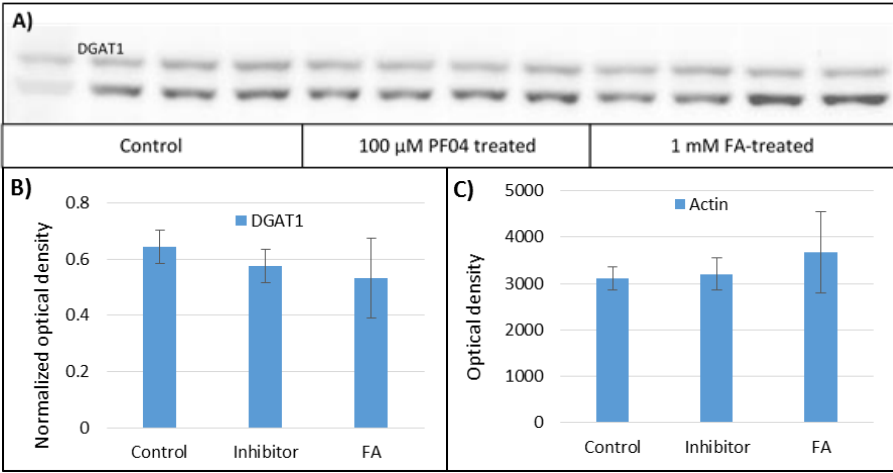


Figure 30: Western blot for DGAT1 in HepG2 cells treated with 100 μ M PF04. Each set of four lanes represents a set of four biological replicates treated as indicated. Lane 1 (from left) showed poor transfer and was not included in the calculations. Actin was used as the loading control. Normalized optical density was calculated as the ratio of optical density of the DGAT enzyme band to the optical density of actin in the same lane. No statistically significant ($p < 0.05$) changes in protein expression were observed.

3.3. Summary

Lipid accumulation in cells treated with 1 mM FA mixture for 6 h was found to be 1.4 (\pm 0.01) fold of control cells. This was reduced to about 1.2 fold in the presence of 100 μ M of the DGAT1 inhibitor PF04. Although the DGAT1 inhibitor A92 was effective in initial experiments, it later had no effect under the same conditions. The DGAT2 inhibitor PF06 also had no effect on lipid accumulation in HepG2 cells treated with 1 mM FA mixture for 6 h. While A92 did show an effect in *in vitro* assays, the DGAT2 inhibitor PF06 was unable to inhibit the microsomal DGAT activity, or the activity in cultured cells. Further use of the DGAT2 inhibitor was discontinued.

The treatment conditions described in this chapter were the basis of a proteomic analysis of the cell-culture model for lipid accumulation. Untreated cells represented the normal proteome of HepG2 cells and served as a negative control for the effects of lipid accumulation on protein expression. Cells treated with 1 mM FA mixture for 6 h represented a proteome that allowed cells to cope with excess lipid accumulation. Treating cells with 100 μ M PF04 in the presence of 1 mM FA mixture for 6 h was used to test whether the proteomic changes due to accumulation of FA could be reversed by inhibition of DGAT1 in the cell. Cells treated with 100 μ M A92 were also included in one of the initial proteomics datasets, as described in the following chapter4.

CHAPTER 4: PROTEOMICS ANALYSIS – PART I

4.1. Introduction

The primary objective of this study was to identify proteomic changes induced as a result of excessive lipid accumulation. It also aimed at testing the effect of DGAT inhibition on the changes caused by lipid accumulation. Based on the effects described in chapter 3, HepG2 cells treated with 1 mM FA mixture for 6 h were assumed to represent a state in which the lipid accumulation was significant but protective mechanisms at play were able to ensure cell survival. An analysis of these proteomic changes using “bottom-up” shotgun proteomics and label-free quantification was performed to investigate these protective mechanisms. This approach was considered useful since it aims to identify as many changes as possible and, instead of being driven by a hypothesis, the analysis aimed at generating hypotheses for testing through targeted studies in the future. As described in section 2.6., the approach involved digestion of proteins by trypsin followed by MS analysis of the complex mixtures of peptides used to identify the proteins present (Figure 31).

Sample preparation, separation of analytes and their detection are the key determinants in the number of proteins and changes in abundance that can be identified. Sample preparation includes all the steps taken to ensure the extraction and solubilisation of proteins and these would include the use of various cell lysis techniques [276, 277]. The

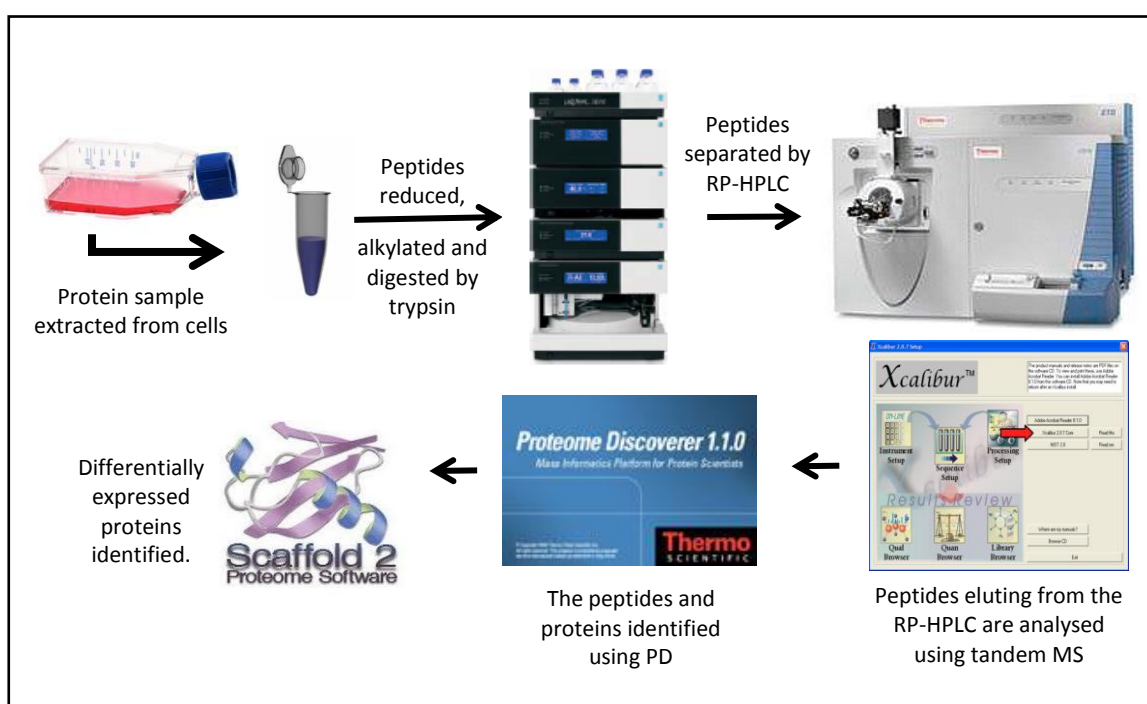


Figure 31: Schematic representation of proteomics workflow. The schematic shows an overview of steps involved in a proteomics experiment: proteins are extracted from the cell, digested by trypsin and analysed using LC-MS/MS. Appendix Table 20 describes the file formats used by each software.

choice of technique depends on the experimental design and has to be mindful of chemical compatibility with the LC-MS process [179]. Reverse-phase LC is the predominant technique used for separation of the peptides since it is capable of high-resolution separation capacity and the mobile phase solvents such as acetonitrile and methanol are compatible with electrospray ionization method used to generate the analyte ions [278, 279]. Chromatographic peak capacity is defined as “the maximum number of resolvable peaks” [280] and provides a measure of how well the analytes have separated [281]. Since each peak represents an MS input that can be translated into a peptide and the peptide in turn assigned to a protein for identification, the number of peptides and proteins identified in a LC-MS/MS run have been considered a measure of the quality of the LC-MS data in this work. Assuming adequate separation of peptides, the relevant MS settings and the parameters of the database protein search are key determinants of the quality and quantity of data acquired for a given sample. This chapter describes the results of the LC-MS/MS dataset LC1 that set up the need to improve the number of proteins identified. The improvement was brought about by modifying the parameters used for data collection. The increase in the number of proteins is reflected in the number of proteins identified in the subsequent analysis of dataset LC2 also described in this chapter. The samples for dataset LC1 were prepared from lysates of cells grown from the ATCC stock vial. Based on the effect shown by the DGAT inhibitor A92 in this culture system, the cells treated with 100 μ M A92 and 1 mM FA mix were included in the analysis of dataset LC1. Dataset LC2 was generated from samples prepared from lysates of cells that originate in stock vial sourced from the Malaghan Institute. As described in detail in Chapter 3, the effect of A92 seen in the ATCC cells could not be reproduced and hence dataset LC2 does not contain the inhibitor-treated cells as a treatment group. PF04 was not commercially available for testing at the time.

4.2. Results

4.2.1. Selection of a lysis buffer for sample preparation

Efficient cell lysis and good solubilisation of a wide range of proteins are vital to generating data representative of the sample. Cell pellets of untreated HepG2 cells lysed using different lysis buffers – glycerol-Triton X-100 buffer, RIPA buffer, SDS-DTT buffer and urea-CHAPS buffer – were analysed using LC-MS/MS. The total number of proteins identified for each buffer was used as the primary criterion for selecting the lysis buffer. Integral membrane proteins are particularly difficult to solubilize and so, the number of membrane proteins

detected in the sample was considered a measure of how well the buffer could extract and solubilize proteins in general and membrane proteins in particular. The number of potential membrane proteins was estimated using the TMHMM algorithm which predicts transmembrane domains using a hidden Markov model [114]. The average number of proteins identified for glycerol-Triton X-100, RIPA, SDS-DTT and urea-CHAPS buffer was 485, 492, 492 and 510 respectively ($n = 2$; LC-MS/MS runs for each sample preparation). There was no statistically significant difference in the mean values of the different lysis buffers (Figure 32A). Of the 485 proteins identified in samples prepared using glycerol-Triton X-100 buffer, 62 of the proteins were predicted to have transmembrane domains. This was considerably more than the 39, 45 and 42 predicted for RIPA, SDS-DTT and urea-CHAPS buffers, respectively (Figure 32B).

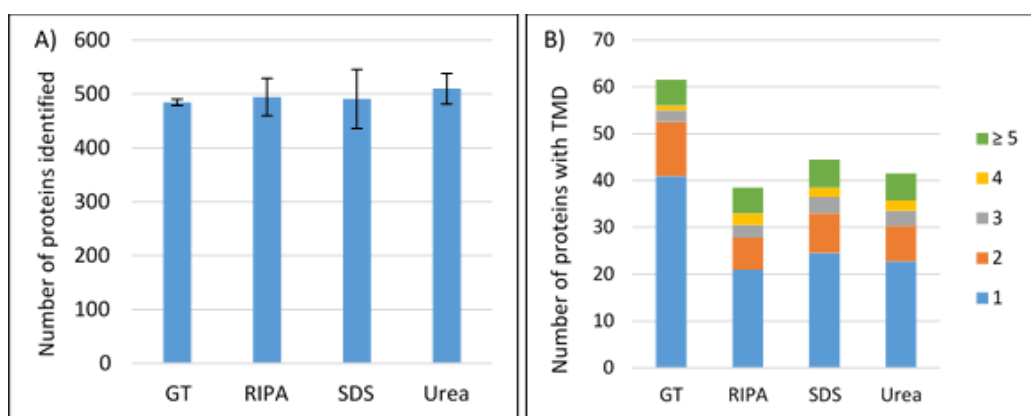


Figure 32: Comparison of lysis buffer used for sample preparation. Buffers tested include glycerol-Triton X-100 buffer (GT), RIPA buffer (RIPA), SDS-DTT buffer (SDS) and urea-CHAPS buffer (Urea). **A)** Average number of proteins identified for each buffer over two LC-MS/MS runs. **B)** Proteins with potential trans-membrane domains predicted by TMHMM (v 2.0). The number of trans-membrane domains predicted are represented by the different colours of the stacked bars.

Triton X-100 is a non-ionic detergent with a low critical micelle concentration (0.24 mM) [282]. Critical micelle concentration is a measure of surfactant efficiency – a lower concentration indicates less surfactant is needed to saturate interfaces and form lipid micelles. Triton X-100 inserts itself and rapidly equilibrates between the two leaflets of the bilayer [283]. Several holes open up as the detergent crosses the solubility threshold and the lipid bilayer then gradually vanishes. A detergent like SDS on the other hand, has a higher critical micelle concentration than Triton X-100 and tends to lyse cells by increasing the curvature of the liposomal membrane [284]. The increased curvature increases the stress on the liposomal vesicles and causes it to burst [284, 285]. Given their modes of action, Triton X-100 was considered more favourable to stability and extraction of membrane proteins.

This was reflected in the number of potential membrane proteins identified using TMHMM. The average number of proteins identified using glycerol-Triton X-100 buffer was not significantly different ($p > 0.05$) from the number identified for other buffers. However, the presence of glycerol to stabilize the proteins and the mode of action of Triton X-100 as a detergent made it suitable for solubilizing a wide range of proteins, including membrane proteins. Because of this, glycerol-Triton X-100 buffer was used for all subsequent LC-MS/MS experiments.

4.2.2. Dataset LC1

4.2.2.1. LC-MS/MS parameters

Cell pellets of four biological replicates representing untreated HepG2 cells (hereafter referred to as 'control' cells) and four additional cell pellets representing cells treated with 1 mM FA for 6 h (hereafter referred to as 'FA-treated cells') were lysed using the glycerol-Triton X-100 buffer. Four biological replicates of HepG2 cells treated with 100 μ M A922500 and 1 mM FA for 6 h (hereafter referred to as 'A92-treated' cells) were also included in this dataset. These cells were derived from the ATCC stock vial used in the initial experiments in which A92 was able to limit lipid accumulation. Consequently, exclusion of the A92 data from analysis could not have been foreseen at the time. The experimental design for dataset LC1 included three groups – control, FA-treated and A92-treated and each group was represented by one pooled sample. Data attributed to a pooled sample was gathered over five LC-MS/MS runs.

The method used to generate the LC-MS/MS data included the separation of peptides using the multi-step LC gradient (Figure 33). The gradient started with 2 % Solvent B in the mobile

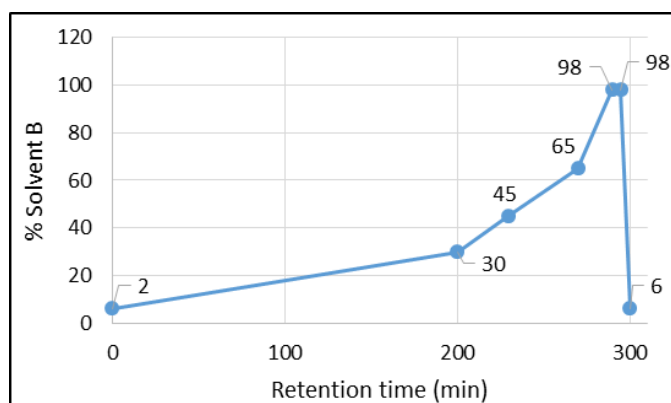


Figure 33: LC gradient used for separation of peptides analysed in dataset LC1. The peptides analysed for LC1 were separated over a period of 300 min using a multistep gradient of solvent A and B. Solvent A was 0.1 % formic acid and solvent B was 0.1 % formic acid in 80 % acetonitrile. The increase in the percentage of solvent B (% Solvent B) over time increased the hydrophobicity of the solvent mixture and so the least hydrophobic peptides eluted first and the most hydrophobic peptides eluted last.

phase and in step 1, the percent of solvent B was increased to 30 % in 200 min. In subsequent steps, the percent of Solvent B was increased to 45 % over 30 min followed by an increase to 65 % in 40 min. In step 4, the percent of Solvent B was increased from 65 to 98 % over 20 min and kept constant at 98 % for 5 min before reducing it to 6 % to start another run. The LTQ Orbitrap XL was set to perform an Nth order double play experimental protocol as described in section 2.6.3.3. The 'N' for scan event 2 was set to 6 for the identification and quantification of proteins, and the method using this setting is referred hereafter as the 'Top 6' method. The Top 6 method selected the six most intense ions of the master scan for data-dependent MS/MS fragmentation with no selection of monoisotopic precursors, no charge-state screening and a dynamic exclusion setting set at a repeat count of 2. This method accepted the most abundant isotopic peak for fragmentation without correcting its mass to the monoisotopic value. It did not determine the charge state of the ions analysed in the Orbitrap. A repeat count of 2 would have allowed the same MS precursor to be fragmented twice before being excluded from MS/MS fragmentation.

4.2.2.2. Protein identification

The LC-MS/MS data of each run were processed using the workflow of the PD wizard for SEQUEST-HT search to identify the peptides and the proteins (Figure 34). In this workflow, the input spectrum was generated by the 'Spectrum Selector' node which also determined the best precursor peak for a given MS/MS spectrum from the corresponding MS1 spectrum in the master scan. The 'Sequest HT' node matched the spectra to a peptide and used its scoring function to evaluate the quality of the match [205]. The peptide-spectra matches were further validated by the 'Target Decoy PSM Validator' node which calculated FDR for the matches. The average number of input spectra detected for the control sample was 16,471 and the average peptide-spectra matches possible was 6,139. For the FA-treated sample, an average of 17,219 spectra were detected, and they were assigned to an average of 5,916 peptide matches. An average of 17,013 spectra were detected for the five LC-



Figure 34: The pipeline tree used for identification of proteins listed in LC1. SEQUEST HT search node in this workflow generated theoretical spectra of all the proteins in the SwissProt database assuming digestion using trypsin with a maximum of 2 missed cleavage sites, precursor mass tolerance of 10 ppm and fragment mass tolerance of 0.6 Da. Modifications assumed in the search included oxidation of methionine (+15.995 Da) and presence of a carbamidomethyl group on cysteine residues (+57.021 Da). The maximum ΔC_n accepted for a peptide-spectra match was 0.05, and the results were further validated by the 'Target Decoy PSM Validator' node that performed a decoy database search to limit the FDR to 0.01.

MS/MS runs of the A92-treated sample, and the average number of peptide spectra matches were 5,990. The total numbers of peptides identified for control, FA-treated and A92-treated samples were 2,615; 2,458 and 2,553, respectively (Figure 35A). The average number of proteins identified over five LC-MS/MS runs was 408, 387 and 410 proteins for the control, FA-treated sample and A92-treated samples respectively. Each protein was identified with high-confidence ($FDR \leq 0.01$) and has at least 2 peptides per protein (Figure 35B). The proteins identified in the three groups were compared using a consensus list generated in PD for a given group. This consensus list was generated using all five LC-MS/MS runs as input and, after filtering the list for proteins identified with high confidence and at least 2 peptides attributed to it, it contained 656, 627 and 662 proteins for the control, FA-treated and A92-treated groups respectively. The consensus list represented the total number of unique proteins identified in each group tested. Of these 548 proteins were common to all three groups (Figure 35C). The rather low number of proteins identified in this dataset merited further optimization of the LC-MS/MS to increase the number of proteins. Different LC-MS/MS parameters tested to achieve this are described in section 4.2.3.

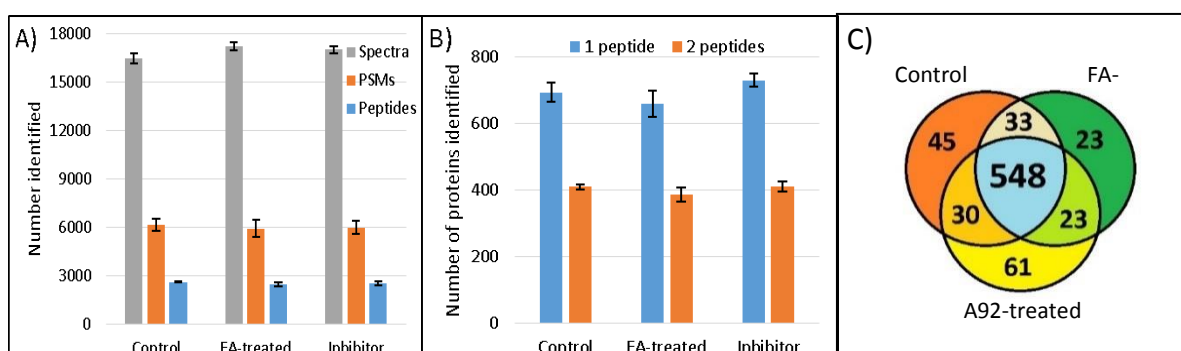


Figure 35: Proteins identification for dataset LC1 in Proteome Discoverer. The mean values (\pm SD) of these five LC-MS/MS runs performed for each pooled sample are indicated A and B. **A)** Input spectra detected, peptide-spectra matches (PSM) and the number of distinct peptides identified **B)** The average number of proteins identified with high confidence ($FDR < 0.01$) and with at least 1 or 2 peptides per protein. **C)** Extent of overlap seen among the list of proteins identified for each treatment group.

4.2.2.3. Label-free quantification

Despite the limited number of proteins identified in dataset LC1, the data from the three groups was compared in Scaffold for preliminary label-free quantification. The *in silico* experimental setup is described in section 2.6.5. Since the experimental design included pooled samples, the three groups were represented by one biosample each and each biosample was represented by five LC-MS/MS runs. Peptide identifications were accepted if they could be established at greater than 99.0 % probability by the Peptide Prophet

algorithm. Protein identifications were accepted if they could be established at greater than 99.9 % probability and contained at least 2 identified peptides. A total of 523 proteins with no decoys were identified for the three conditions (Appendix Table 3). Of the 523 proteins identified in Scaffold, 511 proteins were found in all three groups and 11 proteins were identified in at least two of the three groups (Figure 36A). The number of peptides common to all three groups was 1,832 and this corresponded to 2,243 unique spectra that were common to all three groups (Figure 36B and 36C). A comparison of the QV was performed using Fisher's exact test with multiple testing correction by the Hochberg-Benjamini method. Two categories were compared at a time, and the significance level was set at 95 %. For comparing the control and FA-treated groups, the difference in the QV was considered significant for $p \leq 1.43 \times 10^{-3}$ and 14 proteins showed statistically significant changes in abundance. The difference in QV was considered statistically significant for 21 proteins with $p \leq 2.1 \times 10^{-3}$ in the comparison of the FA-treated group to the A92-treated group. The difference in QV was statistically significant for 48 proteins with $p \leq 4.68 \times 10^{-3}$ in the comparison of A92-treated group to the control group. These proteins are listed in Table 4.

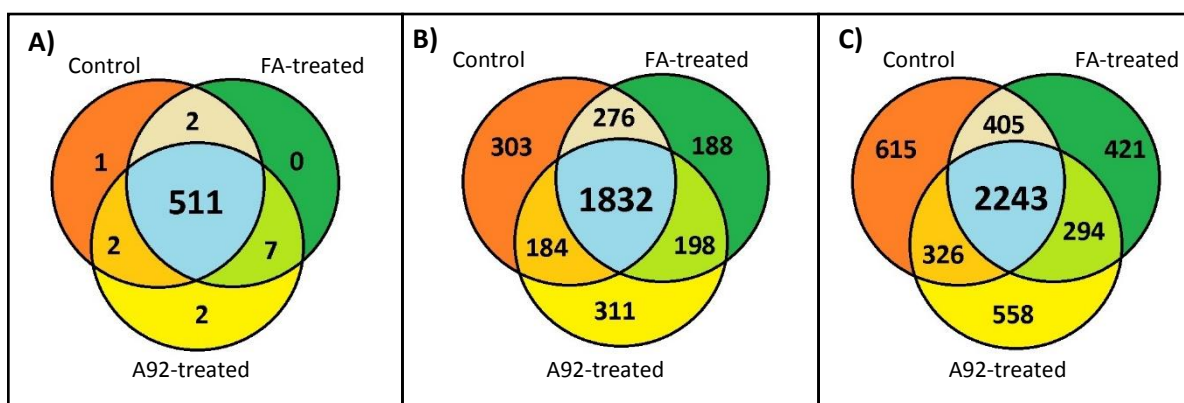


Figure 36: Identifications of proteins, peptides and MS/MS spectra in different treatment groups in Scaffold.
A) Proteins identified in each treatment group. **B)** Total unique peptide counts for all proteins in the dataset.
C) The total number of unique spectra detected for all the peptides in the dataset.

Table 4: Proteins with significant changes in abundance from dataset LC1. 'C', 'F' and 'I' represent the control, FA-treated and A92-treated samples respectively. 'C vs F' indicates the comparison of proteins in the control and the FA-treated groups. Similarly, 'FA vs I' represents comparison of the FA-treated sample to the A92-treated sample and 'C vs I' represents comparison of the control to the A92-treated sample. The QV were compared using Fisher's exact test. The 'fold change' represents the ratio of the QV for the groups as indicated. 'Unique peptides' indicates the number of peptides that are unique and exclusive to the indicated protein

| Protein name | Gene name | Fold change | | | P values determined by Fisher exact test | | | Quantitative value | | | Unique peptides | | |
|--|-----------|-------------|--------|--------|--|----------|----------|--------------------|-------|--------|-----------------|----|----|
| | | C vs F | I vs F | I vs C | C vs F | I vs F | I vs C | C | F | I | C | F | I |
| Actin, cytoplasmic 2 | ACTG1 | 1.20 | 0.8 | 1.0 | 0.0067 | 0.0014 | 0.32 | 481.5 | 567.6 | 470.9 | 20 | 20 | 20 |
| Isoform 2 of Actin-related protein 2 | ACTR2 | INF | 5.4 | INF | 0.25 | 0.021 | 0.0011 | 0.0 | 1.9 | 10.2 | 0 | 1 | 2 |
| Apoptosis-inducing factor 1, mitochondrial | AIFM1 | INF | INF | INF | 1 | 0.0042 | 0.0042 | 0.0 | 0.0 | 8.0 | 0 | 0 | 2 |
| Aldehyde dehydrogenase, mitochondrial | ALDH2 | 0.70 | 2.2 | 1.5 | 0.071 | 0.00022 | 0.025 | 46.5 | 31.9 | 69.3 | 8 | 9 | 11 |
| Fructose-bisphosphate aldolase C | ALDOC | 0.20 | 2.6 | 0.5 | < 0.0001 | 0.024 | 0.0059 | 38.1 | 7.0 | 18.2 | 2 | 2 | 1 |
| Aminopeptidase N | ANPEP | 0.90 | 4.4 | 4.0 | 0.5 | < 0.0001 | < 0.0001 | 11.0 | 9.8 | 43.5 | 3 | 5 | 7 |
| Apolipoprotein B-100 | APOB | 3.80 | 6.4 | 25 | 0.19 | < 0.0001 | < 0.0001 | 1.1 | 4.2 | 27.0 | 1 | 3 | 8 |
| Apolipoprotein E | APOE | 0.90 | 0.5 | 0.5 | 0.33 | 0.013 | 0.0028 | 41.9 | 36.4 | 19.3 | 6 | 7 | 6 |
| Isoform 2 of Transcription factor BTF3 | BTF3 | 0.60 | 0.7 | 0.4 | 0.06 | 0.17 | 0.0039 | 30.9 | 19.4 | 13.6 | 4 | 3 | 2 |
| Complement component 1 Q subcomponent-binding protein, mitochondrial | C1QBP | 1.40 | 1.1 | 1.6 | 0.015 | 0.31 | 0.0029 | 64.7 | 93.0 | 101.8 | 5 | 5 | 7 |
| Isoform 3 of Calumenin | CALU | 0.70 | 0.2 | 0.1 | 0.24 | 0.0095 | 0.00064 | 18.1 | 12.7 | 2.6 | 3 | 2 | 1 |
| Cytoplasmic dynein 1 heavy chain 1 | DYNC1H1 | 1.60 | 2.4 | 3.9 | 0.29 | 0.02 | 0.0024 | 5.0 | 8.1 | 19.4 | 3 | 2 | 7 |
| Eukaryotic initiation factor 4A-I | EIF4A1 | 0.80 | 0.8 | 0.6 | 0.067 | 0.069 | 0.0011 | 100.3 | 79.7 | 61.0 | 14 | 14 | 11 |
| Eukaryotic translation initiation factor 4B | EIF4B | 0.50 | 0.1 | 0.07 | 0.084 | 0.059 | 0.00081 | 13.0 | 5.9 | 0.9 | 3 | 1 | 1 |
| Eukaryotic translation initiation factor 4 gamma 1 | EIF4G1 | 0.10 | 22 | 2.6 | 0.035 | < 0.0001 | 0.016 | 7.4 | 0.9 | 18.9 | 3 | 1 | 3 |
| Squalene synthase | FDFT1 | 3.90 | 0.4 | 1.5 | < 0.0001 | < 0.0001 | 0.13 | 14.2 | 56.0 | 21.2 | 5 | 8 | 6 |
| Isoform 2 of Filamin-A | FLNA | 0.50 | 0.6 | 0.3 | 0.0023 | 0.062 | < 0.0001 | 56.7 | 29.8 | 18.5 | 16 | 11 | 4 |
| Glutaminase kidney isoform, mitochondrial | GLS | 1.00 | 8.3 | 8.0 | 0.69 | 0.00043 | 0.00042 | 2.2 | 2.1 | 17.3 | 2 | 1 | 3 |
| Glutamate dehydrogenase 1, mitochondrial | GLUD1 | 1.00 | 0.7 | 0.7 | 0.52 | 0.0042 | 0.0043 | 105.9 | 105.8 | 69.8 | 12 | 14 | 9 |
| Histone H2AX | H2AFX | 1.70 | 0.4 | 0.60 | 0.016 | < 0.0001 | 0.067 | 30.6 | 50.5 | 18.8 | 3 | 3 | 2 |
| Isoform 2 of Heat shock protein HSP 90-alpha | HSP90AA1 | 1.50 | 0.8 | 1.20 | < 0.0001 | 0.00062 | 0.011 | 305.2 | 463.2 | 369.0 | 18 | 17 | 13 |
| 60 kDa heat shock protein, mitochondrial | HSPD1 | 1.10 | 1.2 | 1.30 | 0.051 | < 0.0001 | < 0.0001 | 749.0 | 816.2 | 1007.7 | 36 | 36 | 33 |
| Importin subunit alpha-1 | KPNA2 | 0.40 | 0.4 | 0.10 | 0.0081 | 0.11 | < 0.0001 | 22.3 | 8.0 | 3.2 | 6 | 4 | 2 |
| Keratin, type II cytoskeletal 1 | KRT1 | 1.40 | 7.2 | 10.0 | 0.39 | < 0.0001 | < 0.0001 | 5.1 | 7.0 | 50.7 | 2 | 2 | 8 |
| Keratin, type I cytoskeletal 10 | KRT10 | 0.90 | 18 | 16.0 | 0.5 | < 0.0001 | < 0.0001 | 6.1 | 5.2 | 94.5 | 2 | 3 | 11 |
| Keratin, type II cytoskeletal 2 epidermal | KRT2 | INF | INF | INF | 1 | 0.00014 | 0.00014 | 0.0 | 0.0 | 13.3 | 0 | 0 | 2 |

Table 4 continued...

| Protein name | Gene name | Fold change | | | P values determined by Fisher exact test | | | Quantitative value | | | Unique peptides | | |
|---|-----------|-------------|--------|--------|--|----------|----------|--------------------|-------|-------|-----------------|----|----|
| | | C vs F | I vs F | I vs C | C vs F | I vs F | I vs C | C | F | I | C | F | I |
| Leucine-rich PPR motif-containing protein, mitochondrial | LRPPRC | 1.10 | 1.4 | 1.50 | 0.16 | 0.0019 | < 0.0001 | 136.6 | 155.7 | 210.8 | 18 | 19 | 26 |
| Matrin-3 | MATR3 | 0.60 | 0.6 | 0.40 | 0.014 | 0.063 | < 0.0001 | 57.7 | 35.6 | 22.2 | 7 | 4 | 3 |
| Isoform 2 of Mannosyl-oligosaccharide glucosidase | MOGS | 3.70 | 1.2 | 4.40 | 0.0095 | 0.33 | 0.0015 | 4.1 | 15.2 | 18.4 | 2 | 3 | 4 |
| Nuclear transport factor 2 | NUTF2 | 0.20 | 0.4 | 0.08 | 0.011 | 0.3 | 0.00081 | 13.1 | 2.9 | 1.0 | 3 | 2 | 1 |
| Proliferation-associated protein 2G4 | PA2G4 | 0.90 | 0.5 | 0.40 | 0.36 | 0.017 | 0.0045 | 35.0 | 31.2 | 15.6 | 7 | 5 | 4 |
| Poly(rC)-binding protein 1 | PCBP1 | 0.80 | 0.7 | 0.50 | 0.13 | 0.052 | 0.0023 | 55.1 | 43.4 | 28.2 | 9 | 7 | 5 |
| Protein disulfide-isomerase A3 | PDIA3 | 1.20 | 0.6 | 0.80 | 0.064 | < 0.0001 | 0.015 | 136.2 | 164.2 | 102.3 | 19 | 20 | 17 |
| Perilipin-2 | PLIN2 | INF | 0.05 | INF | < 0.0001 | < 0.0000 | 0.5 | 0.0 | 21.3 | 1.1 | 0 | 4 | 1 |
| DNA-dependent protein kinase catalytic subunit | PRKDC | 2.00 | 1.7 | 3.30 | 0.075 | 0.072 | 0.0011 | 8.2 | 16.1 | 26.8 | 2 | 4 | 7 |
| 26S proteasome non-ATPase regulatory subunit 2 | PSMD2 | 0.40 | 13 | 5.50 | 0.34 | < 0.0000 | 0.00032 | 4.0 | 1.7 | 21.8 | 2 | 2 | 4 |
| Ras-related protein Rab-11A | RAB11A | INF | 0.5 | INF | 0.00097 | 0.14 | 0.033 | 0.0 | 10.3 | 5.1 | 0 | 1 | 2 |
| Ribonuclease inhibitor | RNH1 | 0.30 | 4.2 | 1.30 | 0.013 | 0.0015 | 0.28 | 16.4 | 4.9 | 20.7 | 3 | 3 | 5 |
| 60S acidic ribosomal protein P2 | RPLP2 | 0.60 | 0.7 | 0.40 | 0.0012 | 0.13 | < 0.0001 | 77.2 | 43.6 | 32.5 | 6 | 6 | 6 |
| Large neutral amino acids transporter small subunit 1 | SLC7A5 | 0.00 | INF | 3.40 | 0.031 | < 0.0001 | 0.0096 | 5.0 | 0.0 | 17.1 | 2 | 0 | 2 |
| Superoxide dismutase [Cu-Zn] | SOD1 | 0.80 | 0.7 | 0.50 | 0.16 | 0.052 | 0.003 | 54.8 | 43.4 | 29.1 | 4 | 4 | 3 |
| Spectrin alpha chain, non-erythrocytic 1 | SPTAN1 | INF | 1.5 | INF | 0.00048 | 0.19 | < 0.0001 | 0.0 | 10.7 | 16.2 | 0 | 5 | 5 |
| Activated RNA polymerase II transcriptional coactivator p15 | SUB1 | 0.30 | 0.2 | 0.07 | 0.025 | 0.18 | 0.00081 | 13.2 | 3.8 | 0.9 | 2 | 2 | 1 |
| Tubulin-specific chaperone A | TBCA | 0.06 | 0.0 | 0.00 | 0.00026 | 0.49 | < 0.0001 | 15.0 | 0.9 | 0.0 | 2 | 1 | 0 |
| Isoform 2 of Tropomyosin alpha-3 chain | TPM3 | 0.50 | 0.6 | 0.30 | 0.00025 | 0.036 | < 0.0001 | 66.2 | 31.4 | 17.4 | 8 | 5 | 5 |
| Tropomyosin alpha-4 chain | TPM4 | 0.09 | 2.0 | 0.20 | < 0.0001 | 0.2 | < 0.0001 | 43.4 | 4.0 | 8.1 | 2 | 0 | 2 |
| Heat shock protein 75 kDa, mitochondrial | TRAP1 | 8.30 | 0.9 | 7.30 | < 0.0001 | 0.29 | < 0.0001 | 5.1 | 42.6 | 37.4 | 1 | 3 | 3 |
| Tubulin beta chain | TUBB | 0.90 | 0.9 | 0.80 | 0.14 | 0.022 | 0.00081 | 416.7 | 384.8 | 329.5 | 24 | 20 | 19 |
| Tubulin beta-4B chain | TUBB4B | 0.90 | 0.8 | 0.80 | 0.12 | 0.025 | 0.00071 | 333.2 | 302.8 | 255.1 | 4 | 3 | 2 |
| Voltage-dependent anion-selective channel protein 2 | VDAC2 | 1.20 | 4.8 | 5.80 | 0.5 | 0.00055 | 0.00019 | 4.0 | 4.9 | 23.2 | 1 | 1 | 5 |
| Nuclease-sensitive element-binding protein 1 | YBX1 | 0.30 | 1.9 | 0.60 | < 0.0001 | 0.032 | 0.0059 | 53.2 | 16.4 | 30.8 | 8 | 6 | 4 |

4.2.2.4. Validation of proteins detected

Six proteins were selected for validation using Western blot (Table 5). These targets were selected on the basis of their biological functions and to corroborate the statistical analysis of the dataset. Perilipin 2 and mitochondrial 75 kDa heat shock protein (TRAP1) showed statistically significant changes in abundance. Perilipin 2 was not detected in the control sample but 4 peptides were assigned to the protein in the FA-treated sample and 1 peptide was assigned to it in the A92-treated sample. Perilipin 2 is a marker for LD and given its absence in the control sample was considered a good target for validation. TRAP1 was identified in all samples and its abundance in the FA-treated and A92-treated samples was 8.3 and 7.3 times that in control. This increase was statistically significant. TRAP1 is a mitochondrial chaperone.

Glutathione reductase plays a vital role in maintaining the redox state of the cells and given the role of ROS in disease progression as described in literature, it was considered for validation [286]. Its abundance in the FA-treated sample was 5 times lower than that of control and its abundance in A92-treated sample was 4.4 times that of the FA-treated. Both changes in abundances were assigned $p < 0.05$ but were not statistically significant after multiple testing corrections. Glutaredoxin 3 is a cytoplasmic protein that also plays an important role in maintenance the redox state of the cell [287]. Fatty acid synthase (FAS) catalyses the *de novo* synthesis of FA in the cell and inorganic pyrophosphatase 1 (PPA1) hydrolyses pyrophosphate to inorganic phosphate that is used in biosynthesis of nucleic acids and proteins [288]. FAS showed almost no change in abundance across the three samples while the changes in PPA1 were assigned $p > 0.05$. The abundance of DGAT1 was also tested since A92, a DGAT1 inhibitor was part of the analysis.

Table 5: Target proteins from dataset LC1 for Western blot validation. Fold change is the ratio of QV of groups as indicated. So, for example, 'F: C' is the ratio of QV of the protein in the FA-treated sample (F) to that in control sample (C). 'I' indicates A92-treated sample.

| Protein name | Gene name | M.W (kDa) | Fold Change | | | P value | | |
|--|-----------|-----------|-------------|-----|------|----------|----------|----------|
| | | | F:C | I:F | I:C | F:C | I:F | I:C |
| Inorganic pyrophosphatase | PPA1 | 33 | 0.71 | 1.6 | 1.1 | 0.23 | 0.15 | 0.45 |
| Fatty acid synthase | FASN | 273 | 1 | 0.9 | 0.9 | 0.47 | 0.098 | 0.08 |
| Perilipin-2 | PLIN2 | 48 | INF* | -20 | INF* | < 0.0001 | < 0.0001 | 0.5 |
| 75 kDa Heat shock protein | TRAP1 | 74 | 8.3 | 0.9 | 7.3 | < 0.0001 | 0.29 | < 0.0001 |
| Glutathione reductase | GSR | 53 | 0.2 | 4.4 | 0.9 | 0.019 | 0.035 | 0.48 |
| Glutaredoxin-3 | GLRX3 | 37 | 5.2 | 0.9 | 4.9 | 0.011 | 0.48 | 0.021 |
| * Perilipin 2 was not detected in control samples; the QV for this protein in the FA-treated and A92-treated sample was 21.1 and 1.1 respectively. Its fold change is shown as 'INF' since the QV for the protein in the control sample is zero. | | | | | | | | |

None of target proteins showed a statistically significant changes in abundance in the Western blots (Figure 37). Since the difference in the QV for FAS, PPA1, glutathione reductase and glutaredoxin 3 were not statistically significant, the lack of change observed in the Western blots supported the statistical test. However, the Western blots for proteins with statistically significant changes – perilipin 2 and TRAP1 – did not show any significant changes in abundance either. In fact, perilipin 2 was detected in the control samples used for Western blots in spite of no peptides being assigned to it in the LC-MS/MS data. Actin was detected in the LC-MS/MS data with the same number of peptides assigned to all three samples; the difference in the QV of the three samples however, was considered statistically significant.

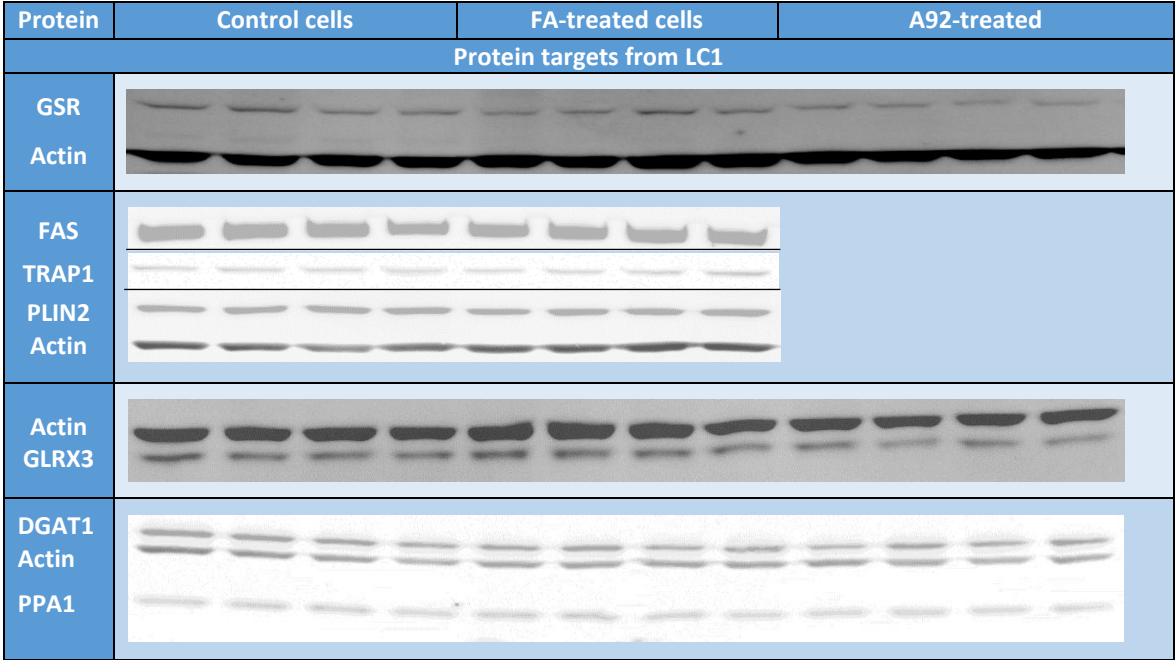


Figure 37: Western blots for the validation of dataset LC1. The proteins validated through Western blots for dataset LC1 included glutathione reductase (GSR), actin, fatty acid synthase (FAS), TRAP1, perilipin 2 (PLIN2), glutaredoxin 3 (GLRX3) and DGAT1. Four biological replicates of each condition were tested- first four lanes represent the control replicates, middle four represent FA-treated replicates and last four represent A92-treated replicates. Some images have been cropped for want of space and indicated by a line separating the target proteins. (Appendix figure 71 shows uncropped images.) * The blot for FAS, TRAP1 and PLIN2 did not include the A92-treated sample due to unavailability of sample.

These Western blot results validate the presence of proteins identified in the samples by LC-MS/MS but fail to corroborate the changes in abundances suggested by the LC-MS/MS data. Validation of LC-MS/MS targets is discussed further detail in section 5.2.4. The outcome of this analysis also suggested that greater robustness of the experimental design and an increase the number of proteins identified was needed. The next dataset generated included individual biological replicates instead of the pooled sample used in this dataset.

4.2.3. Improving protein identification

4.2.3.1. Comparison of LC-MS/MS parameters

Two LC gradients and various data-dependent MS/MS settings were tested to improve the number of proteins identified. The effectiveness of changing the LC-MS/MS parameters was determined on the basis of the number of input spectra accepted by the ‘Spectrum Selector’ node of the PD workflow was used as a measure of the number of the spectra generated using the parameter tested. The quality of the spectra was gauged by the number of peptides these spectra could be matched to. Since multiple spectra can be assigned to the same peptide, the number of distinct peptides identified on the basis of the data was used a second measure of quality. Biological relevance of the peptides requires inferences to be made at the protein level and so, the peptides identified were assigned to their respective proteins and the total number of protein identified was the third measure used to gauge the quality of the spectral information collected.

4.2.3.1.1. Optimization of LC gradient

The LC gradient used for dataset LC1 was generated using a step-wise increase in solvent B from 2 to 98 over 300 min. For all samples tested, the maximum number of peaks were detected within 200 min of this gradient run (Figure 38 shows a representative chromatogram). This corresponded to an increase of solvent B from 2 % to 30 % of mobile phase. Given the limited number of proteins identified in dataset LC1, the total run-time of the LC was increased to 352 min; this gradient is hereafter referred to as ‘Gradient 1’ (Table 6). Four LC-MS/MS runs were performed for a sample of untreated HepG2 cells using Gradient 1 and a modified Top 6 method described in section 4.2.3.2.

The chromatograms of this sample showed a similar pattern to that observed in dataset LC1 chromatograms – maximum number of peaks (approximately 81%) were detected in a step that corresponded to an increase of solvent B from 2 % to 30 % (Figure 39 shows a representative chromatogram). Given this elution pattern, an attempt was made to spread the peaks more evenly over the run. To this end, the data generated from the four LC-MS/MS runs using Gradient 1 was analysed using GOAT[®]. GOAT[®] is a gradient optimization tool that suggested a multi-step gradient such that all steps would contain an equal fraction of the MS/MS spectra acquired in the LC-MS/MS runs performed using Gradient 1 [241]. The gradient suggested by GOAT[®] is hereafter referred to as Gradient 2.

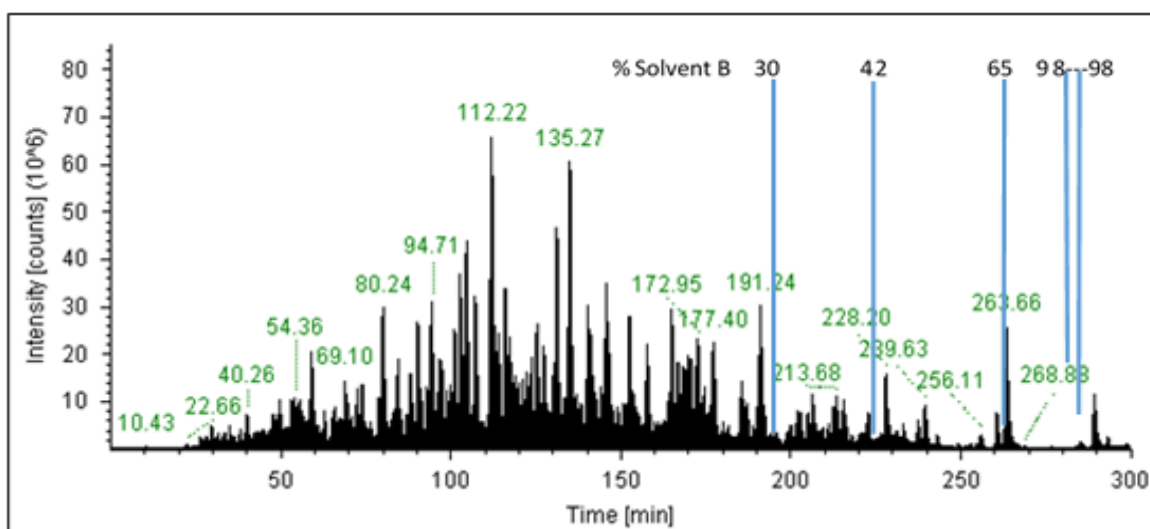


Figure 38: Representative chromatogram for Top 6 method using standard LC gradient. The figure shows the peaks detected over the 300 min LC-MS/MS run. The multi-step gradient set-up is indicated on the timeline with the corresponding percent of Solvent B in the mobile phase.

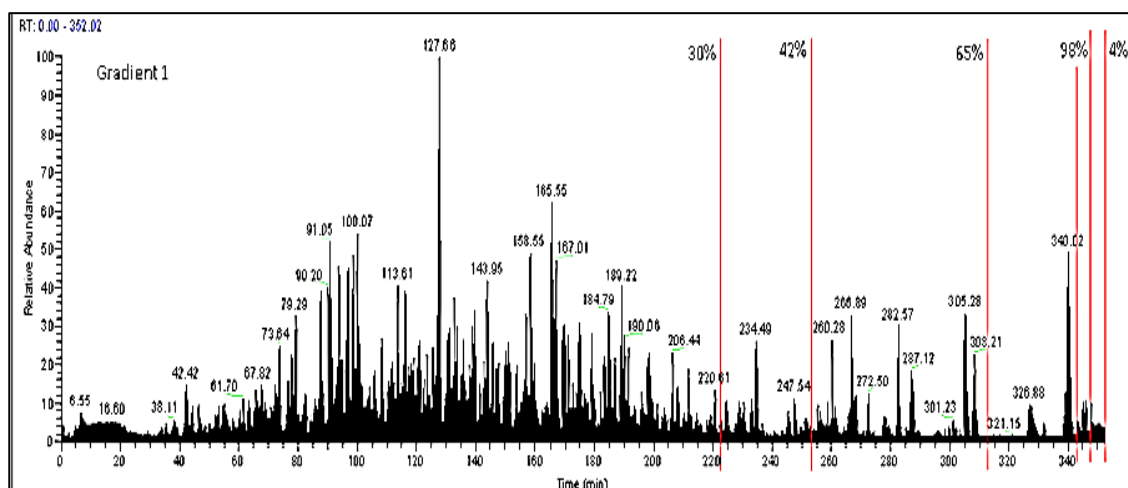


Figure 39: Representative chromatogram for the modified Top 6 method using gradient 1. The chromatogram shows the elution pattern of analytes over a period of 352 min. The % solvent B at the end-points of each step in the multi-step gradient are indicated along the red lines.

Table 6: Comparison of LC-MS/MS gradients. The gradient used to separate peptides analysed in dataset LC1 were separated using gradient labelled 'LC1 gradient'. Gradient 1 used a mobile phase with the same percent of Solvent B at the end of each step as the LC1 gradient; the difference between the two gradients was that some steps in Gradient 1 were longer than that the corresponding one in LC1 gradient.

| % Solvent B | | 2 | 30 | 45 | 65 | 98 | 98 | 6 |
|----------------|--------------|---|-----|-----|-----|-----|-----|-----|
| Retention time | LC1 Gradient | 0 | 200 | 230 | 270 | 290 | 295 | 300 |
| | Gradient 1 | 0 | 222 | 252 | 312 | 342 | 347 | 352 |

The mean value of % Solvent B suggested by GOAT[®] for the four replicates at each time point was considered as the final value used to define Gradient 2 (Figure 40A). Gradient 2 was run for the same duration as Gradient 1 (352 min) and the MS/MS was performed using the same modified Top 6 method that was used to test Gradient 1. Unlike the LC1 gradient and Gradient 1, Gradient 2 computed by GOAT had a maximum 43 % Solvent B in its mobile phase at 347 min of the run (Figure 40B). The performance of the two gradients was compared on the basis of total number of proteins identified. Protein identification workflow Combined 1 described in section 2.6.4 was used to generate a list of proteins based on the LC-MS/MS data.

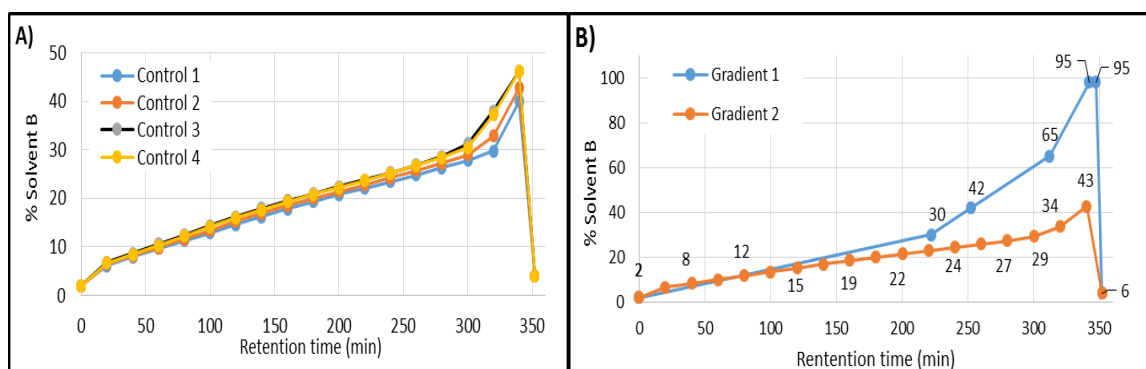


Figure 40: Optimization of LC gradient using GOAT. A) Four LC-MS/MS runs representing control samples were analysed using GOAT[®]. B) The average value of % solvent B at a given time point was used to construct Gradient 2.

Changing the gradient did produce a statistically significant increase ($p \leq 0.05$) in the average input spectra – an average of 17,267 input spectra were detected for Gradient 1 and the average for Gradient 2 was 19,432. The number of spectra matched to peptides also increased significantly from 14,484 for Gradient 1 to 16,992 for Gradient 2 (Figure 41A). This increase could be attributed to the better distribution of analytes over the run time as GOAT[®] was intended to do and would, as such, fulfil the objective of using GOAT[®]. However, because this distribution was done on the basis of the number of MS/MS inputs alone, the increased number of input spectra did not translate into an increase in the number of proteins identified. In fact, the total number of peptides and proteins identified using Gradient 2 was less than that for Gradient 1 (Figure 41B). This reduction could be in part because the highest % Solvent B in Gradient 2 was 43 %. This could have excluded the analytes that would have eluted at higher percentages used in Gradient 1. Retaining the elution pattern of Gradient 2 and increasing the run time to include higher % Solvent B in the mobile phase like in Gradient 1 would have been the ideal choice to hold onto the

advantages of both gradients tested but given the number of replicates, an increase in run time would increase the total machine time required significantly. And so, all subsequent experiments were carried out using Gradient 1 for separation of analytes since it was able to produce more protein identifications in the same run time as Gradient 2.

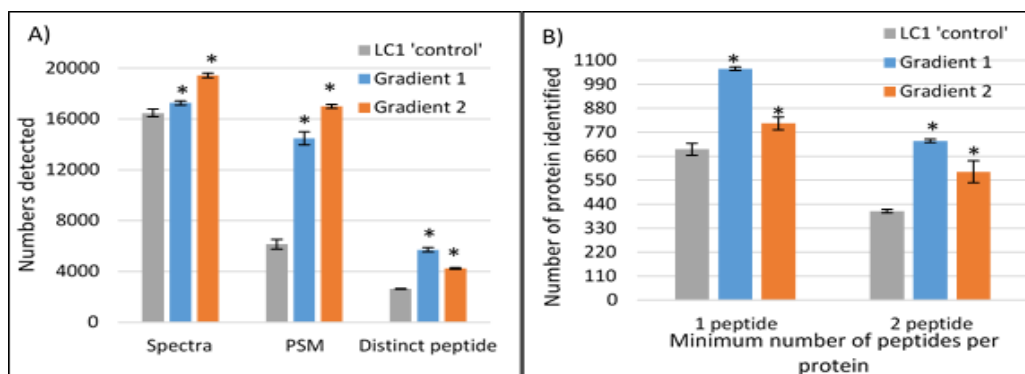


Figure 41: Comparison of gradients used for Top 6 methods. The gradients were compared on the basis of **A)** Average input spectra; the average number peptide-spectra matches (PSM), the average number of distinct peptides and **B)** number of proteins identified with high confidence (FDR < 0.01). The difference in the mean values of each selection criteria was tested using one-way ANOVA and was considered significantly different for $p < 0.05$ (*).

4.2.3.1.2. Parameters for data-dependent MS/MS fragmentation

The data-dependent MS/MS settings of Top 6 method used for the LC-MS/MS analysis of the Gradient 1 runs differed from those of the Top 6 method used for generation of dataset LC1. Unlike the Top 6 method described in section 4.2.2, the modified Top 6 method used to test Gradient 1 and Gradient 2 allowed for charge-mass screening and selection of monoisotopic precursors. The repeat count in the dynamic exclusion settings was set to 1; this meant that a MS precursor would be fragmented only once before being excluded from fragmentation. The number of input spectra detected for dynamic exclusion after a repeat count of 2 was not affected by the selection of the monoisotopic precursor; however, the number of peptides that could be matched to the spectra and the number of peptides and protein identified dropped significantly when monoisotopic peak selection was disabled (Figure 42). Therefore, selection of monoisotopic peaks was enabled for all methods tested. Lowering the repeat count to 1 did increase the number of input spectra detected but the number of peptides and proteins identified was comparable to the repeat count of 2. Since there was no advantage in performing two MS/MS fragmentation scans per precursor before excluding it, dynamic exclusion after a repeat count of 1 was used for the subsequent methods.

Increasing the number of MS precursors selected for MS/MS fragmentations was considered another way of improving the quality of the data and the number of proteins identified. This

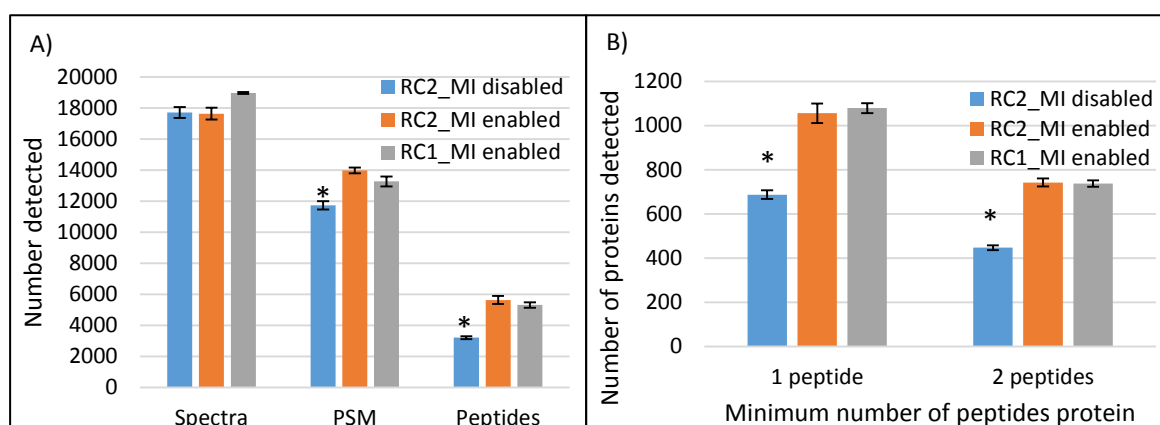


Figure 42: Comparison of data-dependent MS/MS settings. Data was collected over four LC-MS/MS runs and the figure represents the mean values (\pm SD) of the four runs. **A)** Input spectra, peptide-spectra matches (PSM) and number of distinct peptides detected for the three methods. **B)** Number of proteins identified for each with method with high confidence (FDR < 0.01) and at least 1 or 2 peptides per proteins in PD. The mean values were compared using one-way ANOVA and the difference was considered significantly for $p < 0.05$ (*). Legend: RC2_MI disabled: Repeat count 2, monoisotopic precursor selection disabled; RC2_MI enabled: Repeat count 2, monoisotopic precursor selection enabled; RC1_MI enabled: Repeat count 1, monoisotopic precursor selection enabled.

led to modifying the method to select 8 or 10 scan 1 precursors instead of the 6 used for the Top 6 method. These methods are referred to as 'Top 8' and 'Top 10' methods respectively. Settings tested for the modified Top 6 method were retained and the only difference between the methods was the number of precursor ions selected for MS/MS fragmentation. All references to the Top 6 method hereafter refer to the modified Top 6 method. Four LC-MS/MS runs were performed for each method and the average input spectra for Top 6, Top 8 and Top 10 methods was 17,267; 17,665 and 17,358 respectively. The number of peptide-spectra matches was 14,484; 11,193 and 9,835 for Top 6, Top 8 and Top 10 methods respectively and these corresponded to 5,703; 5,140 and 3,931 distinct peptides for their respective methods. So, in spite of Top 8 and Top 10 methods being able to generate more input spectra than Top 6, the number of peptide-spectra matches and the number of distinct peptides identified for Top 6 method were significantly more than that of the other two methods (Figure 43A).

And, although the number of proteins identified using Top 8 method was found to be significantly more than that found in the Top 6 method using the criteria of at least 1 peptide per protein, the difference between the two methods was not significant if the selection criteria is made more stringent to include only those proteins identified with at least 2 peptides per protein (Figure 43B). The average number of proteins identified with high confidence (FDR < 0.01) and at least 2 peptides per protein was 731 (\pm 8), 755 (\pm 21) and 609 (\pm 4) for Top 6, Top 8 and Top 10 respectively. Assuming all peptides identified were

assigned to a protein or protein group, the 731 protein identified by Top 6 method would be assigned 5,703 peptides among them while the 755 proteins identified by Top 8 method would be assigned 5,140 peptides. These peptides in turn reflect the number of spectra that could be assigned to a protein. The Top 6 method had more peptide-spectra matches than the other two methods. Based on this information, it was inferred that the protein identified by the Top 6 method are likely to be represented by more spectral counts than the same protein detected by the other two methods. Given that label-free spectral counting was the quantification method to be used for this project, the Top 6 method used for subsequent experiments.

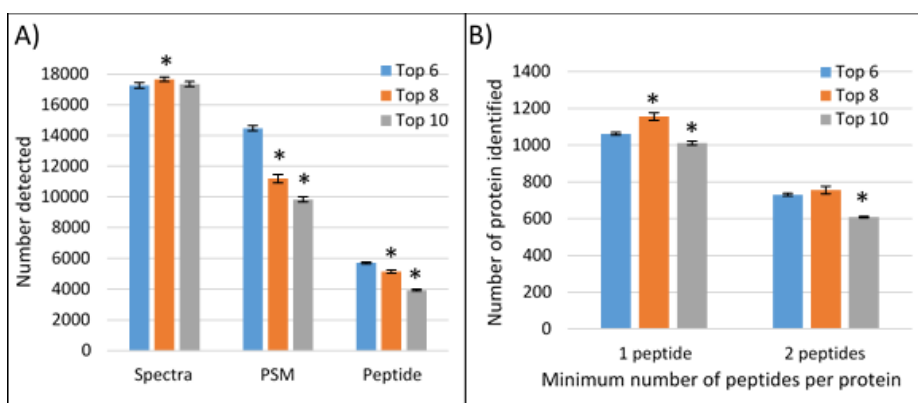


Figure 43: Comparison of Top N methods. The methods ‘Top 6’, ‘Top 8’ and ‘Top 10’ were named according to the number of MS precursors selected for MS/MS fragmentation. Data was collected over four LC-MS/MS runs and the figure represents the mean values (\pm SD) of the four runs. **A)** Input spectra, peptide-spectra matches (PSM) and number of distinct peptides detected for the three methods. **B)** Number of proteins identified for each with method with high confidence (FDR < 0.01) and at least 1 or 2 peptides per proteins. The mean values of Top 8 and Top 10 were compared to that of the Top 6 methods using one-way ANOVA and the difference was considered significantly for $p < 0.05$ (*).

4.2.3.2. Parameters for protein identification

Having defined the parameters of the LC-MS/MS method, the data collected was put through various PD workflows to ensure the search parameters used to identify the proteins would allow the identification of maximum number of proteins with high confidence. The MS/MS data generated was cumulated in a file of ‘.raw’ format and each file represented a technical replicate. Every workflow in PD requires the ‘Spectrum files’ node to accept the LC-MS/MS data. This node is followed by the ‘Spectrum Selector’. The ‘Sequest HT’ and ‘Mascot’ nodes were used to match the spectra to possible peptides and the output of these search nodes was validated using the ‘Percolator’ node. The node settings and the different workflows they were arranged in for this project has been described in section 2.6.4. An

average of 18,959 (± 62.8) input spectra were detected for the four LC-MS/MS runs analysed. Since the input spectra used for each workflow was the same, any difference in the number and identity of proteins was attributed to the workflow and the manner in which the data was processed by it. This section describes the comparison of different workflows and databases used to select a combination that would identify the highest number of protein.

4.2.3.2.1. Comparison of Proteome Discoverer workflows

The combination of SEQUEST HT and MASCOT was able to generate a significantly higher number of peptides-spectra matches than either of the two search engines alone. The 'SEQUEST only' workflow was able to produce 6,848 peptide spectra matches while; 'MASCOT only' workflow produced 6,418 matches. 'Combined 1' and 'Combined 2' on the other hand were able to generate 13,266 and 13,291 peptide-spectra matches, respectively (Figure 44A). In spite of this, the average numbers of peptides and proteins identified by each workflow were comparable – 5,120; 4,861; 5,309 and 5,320 peptides and 724, 698, 737 and 737 protein with high confidence (FDR < 0.01) and at least 2 peptides were identified using 'SEQUEST only', 'MASCOT only', 'Combined 1' and 'Combined 2' respectively (Figure 44B).

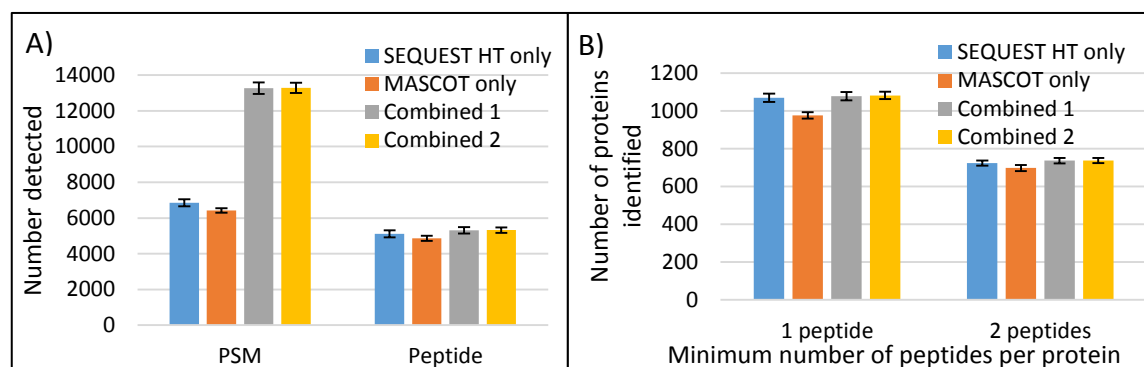


Figure 44: Comparison of number of proteins identified by different PD workflows. Data was collected over four LC-MS/MS runs and the figure represents the mean values (\pm SD) of the four runs with an average input spectra of 18,959 (± 62.8). **A)** Peptide-spectra matches (PSM) and number of distinct peptides detected for the four workflows. **B)** Number of proteins identified for each with method with high confidence (FDR < 0.01) and at least 1 or 2 peptides per proteins

A consensus list of proteins was generated to represent all proteins identified by a workflow across all four technical replicates. A comparison of this consensus list using Venny 2.1 (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>) showed that most of proteins identified were common to all the workflows tested [277] (Figure 45). A distinction was

made between the 'Combined' workflows and the workflows with either SEQUEST HT or MASCOT node on the basis of the number of peptide-spectra matches generated by the workflow. The 'MS2 – Spectrum Processor' node used in workflow 'Combined 2' was included in the workflow to perform charge state deconvolution and deisotoping of the data with the aim of improving the number of proteins identified. Since the peptide-spectra matches and the number of proteins identified for the Combined workflows were not significantly different, the inclusion of this node was not deemed necessary and workflow Combined 1, which had the same nodes as Combined 2, barring the 'MS2 –Spectrum Processor', was used for all subsequent experiments.

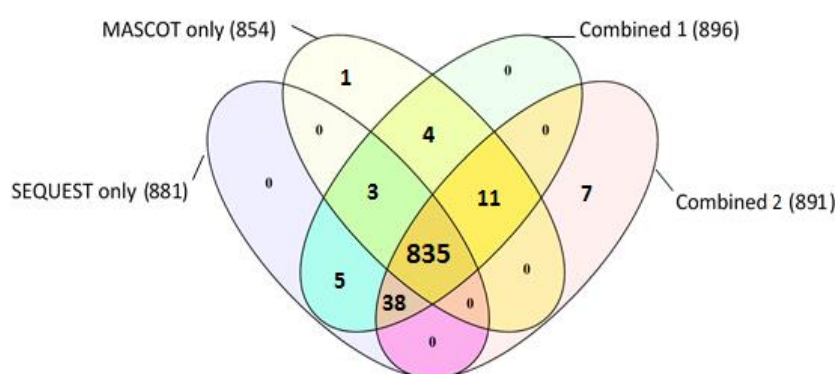


Figure 45: Comparison of proteins identified by the different PD workflows. A multi-file report was generated in PD to list all the proteins identified in all the four LC-MS/MS runs tested using a given PD workflow. The proteins identified using a given workflow were compared to the proteins identified by other workflows using the Venn diagram tool Venny 2.1.

4.2.3.2.2. Comparison of protein databases

UniProt is a comprehensive resource for protein sequence and annotation [115]. At the time of analysis, the UniProtKB database had 148,986 entries for human proteins and of these, 20,196 were manually annotated and reviewed as SwissProt entries. Evidence of existence was available for 51,543 human proteins in UniProtKB, and of these, 14,685 were manually annotated and reviewed as SwissProt entries. The 'SEQUEST HT' and 'Mascot' search nodes in 'Combined 1' workflow were set-up to search these databases. The average number of proteins identified and the individual proteins identified in each database were about the same for all databases tested. An average of 737, 735, 734 and 733 proteins were identified with high confidence ($FDR < 0.01$) from SwissProt (with evidence), all SwissProt, UniProt (with evidence) and all UniProt with at least 2 peptides attributed to each protein (Figure 46A). A comparison of proteins identified in the consensus lists (derived from the PD multi-

file report that cumulates data from all four LC-MS/MS runs into one report for each database tested) showed that 823 proteins were common to all four database searches (Figure 46B). Given these results, the SwissProt database with evidence of protein existence was selected for the quantification experiments because of its entries being reviewed and verified by experimental evidence.

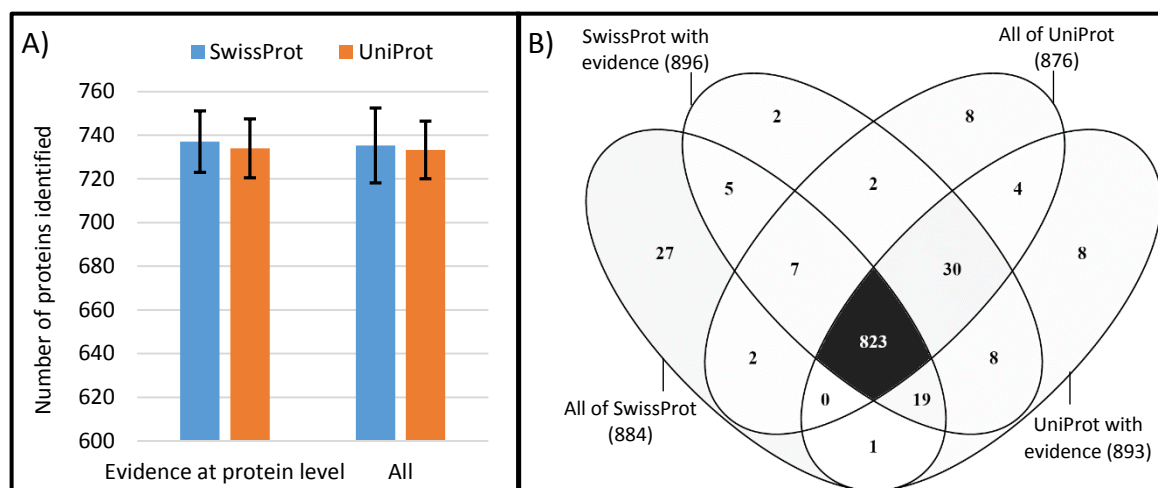


Figure 46: Comparison of protein databases used for protein identification. A) The number of proteins identified using the different databases. **B)** Extent of overlap between the protein lists generated by the different databases. Protein identification data from four technical replicates was accumulated into one multi-consensus report to give a proteins list representative of the sample. The consensus lists of proteins were compared using Venny 2.1.

4.2.4. Dataset LC2

4.2.4.1. Protein identification and quantification

Having tested the different parameters that could increase the number of proteins identified, a new batch of HepG2 cells was treated with 1 mM FA for 6 h for preparation of LC-MS/MS samples. These cells originated in the stock vial obtained from the Malaghan Institute. The DGAT inhibitor A92 was also tested in cell culture as described in section 3.2.2.1. but its effect on lipid accumulation was not reproducible. As a result, it was not included in the proteomics analysis performed to generate dataset LC2. The commercial unavailability of other DGAT inhibitors at the time restricted the number of treatment groups to that of control cells and FA-treated cells. The LC-MS/MS runs for these samples were performed using Gradient 1 for separation of peptides (described in section 4.2.3.1.) and the modified Top 6 method (described in section 4.2.3.2.). Dataset LC2 was generated using four biological replicates of untreated control cells and another four biological replicates of FA-treated cells. Unlike dataset LC1, the samples were not pooled, and four LC-

MS/MS runs were performed for each replicate. The average number of proteins identified for each biological replicate and for each condition in general is described in Figure 47.

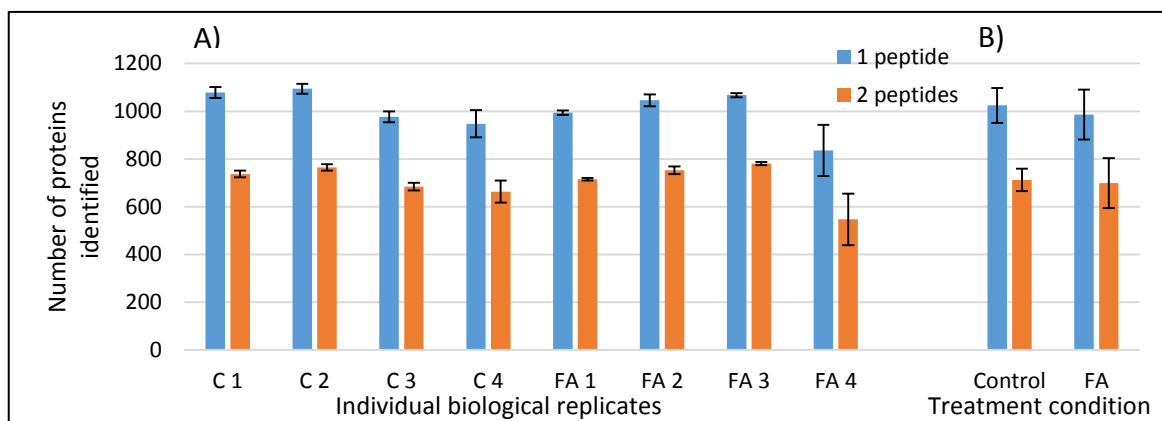


Figure 47: Number of proteins identified for each biological replicate of dataset LC2 in PD. The figure represents the average number of proteins identified. **A)** Mean values (\pm SD) for four LC-MS/MS runs of each biological replicate. **B)** Mean value for all the 16 LC-MS/MS runs that represent the treatment condition. *Legend: 1 peptide and 2 peptides represent the protein identified with high confidence ($FDR < 0.01$) and least 1 or 2 peptides per proteins attributed to it in PD; #Scaffold is the number of proteins identified in Scaffold with $p < 0.001$ and contained at least two peptides identified with $p < 0.01$. Key: C: Control samples; FA: Sample treated with 1 mM FA for 6 h.*

The *in silico* experimental setup for dataset LC2 in Scaffold also differed from dataset LC1. Each category had four biosamples, and each biosample was analysed by four LC-MS/MS runs. The data from the four runs were condensed to a single value for each protein in the biosample. Peptide identifications were accepted if they could be established at greater than 99.0 % probability by the Peptide Prophet algorithm. Protein identifications were accepted if they could be established at greater than 99.9 % probability and contained at least 2 identified peptides. With these identification criteria, a total of 998 proteins and no decoys were identified for dataset LC2 (Table 21 in Appendix). An average of 846 proteins could be identified for the four control group biosamples, and an average of 833 proteins were identified for the four FA-treated group biosamples. The statistical difference between the mean QV for the control group was compared to that of the FA-treated group for each protein using a *t*-test with a Hochberg-Benjamini correction carried out. The significance level was set at 95 %. Although there were 94 proteins that showed a fold change of ≥ 1.3 with a $p \leq 0.05$ (Table 7), none of these were statistically significant after the multiple testing correction that required $p \leq 5 \times 10^{-4}$ to achieve the significance level of 95 %.

4.2.4.2. Comparison of datasets

The proteins identified in Scaffold for dataset LC2 were compared to the proteins identified in Scaffold for dataset LC1, and almost all the proteins identified in LC1 were also detected in

LC2 (Figure 48A). However, the same overlap was not seen in the proteins that showed changes in abundance with $p \leq 0.05$ (Figure 48 B). The lists of proteins identified for LC1 and LC2 were filtered to retain proteins that showed a change in abundance with $p \leq 0.05$ for the comparison between control and FA-treated samples. Only 7 proteins were common to both lists. These proteins were L-lactate dehydrogenase A chain (LDH-A), thioredoxin reductase 1 (TR), 14-3-3 protein, perilipin-2, proteasome activator complex subunit 3 (PSME3), activated RNA polymerase II transcriptional coactivator p15 (PC4), and eukaryotic translation initiation factor 4 gamma 1. The functions of these proteins varies greatly – LDH-A carries out the reversible conversion of lactate to pyruvate using $\text{NADH} + \text{H}^+$ as the proton donor [289] while the glutaredoxin activity as well as thioredoxin reductase activity of TR contribute towards the maintenance of the redox environment of the cell [290]. Perilipin 2 is a lipid-droplet marker [291]. PC4 is capable of binding single-stranded DNA and mediating activator-dependent transcription [292] while PSME3 is part of the proteosomal regulator complex that activates the catalytic subunits of the proteasome [293]. Eukaryotic translation initiation factor 4 gamma 1 as the name suggests is part of the translation initiation complex while the 14-3-3 protein has been identified as regulatory element in intracellular signalling pathways [294]. The change in the abundance of these proteins was not consistent between the two datasets. For example, the abundance of LDH-A in FA-treated sample of dataset LC1 was 1.4 times of control while in dataset LC2, its abundance was down to 0.7 times that of control. Moreover, perilipin 2 was exclusive to the FA-treated sample in dataset LC1 while it was detected in both control and FA-treated samples of datasets LC2.

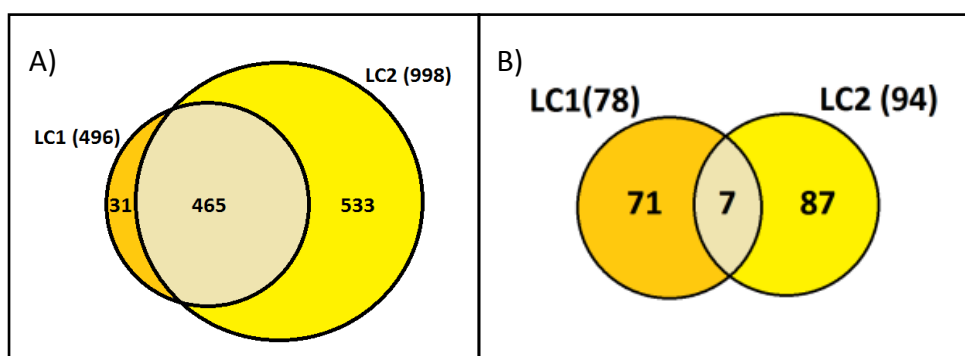


Figure 48: Comparison of datasets LC1 and LC2. A) The list of proteins identified in dataset LC1 was compared to the proteins identified in LC2 using Venny 2.1. B) The proteins that showed changes in abundance with $p < 0.05$ did not show the same extent of overlap as the list of total proteins.

The limited number of overlapping proteins (with change in abundance with $p < 0.05$) and the disparity in the extent of change seen in these proteins may be attributed to the experimental design and perhaps the change of cell culture stock. The biological variance in

dataset LC1 may be considered as being masked by the pooling of cell lysates into one sample. The experimental design of dataset LC2 included the use of unpooled samples which would reflect the biological variance among the four biological replicates. The biological variance and the inherent technical variations introduced during LC-MS/MS data generation are likely to have had an adverse effect on the statistical significance assigned to proteins in dataset LC2. The co-efficient of variance (CV) associated with a protein for each category is presented in Table 7. A direct comparison of datasets LC1 and LC2 is made difficult by the fact that they were generated from cell culture samples that originate in stock vials sourced from different institutes. This is evident by the lack of effect of DGAT inhibitor A92 on cells used for dataset LC2. A critical evaluation of the experimental design would require a comparison of individual replicates and a pooled sample originating from the same biological source.

4.3. Summary

HepG2 cells were treated with 1 mM FA mixture for 6 h and changes in protein abundance as a result of the treatment were recorded. Cells treated with the 100 μ M A92 and 1 mM FA mixture for 6 h were also included in the analysis. Of the four lysis buffers tested, the glycerol-Triton X-100 buffer was selected for preparing samples for proteomics because of the mild non-ionic nature of Triton X-100 and its ability to extract more membrane proteins than the other buffers tested. Various parameters of LC-MS/MS and protein identification were compared and modified to determine a method that would give the highest number of proteins identified in the dataset. Dataset LC1 was generated from a pooled sample and although proteins relevant to fat metabolism were identified, none of the proteins could be validated by Western blot. Dataset LC2 was then generated from unpooled, individual replicates to improve the robustness of the experimental design. A92-treated cells were not included in this dataset since the previous effects of the molecule on lipid accumulation could not be reproduced. No statistically significant protein changes could be identified in dataset LC2, possibly because of the additional variance introduced in the analysis by using individual biological samples rather than a pooled sample. To ascertain the impact of pooling samples, a new batch of HepG2 cells was prepared and treated with 1 mM FA. The biological replicates produced in this preparation were analysed in datasets LC3 and LC4 and are described in the next chapter.

Table 7: List of proteins identified in dataset LC2 with p < 0.05. The ‘fold change’ represents the ratio of the quantitative value for FA-treated group to that of the control group. ‘Quantitative value’ is the normalized spectral count for the protein indicated. ‘CV (%)’ is the coefficient of variance expressed as a percentage for the four biological replicates in each category.

| Protein name | Gene name | FC | P value | Quantitative value | | | | | | | | Exclusive peptides | | | | | | | | CV (%) | |
|---|-----------|-----|---------|--------------------|--------|--------|--------|--------|--------|--------|-------|--------------------|----|----|----|----|----|----|----|--------|----|
| | | | | C1 | C2 | C3 | C4 | F1 | F2 | F3 | F4 | C1 | C2 | C3 | C4 | F1 | F2 | F3 | F4 | C | F |
| Oxygen-dependent coproporphyrinogen-III oxidase | CPOX | 1.5 | 0.00033 | 11.48 | 9.57 | 9.34 | 10.32 | 15.53 | 14.47 | 15.36 | 17.19 | 4 | 4 | 3 | 4 | 6 | 5 | 4 | 5 | 9 | 7 |
| Ubiquitin carboxyl-terminal hydrolase 5 | USP5 | 2.6 | 0.0005 | 3.03 | 2.18 | 0.99 | 2.83 | 5.48 | 6.81 | 5.98 | 5.63 | 2 | 1 | 1 | 2 | 2 | 3 | 3 | 1 | 41 | 10 |
| Pyrroline-5-carboxylate reductase 1 | PYCR1 | 0.7 | 0.0014 | 23.84 | 24.54 | 28.28 | 23.43 | 19.08 | 15.42 | 17.95 | 15.36 | 7 | 7 | 10 | 6 | 8 | 5 | 6 | 5 | 9 | 11 |
| DNA replication licensing factor MCM4 | COX5A | 3.8 | 0.0026 | 3.08 | 3.22 | 0.00 | 1.83 | 9.98 | 6.00 | 7.80 | 6.85 | 2 | 1 | 0 | 2 | 5 | 4 | 6 | 3 | 73 | 22 |
| Cytochrome c oxidase subunit 5A | MCM4 | INF | 0.0026 | 0.00 | 0.00 | 0.00 | 0.00 | 2.71 | 4.28 | 3.42 | 1.43 | 0 | 0 | 0 | 0 | 2 | 2 | 2 | 1 | - | 41 |
| Histidine triad nucleotide-binding protein 2, mitochondrial | HINT2 | 4.1 | 0.0029 | 0.00 | 2.14 | 2.05 | 0.00 | 3.58 | 4.27 | 4.29 | 5.01 | 0 | 1 | 1 | 0 | 2 | 2 | 2 | 2 | 120 | 14 |
| NAD kinase 2, mitochondrial | NADK2 | 3.2 | 0.0032 | 2.02 | 4.23 | 3.30 | 3.75 | 6.34 | 11.94 | 12.88 | 11.05 | 1 | 2 | 3 | 1 | 3 | 5 | 6 | 4 | 29 | 28 |
| Dehydrogenase/reductase SDR family member 4 | DHRS4 | 5.6 | 0.004 | 1.08 | 0.00 | 0.00 | 0.96 | 2.76 | 3.42 | 1.67 | 3.47 | 1 | 0 | 0 | 1 | 2 | 1 | 1 | 2 | 120 | 30 |
| Endoplasmic reticulum resident protein 44 | ERP44 | 5.8 | 0.004 | 0.00 | 6.36 | 0.99 | 0.00 | 13.62 | 10.30 | 7.80 | 10.74 | 0 | 2 | 1 | 0 | 5 | 3 | 4 | 4 | 170 | 22 |
| Vesicular integral-membrane protein VIP36 | LMAN2 | 0.6 | 0.0049 | 10.21 | 9.58 | 13.67 | 12.20 | 5.51 | 6.80 | 6.02 | 8.18 | 4 | 3 | 4 | 3 | 2 | 3 | 2 | 2 | 16 | 18 |
| Peroxiredoxin-1 | PRDX1 | 0.8 | 0.0054 | 91.39 | 92.41 | 99.26 | 75.88 | 70.67 | 64.99 | 63.35 | 70.28 | 19 | 21 | 17 | 16 | 17 | 19 | 16 | 17 | 11 | 6 |
| 28S ribosomal protein S23 | MRPS23 | 16 | 0.0057 | 0.00 | 1.08 | 0.00 | 0.00 | 6.29 | 2.56 | 5.15 | 2.86 | 0 | 1 | 0 | 0 | 2 | 1 | 3 | 2 | 200 | 43 |
| Valine--tRNA ligase | VAR5 | 0.2 | 0.0058 | 20.75 | 14.95 | 9.44 | 23.34 | 0.00 | 2.56 | 5.98 | 3.99 | 8 | 6 | 5 | 10 | 0 | 2 | 3 | 2 | 36 | 80 |
| 60S ribosomal protein L29 | RPL29 | 0.9 | 0.0062 | 4.16 | 4.26 | 4.18 | 3.75 | 3.63 | 3.42 | 3.43 | 3.68 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 6 | 4 |
| Isoform 2 of Nucleobindin-2 | NUCB2 | 2.2 | 0.0067 | 5.24 | 1.10 | 3.08 | 3.75 | 7.26 | 7.67 | 6.00 | 8.49 | 1 | 1 | 1 | 1 | 3 | 3 | 2 | 2 | 52 | 14 |
| Thioredoxin reductase 1 | TXNRD1 | 0.5 | 0.0074 | 15.51 | 20.19 | 21.00 | 12.22 | 9.10 | 9.38 | 9.44 | 7.06 | 4 | 7 | 8 | 3 | 3 | 3 | 3 | 2 | 24 | 13 |
| Transitional endoplasmic reticulum ATPase | VCP | 1.2 | 0.0079 | 109.19 | 113.38 | 106.34 | 102.11 | 130.47 | 123.25 | 114.81 | 128.1 | 24 | 31 | 28 | 27 | 29 | 33 | 28 | 27 | 4 | 6 |
| L-lactate dehydrogenase A chain | LDHA | 0.7 | 0.0082 | 54.22 | 57.43 | 46.95 | 68.50 | 38.95 | 37.56 | 42.17 | 35.81 | 16 | 16 | 14 | 18 | 13 | 13 | 15 | 9 | 16 | 7 |
| 3-mercaptopyruvate sulfurtransferase | MPST | 1.4 | 0.0082 | 11.41 | 13.80 | 14.77 | 13.10 | 17.24 | 21.35 | 18.00 | 16.47 | 4 | 3 | 6 | 5 | 5 | 8 | 8 | 6 | 11 | 12 |
| Protein disulfide-isomerase A4 | PDIA4 | 1.2 | 0.0085 | 117.61 | 126.88 | 138.13 | 135.89 | 163.27 | 154.67 | 148.15 | 179.7 | 31 | 26 | 32 | 32 | 40 | 37 | 35 | 33 | 7 | 8 |
| Proteasome subunit beta type-5 | PSMB5 | 0.5 | 0.0086 | 18.59 | 17.03 | 12.52 | 19.64 | 8.17 | 6.02 | 12.01 | 10.32 | 7 | 5 | 5 | 6 | 3 | 3 | 7 | 3 | 19 | 28 |
| Bifunctional purine biosynthesis protein PURH | ATIC | 0.8 | 0.0092 | 62.35 | 60.73 | 65.75 | 65.60 | 42.55 | 53.81 | 57.29 | 48.98 | 16 | 19 | 22 | 18 | 14 | 21 | 19 | 16 | 4 | 13 |
| Enoyl-CoA hydratase, mitochondrial | ECHS1 | 1.4 | 0.0095 | 23.88 | 33.02 | 26.15 | 28.99 | 37.15 | 45.28 | 38.63 | 35.18 | 8 | 7 | 9 | 9 | 10 | 11 | 9 | 8 | 14 | 11 |
| 26S protease regulatory subunit 6B | PSMC4 | 1.6 | 0.0095 | 8.32 | 10.68 | 13.60 | 7.49 | 19.07 | 14.55 | 15.46 | 16.26 | 3 | 5 | 6 | 3 | 6 | 6 | 6 | 5 | 27 | 12 |
| 60S ribosomal protein L27a | RPL27A | 2.1 | 0.01 | 8.26 | 6.41 | 4.20 | 3.75 | 14.52 | 9.36 | 12.82 | 9.83 | 4 | 2 | 2 | 2 | 4 | 4 | 5 | 3 | 37 | 21 |

Table 7 continued...

| Protein name | Gene name | FC | P value | Quantitative value | | | | | | | | Exclusive peptides | | | | | | | | CV (%) | |
|---|-----------|------|---------|--------------------|--------|--------|--------|--------|--------|--------|--------|--------------------|----|----|----|----|----|----|----|--------|-----|
| | | | | C1 | C2 | C3 | C4 | F1 | F2 | F3 | F4 | C1 | C2 | C3 | C4 | F1 | F2 | F3 | F4 | C | F |
| Cytoplasmic dynein 1 heavy chain 1 | DYNC1H1 | 0.04 | 0.011 | 12.42 | 16.96 | 8.26 | 28.14 | 0.00 | 0.00 | 0.00 | 2.56 | 6 | 7 | 5 | 14 | 0 | 0 | 0 | 2 | 52 | 200 |
| ADP-ribosylation factor 3 | ARF3 | 0.2 | 0.011 | 13.49 | 5.29 | 12.61 | 16.91 | 2.76 | 5.09 | 1.73 | 1.43 | 3 | 1 | 4 | 3 | 2 | 3 | 1 | 1 | 40 | 60 |
| Activated RNA polymerase II transcriptional coactivator p15 | SUB1 | 1.8 | 0.011 | 11.47 | 4.30 | 12.55 | 11.24 | 16.35 | 16.20 | 17.17 | 20.48 | 4 | 1 | 3 | 4 | 4 | 4 | 4 | 3 | 38 | 11 |
| Eukaryotic translation initiation factor 3 subunit A | EIF3A | 0.6 | 0.013 | 9.33 | 7.48 | 7.27 | 9.35 | 6.34 | 2.52 | 5.91 | 3.68 | 3 | 3 | 2 | 4 | 2 | 3 | 3 | 3 | 14 | 39 |
| Transaldolase | TALDO1 | 0.7 | 0.013 | 24.96 | 24.47 | 30.25 | 21.52 | 15.42 | 14.52 | 18.82 | 20.46 | 9 | 8 | 10 | 8 | 6 | 4 | 6 | 6 | 14 | 16 |
| Alanine aminotransferase 2 | GPT2 | 1.2 | 0.013 | 29.00 | 31.96 | 27.10 | 33.76 | 36.23 | 35.02 | 37.55 | 35.38 | 9 | 7 | 8 | 11 | 13 | 12 | 11 | 8 | 10 | 3 |
| Ezrin | EZR | 0.8 | 0.014 | 35.45 | 30.82 | 32.18 | 29.05 | 19.99 | 25.64 | 28.26 | 23.23 | 10 | 9 | 12 | 9 | 8 | 11 | 10 | 7 | 8 | 15 |
| Isoform 2 of Filamin-B | FLNB | 0.8 | 0.016 | 202.26 | 181.10 | 159.48 | 212.65 | 138.10 | 148.85 | 131.11 | 159.75 | 62 | 58 | 54 | 65 | 49 | 61 | 47 | 47 | 12 | 9 |
| Coatomer subunit delta | ARCNI | 1.4 | 0.016 | 12.41 | 11.83 | 11.34 | 9.31 | 14.47 | 16.25 | 18.93 | 13.51 | 5 | 6 | 5 | 5 | 6 | 6 | 9 | 3 | 12 | 15 |
| Annexin A6 | ANXA6 | 2.5 | 0.016 | 15.43 | 14.92 | 7.33 | 11.18 | 44.39 | 32.44 | 26.52 | 19.65 | 7 | 8 | 2 | 6 | 16 | 13 | 11 | 8 | 31 | 34 |
| Multifunctional protein ADE2 | PAICS | 0.7 | 0.017 | 27.05 | 35.17 | 28.33 | 35.58 | 19.99 | 21.35 | 27.42 | 20.24 | 10 | 11 | 10 | 11 | 9 | 8 | 11 | 6 | 14 | 16 |
| Proteasome activator complex subunit 3 | PSME3 | 1.5 | 0.018 | 12.48 | 9.56 | 14.70 | 11.25 | 17.32 | 21.40 | 18.74 | 14.21 | 3 | 3 | 5 | 3 | 5 | 6 | 6 | 3 | 18 | 17 |
| Phosphoglycerate mutase 1 | PGAM1 | 0.7 | 0.019 | 36.33 | 45.79 | 38.40 | 30.99 | 27.96 | 24.86 | 30.83 | 24.74 | 10 | 8 | 11 | 9 | 10 | 9 | 10 | 7 | 16 | 11 |
| Keratin, type I cytoskeletal 18 | KRT18 | 1.2 | 0.019 | 125.98 | 110.79 | 119.95 | 109.64 | 139.50 | 149.51 | 122.52 | 143.47 | 20 | 20 | 23 | 19 | 25 | 24 | 24 | 24 | 7 | 8 |
| Isoform D of Eukaryotic translation initiation factor 4 gamma 1 | EIF4G1 | 0.1 | 0.02 | 6.24 | 6.45 | 1.10 | 7.48 | 0.00 | 1.69 | 0.86 | 0.00 | 2 | 4 | 1 | 3 | 0 | 2 | 1 | 0 | 54 | 130 |
| Transmembrane emp24 domain-containing protein 9 | TMED9 | 0.6 | 0.02 | 17.71 | 16.93 | 13.54 | 15.00 | 6.39 | 13.65 | 8.59 | 11.35 | 4 | 5 | 4 | 4 | 3 | 4 | 3 | 5 | 12 | 32 |
| Sialidase-1 | NEU1 | 1.4 | 0.02 | 15.57 | 22.39 | 15.63 | 15.00 | 25.27 | 29.08 | 22.29 | 22.00 | 5 | 6 | 7 | 6 | 7 | 11 | 7 | 6 | 20 | 13 |
| Perilipin-2 | PLIN2 | 2 | 0.02 | 14.43 | 35.12 | 41.71 | 10.32 | 50.74 | 52.16 | 47.16 | 48.89 | 6 | 10 | 13 | 4 | 18 | 16 | 15 | 12 | 61 | 4 |
| Acylpyruvase FAHD1, mitochondrial | FAHD1 | 2.3 | 0.02 | 5.17 | 6.44 | 1.03 | 3.74 | 10.02 | 8.52 | 6.85 | 13.00 | 2 | 3 | 1 | 1 | 3 | 3 | 3 | 2 | 57 | 27 |
| Isoform 3 of Alpha-actinin-1 | ACTN1 | 0.8 | 0.021 | 86.33 | 84.95 | 67.73 | 88.06 | 61.69 | 66.72 | 69.41 | 49.91 | 11 | 11 | 9 | 11 | 9 | 10 | 9 | 6 | 12 | 14 |
| Isoform 2 of Nodal modulator 2 | NOMO2 | 0.4 | 0.022 | 22.82 | 9.62 | 13.60 | 18.77 | 8.17 | 2.54 | 7.77 | 7.15 | 8 | 4 | 5 | 7 | 2 | 2 | 4 | 3 | 36 | 41 |
| Medium-chain specific acyl-CoA dehydrogenase, mitochondrial | ACADM | 1.4 | 0.022 | 9.26 | 10.59 | 7.37 | 10.29 | 13.66 | 14.55 | 15.45 | 10.32 | 4 | 3 | 3 | 4 | 4 | 4 | 5 | 3 | 16 | 17 |
| Transcription intermediary factor 1-beta | TRIM28 | 1.5 | 0.022 | 23.84 | 39.41 | 24.94 | 31.86 | 45.17 | 43.60 | 36.80 | 50.65 | 8 | 14 | 13 | 8 | 15 | 17 | 13 | 12 | 24 | 13 |
| Alpha-actinin-4 | ACTN4 | 0.8 | 0.023 | 129.02 | 159.39 | 142.86 | 134.86 | 117.07 | 123.18 | 124.97 | 107.90 | 32 | 39 | 38 | 36 | 31 | 34 | 36 | 33 | 9 | 7 |
| Trifunctional enzyme subunit alpha | HADHA | 1.3 | 0.023 | 59.34 | 58.83 | 59.43 | 49.74 | 78.86 | 69.17 | 60.72 | 78.98 | 15 | 19 | 19 | 18 | 25 | 21 | 20 | 21 | 8 | 12 |

Table 7 continued...

| Protein name | Gene name | FC | P value | Quantitative value | | | | | | | | Exclusive peptides | | | | | | | | CV (%) | |
|--|-----------|-----|---------|--------------------|-------|-------|-------|-------|-------|-------|-------|--------------------|----|----|----|----|----|----|----|--------|-----|
| | | | | C1 | C2 | C3 | C4 | F1 | F2 | F3 | F4 | C1 | C2 | C3 | C4 | F1 | F2 | F3 | F4 | C | F |
| 3-hydroxyacyl-CoA dehydrogenase type-2 | HSD17B10 | 1.3 | 0.023 | 24.98 | 26.58 | 27.14 | 36.59 | 39.91 | 36.60 | 37.85 | 35.20 | 8 | 8 | 9 | 9 | 11 | 9 | 11 | 9 | 18 | 5 |
| 45 kDa calcium-binding protein | SDF4 | 2.1 | 0.023 | 8.26 | 3.12 | 4.14 | 10.27 | 10.85 | 12.86 | 12.03 | 17.80 | 2 | 1 | 3 | 4 | 5 | 6 | 5 | 4 | 52 | 23 |
| Keratin, type II cytoskeletal 1 | KRT1 | 3.8 | 0.023 | 18.65 | 3.31 | 11.64 | 0.00 | 34.44 | 28.91 | 47.89 | 16.56 | 3 | 1 | 3 | 0 | 6 | 7 | 12 | 4 | 100 | 41 |
| Acyl-CoA synthetase short-chain family member 3, mitochondrial | ACSS3 | 0.2 | 0.024 | 5.17 | 5.24 | 4.10 | 11.25 | 0.92 | 1.67 | 1.73 | 1.74 | 3 | 1 | 4 | 5 | 1 | 2 | 1 | 1 | 50 | 26 |
| DNA-(apurinic or apyrimidinic site) lyase | APEX1 | 0.4 | 0.025 | 11.40 | 10.72 | 5.24 | 6.60 | 3.63 | 4.29 | 0.89 | 4.50 | 3 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 36 | 50 |
| Saccharopine dehydrogenase-like oxidoreductase | SCCPDH | 0.4 | 0.025 | 5.17 | 3.21 | 3.15 | 5.60 | 0.92 | 0.84 | 3.43 | 1.43 | 2 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 30 | 73 |
| Ubiquitin-like modifier-activating enzyme 1 | UBA1 | 0.8 | 0.025 | 78.12 | 71.28 | 68.92 | 63.68 | 47.12 | 59.77 | 64.35 | 53.51 | 19 | 18 | 20 | 20 | 14 | 20 | 20 | 12 | 9 | 13 |
| Perilipin-3 | PLIN3 | 1.4 | 0.025 | 23.90 | 37.31 | 41.65 | 25.25 | 42.65 | 49.49 | 50.42 | 43.15 | 7 | 11 | 14 | 9 | 11 | 15 | 16 | 14 | 27 | 9 |
| Calpain-1 catalytic subunit | CAPN1 | 0.4 | 0.026 | 5.30 | 8.50 | 6.34 | 11.22 | 1.85 | 2.53 | 6.00 | 1.12 | 2 | 2 | 3 | 3 | 2 | 3 | 2 | 1 | 33 | 75 |
| 2',3'-cyclic-nucleotide 3'-phosphodiesterase | CNP | 0.6 | 0.026 | 11.41 | 9.62 | 8.30 | 13.14 | 4.50 | 5.11 | 9.39 | 5.42 | 3 | 4 | 3 | 3 | 2 | 2 | 3 | 3 | 20 | 36 |
| Adapter molecule crk | CRK | 0.1 | 0.027 | 1.01 | 1.08 | 3.11 | 1.86 | 0.00 | 0.85 | 0.00 | 0.00 | 1 | 1 | 2 | 1 | 0 | 1 | 0 | 0 | 55 | 200 |
| Dihydrolipoyl dehydrogenase, mitochondrial | DLD | 1.6 | 0.027 | 10.40 | 18.12 | 14.38 | 17.88 | 24.52 | 23.17 | 18.04 | 29.46 | 4 | 6 | 7 | 6 | 10 | 12 | 6 | 8 | 24 | 20 |
| DNA replication licensing factor MCM7 | MCM7 | 2.9 | 0.027 | 4.09 | 8.51 | 1.10 | 1.88 | 8.18 | 14.47 | 15.37 | 7.87 | 3 | 2 | 1 | 2 | 4 | 6 | 7 | 4 | 85 | 35 |
| 39S ribosomal protein L45 | MRPL45 | 3.2 | 0.027 | 0.00 | 1.05 | 1.10 | 0.94 | 1.79 | 2.54 | 1.68 | 3.99 | 0 | 1 | 1 | 1 | 2 | 1 | 2 | 2 | 67 | 43 |
| Leukocyte elastase inhibitor | SERPINB1 | 0.3 | 0.028 | 5.23 | 3.15 | 7.33 | 4.66 | 0.00 | 0.88 | 4.23 | 0.00 | 2 | 2 | 3 | 2 | 0 | 1 | 2 | 0 | 34 | 160 |
| 14-3-3 protein zeta/delta | YWHAZ | 0.9 | 0.029 | 55.13 | 46.80 | 52.29 | 49.64 | 42.63 | 45.25 | 48.02 | 43.68 | 9 | 9 | 9 | 8 | 8 | 8 | 10 | 7 | 7 | 5 |
| Ras-related C3 botulinum toxin substrate 1 | RAC1 | 0.2 | 0.03 | 4.16 | 5.33 | 0.99 | 6.54 | 0.87 | 0.00 | 0.83 | 1.43 | 1 | 2 | 1 | 2 | 1 | 0 | 1 | 1 | 56 | 75 |
| Cathepsin D | CTSD | 1.2 | 0.03 | 73.22 | 62.46 | 64.24 | 63.61 | 85.38 | 67.40 | 83.03 | 89.73 | 14 | 14 | 11 | 14 | 17 | 15 | 17 | 15 | 8 | 12 |
| LDLR chaperone MESD | MESDC2 | 1.5 | 0.03 | 4.16 | 6.41 | 8.34 | 9.33 | 12.67 | 11.13 | 10.33 | 9.31 | 2 | 3 | 3 | 5 | 4 | 4 | 5 | 3 | 32 | 13 |
| Isoform 2 of Tripeptidyl-peptidase 1 | TPP1 | 1.7 | 0.03 | 8.25 | 7.48 | 9.36 | 6.58 | 19.02 | 10.29 | 11.96 | 12.99 | 2 | 3 | 3 | 2 | 6 | 5 | 4 | 3 | 15 | 28 |
| 60S ribosomal protein L24 | RPL24 | 3.4 | 0.031 | 3.09 | 1.10 | 2.20 | 0.92 | 2.76 | 7.71 | 9.39 | 4.90 | 3 | 1 | 2 | 1 | 1 | 3 | 2 | 3 | 56 | 48 |
| Toll-interacting protein | TOLLIP | 1.7 | 0.032 | 2.08 | 3.22 | 4.07 | 3.75 | 4.50 | 5.98 | 4.29 | 7.06 | 1 | 1 | 3 | 2 | 2 | 2 | 2 | 1 | 27 | 24 |
| Laminin subunit gamma-1 | LAMC1 | 0.6 | 0.035 | 2.08 | 3.12 | 2.09 | 1.90 | 0.92 | 1.69 | 1.73 | 1.12 | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 1 | 24 | 30 |
| Tubulin-tyrosine ligase-like protein 12 | TTLL12 | 0.5 | 0.037 | 11.47 | 10.69 | 8.39 | 14.06 | 7.25 | 5.12 | 8.56 | 0.00 | 4 | 4 | 3 | 4 | 2 | 3 | 4 | 0 | 21 | 72 |

Table 7 continued...

| Protein name | Gene name | FC | P value | Quantitative value | | | | | | | | Exclusive peptides | | | | | | | | CV (%) | |
|---|-----------|-----|---------|--------------------|--------|--------|--------|--------|--------|--------|--------|--------------------|----|----|----|----|----|----|----|--------|----|
| | | | | C1 | C2 | C3 | C4 | F1 | F2 | F3 | F4 | C1 | C2 | C3 | C4 | F1 | F2 | F3 | F4 | C | F |
| Aldose 1-epimerase | GALM | 1.6 | 0.037 | 5.16 | 3.21 | 4.23 | 7.51 | 9.09 | 8.59 | 9.39 | 5.94 | 3 | 1 | 2 | 4 | 4 | 5 | 4 | 1 | 37 | 19 |
| RNA-binding protein Raly | RALY | 3.2 | 0.037 | 1.01 | 4.26 | 2.05 | 0.00 | 9.05 | 5.98 | 3.43 | 4.90 | 1 | 1 | 1 | 0 | 2 | 2 | 1 | 2 | 100 | 41 |
| DnaJ homolog subfamily B member 1 | DNAJB1 | 0.7 | 0.038 | 5.24 | 3.11 | 5.24 | 4.69 | 3.58 | 2.56 | 3.43 | 2.77 | 1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 22 | 16 |
| Villin-1 | VIL1 | 0.9 | 0.038 | 85.06 | 92.57 | 79.65 | 86.13 | 77.90 | 77.71 | 81.40 | 74.04 | 27 | 26 | 25 | 29 | 24 | 25 | 27 | 25 | 6 | 4 |
| 39S ribosomal protein L46, mitochondrial | MRPL46 | 1.5 | 0.038 | 4.16 | 4.26 | 6.29 | 6.56 | 6.33 | 8.54 | 9.48 | 7.06 | 2 | 1 | 3 | 4 | 3 | 3 | 4 | 2 | 24 | 18 |
| Elongation factor 1-gamma | EEF1G | 0.7 | 0.039 | 24.83 | 39.40 | 33.37 | 30.91 | 23.52 | 17.96 | 21.36 | 27.51 | 9 | 10 | 10 | 10 | 9 | 7 | 9 | 7 | 19 | 18 |
| Profilin-1 | PFN1 | 0.8 | 0.039 | 63.50 | 55.37 | 70.94 | 51.54 | 44.41 | 46.06 | 45.42 | 54.12 | 7 | 9 | 11 | 8 | 10 | 8 | 8 | 8 | 14 | 9 |
| AT-rich interactive domain-containing protein 3A | ARID3A | 9.8 | 0.039 | 0.00 | 0.00 | 0.00 | 0.92 | 3.63 | 0.85 | 3.40 | 1.12 | 0 | 0 | 0 | 1 | 1 | 1 | 2 | 1 | 200 | 65 |
| Peptidyl-prolyl cis-trans isomerase FKBP4 | FKBP4 | 1.6 | 0.04 | 5.17 | 11.73 | 7.31 | 6.59 | 10.84 | 11.96 | 16.28 | 11.35 | 2 | 4 | 4 | 3 | 4 | 5 | 7 | 4 | 37 | 20 |
| Elongation factor 2 | EEF2 | 0.9 | 0.042 | 168.61 | 180.76 | 141.05 | 169.62 | 129.74 | 146.87 | 146.33 | 140.65 | 36 | 39 | 35 | 37 | 35 | 35 | 35 | 28 | 10 | 6 |
| High mobility group protein B1 | HMGB1 | 0.7 | 0.043 | 13.42 | 25.40 | 24.02 | 20.56 | 15.48 | 11.96 | 13.71 | 13.91 | 4 | 7 | 6 | 6 | 6 | 4 | 4 | 4 | 26 | 10 |
| Succinyl-CoA ligase [GDP-forming] subunit beta, mitochondrial | SUCLG2 | 1.4 | 0.046 | 13.42 | 17.02 | 13.60 | 19.64 | 23.60 | 27.31 | 23.16 | 16.56 | 4 | 4 | 4 | 7 | 7 | 9 | 8 | 6 | 19 | 20 |
| V-type proton ATPase subunit B, brain isoform | ATP6V1B2 | 1.5 | 0.046 | 12.42 | 9.59 | 10.69 | 19.72 | 17.26 | 22.19 | 19.79 | 18.41 | 6 | 4 | 5 | 8 | 7 | 10 | 8 | 3 | 35 | 11 |
| Lupus La protein | SSB | 1.2 | 0.047 | 34.36 | 42.67 | 41.60 | 37.48 | 44.57 | 43.52 | 47.31 | 57.08 | 9 | 12 | 16 | 13 | 16 | 15 | 16 | 12 | 10 | 13 |
| Acyl-CoA synthetase family member2 | ACSF2 | 1.6 | 0.047 | 6.17 | 7.47 | 8.43 | 6.56 | 10.83 | 8.56 | 10.26 | 16.17 | 2 | 3 | 4 | 2 | 5 | 5 | 5 | 6 | 14 | 29 |
| Arylsulfatase B | ARSB | 8.5 | 0.047 | 0.00 | 0.00 | 0.00 | 0.92 | 2.70 | 2.52 | 2.54 | 0.00 | 0 | 0 | 0 | 1 | 1 | 2 | 2 | 0 | 200 | 67 |
| Elongation factor 1-alpha 1 | EEF1A1 | 0.7 | 0.048 | 126.1 | 97.75 | 161.57 | 119.96 | 104.93 | 79.51 | 96.59 | 78.18 | 18 | 18 | 18 | 19 | 19 | 17 | 17 | 16 | 21 | 15 |
| 40S ribosomal protein S19 | RPS19 | 1.5 | 0.048 | 16.78 | 11.74 | 15.67 | 14.06 | 21.77 | 17.10 | 17.99 | 27.82 | 7 | 5 | 5 | 6 | 9 | 7 | 5 | 7 | 15 | 23 |
| Obg-like ATPase 1 | OLA1 | 0.9 | 0.05 | 12.48 | 11.68 | 14.64 | 14.06 | 10.89 | 11.09 | 11.98 | 11.75 | 4 | 5 | 6 | 5 | 3 | 4 | 4 | 3 | 10 | 5 |
| Ethylmalonyl-CoA decarboxylase | ECHDC1 | 11 | 0.05 | 0.00 | 0.00 | 0.00 | 0.96 | 3.63 | 2.58 | 4.32 | 0.00 | 0 | 0 | 0 | 1 | 2 | 1 | 3 | 0 | 200 | 72 |

CHAPTER 5: PROTEOMICS ANALYSIS – PART II

5.1. Introduction

The analysis of datasets LC1 and LC2 have highlighted the importance of experimental design and LC-MS/MS parameters. The optimization of LC-MS/MS parameters described in chapter 4 increased the total number of proteins identified in LC2, but this improvement did not help identify more proteins with statistically significant changes in protein abundance. The underlying reasons for this could be the experimental design – while dataset LC1 was generated from pooled samples, dataset LC2 was generated from individual biological replicates. A direct comparison of datasets LC1 and LC2 is made difficult by the fact that they were generated from different biological replicates that originate in different stock vials. Moreover, the LC-MS/MS method used to generate LC2 was optimized and therefore different from that used to generate dataset LC1. To overcome these problems and to examine the impact of pooling samples on the proteomics analysis of HepG2 cells with accumulated lipids, two more datasets – LC3 and LC4 were obtained. These two datasets were generated from the same cell lysates – the lysate of each biological replicate was divided into two aliquots and one aliquot was used for each dataset. The lysates prepared for datasets LC3 and LC4 originated in cell cultures grown from a stock vial received from the Malaghan Institute and were different from those used for LC1 and LC2. Dataset LC3 represents data from individual biological replicates while dataset LC4 represents data of a pooled sample of biological replicates. The commercial availability of DGAT1 inhibitor PF04 allowed the inclusion of DGAT inhibition as a treatment variable and was included in the proteomics analysis for dataset LC3 and LC4. This chapter describes the proteomics analysis of these two datasets.

Functional analysis of proteins

The biological relevance of the identified proteins using 'shotgun proteomics' can be assessed through functional analysis of the proteins. Functional analysis or enrichment analysis is based on the rationale that a treatment affecting one protein in a biological pathway is likely to affect other co-functioning proteins in that pathway and so, the increased detection of these co-functioning proteins by a screening method reflects an increased probability of the represented pathway being affected by the treatment [295]. Enrichment analysis tools like GoMiner, DAVID and WebGeStalt are used to assign proteins

detected in proteomic data to defined categories called ‘annotation terms’ in databases like Gene Ontology (GO) and KEGG pathways [296]. Depending on the database, these annotation terms represent biological processes, localization or molecular function of the protein. A comparison of the observed frequency of annotation term to the frequency expected by chance allows for the calculation of enrichment p-values; these p-values denote the probability that the annotation term plays an important role in the biological events resulting from the treatment [297]. Enrichment of annotation terms suggests a high probability of the function described playing an important role in the cellular response to the treatment. The presence of the proteins identified in the enriched pathways can then be verified through orthogonal methods such as Western blotting. This chapter describes the functional analysis of proteins identified as changing in abundance after FA treatment in dataset LC4 and the efforts made to validate the changes in protein abundance reported in datasets LC3 and LC4.

5.2. Results

HepG2 cells treated with 100 μ M PF04 and 1 mM FA mixture for 6 h are referred to as ‘inhibitor-treated’ cells in this chapter. FA-treated cells and control cells refer to HepG2 cells treated with 1 mM FA mixture and untreated cells respectively. Since the objective was to assess the impact of pooling samples as well as to identify the maximum number of protein possible, the optimized LC-MS/MS and protein identification parameters used to generate dataset LC2 were used to generate datasets LC3 and LC4. This included the use of Gradient 1 and Top 6 method with dynamic repeat count set at 1 and monoisotopic peak selection enabled. Combined 1 workflow with SwissProt database filtered to retain proteins with evidence of existence was used for protein identification in PD. The setup for Scaffold depended on the experimental design and is described for each dataset in the sections describing label-free quantification for the respective datasets.

5.2.1. Dataset LC3

5.2.1.1. Protein identification

The three treatment groups – control, FA-treated and inhibitor-treated – were represented by four biological replicates per group and each biological replicate was represented by data collected over four LC-MS/MS runs. These data constituted dataset LC3. All proteins identified in PD were filtered to retain only those identified with high-confidence (FDR < 0.01) and at least 2 peptides per protein. Using the spectral information available from all 16

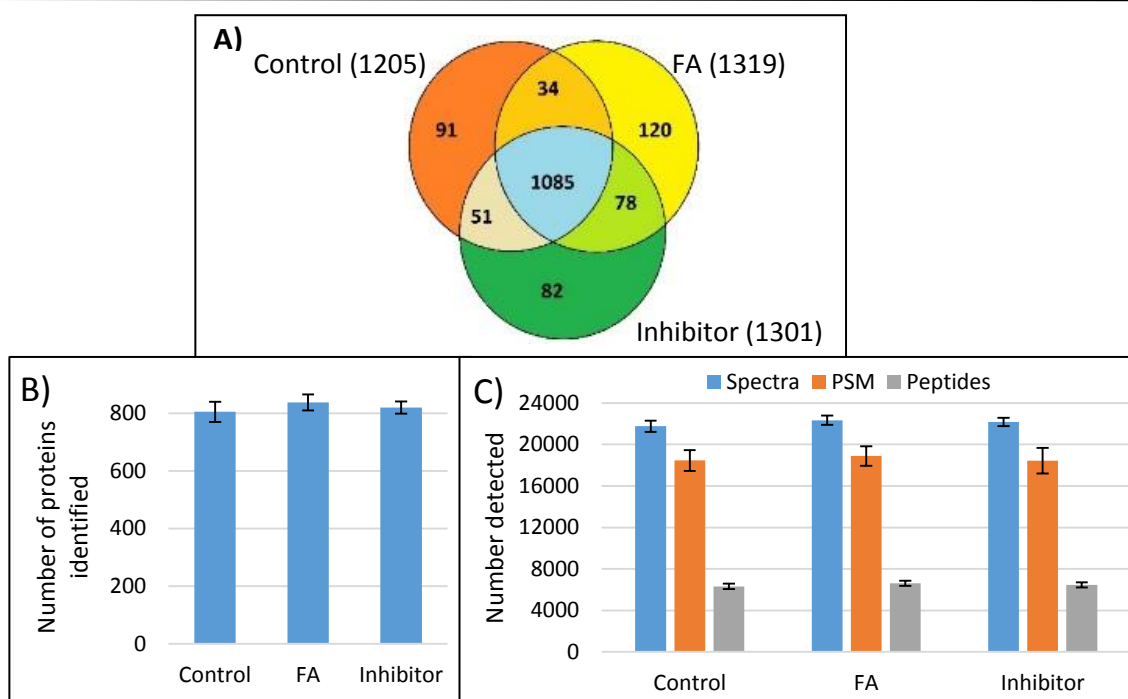


Figure 49: Number of proteins identified for dataset LC3 in Proteome Discoverer. A) Comparison of proteins identified in the three treatment groups – control, FA-treated and inhibitor-treated using Venny 2.1. **B)** The average number of proteins identified for a biological replicate of the treatment groups indicated. **C)** Average number of spectra, peptide-spectra matches (PSMs) and peptides for a biological sample of a treatment group as identified in PD.

LC-MS/MS runs for a given treatment group, the total number of proteins identified for the control group was 1,205. The number of proteins identified for the FA-treated and inhibitor-treated groups was 1,319 and 1,301 respectively. About 70 % of proteins identified in the dataset were present in all 3 treatment groups (Figure 49A). Based on the number of protein identified in each LC-MS/MS run for a given biological replicate, an average of 805, 838 and 821 proteins could be identified in biological replicates of control, FA-treated and inhibitor-treated, respectively. (Figure 49B). On average, 6,321; 6,628 and 6,462 distinct peptides were identified for the biological replicates of the control, FA-treated and inhibitor-treated groups, respectively. These identifications were based on an average of 18,461, 18,888 and 18,440 peptide-spectra matches made in PD for the 21,763, 22,340 and 22,181 spectra detected for the control, FA-treated and inhibitor-treated groups, respectively (Figure 49C).

5.2.1.2. Label-free quantification using spectral counts

The protein lists generated for each biological replicate in PD were loaded onto Scaffold as described in section 4.2.5 for quantification. All three groups – control, FA-treated and inhibitor-treated – were included in the analysis. A total of 1,102 proteins with no decoys were identified for the three groups analysed in Scaffold (Table 21 in Appendix). Peptide identifications were accepted if they could be established at greater than 99.0 % probability

by the Peptide Prophet algorithm, and protein identifications were accepted if they could be established at greater than 99.9 % probability and contained at least 2 identified peptides. One-way ANOVA was used to compare the mean QV for each protein across the three groups. At a significance level of 95 %, the change in abundance of the protein was considered significant for values of $p \leq 5 \times 10^{-5}$ after multiple testing correction using the Hochberg-Benjamini method. However, none of the identified proteins reached this level of significance. The analysis was continued with a comparison of two groups at a time using Student's *t*-tests. Two primary comparisons were made – control vs. FA-treated and FA-treated vs. inhibitor-treated. Of minor importance was a comparison between the control vs the inhibitor-treated. As with one-way ANOVA, at a significance level of 95 %, the difference was considered significant for $p \leq 5 \times 10^{-5}$ after multiple testing correction using the Hochberg-Benjamini method. Again, none of the identified proteins satisfied this test of significance. The data were analysed further as described in the next section to better understand the underlying reasons for this lack of significant changes.

5.2.1.3. Variance associated with data generated from unpooled replicates

Of the 1,102 proteins identified in Scaffold, 1,091 proteins were found in all three groups. Eleven proteins were identified in two of the three groups (Figure 50A). The number of peptides common to all three groups was 6,949, and this corresponded to 8,348 unique spectra that were common to all three groups (Figure 50B and 50C). The total number of unique spectra identified in the control, FA-treated and inhibitor-treated groups was 10,681; 10,514 and 10,621, respectively. The average number of spectra identified for a replicate in the control group was 6,821 and the number was 6,914 and 6,717 for the FA-treated and inhibitor-treated groups, respectively. The difference between the average number of total

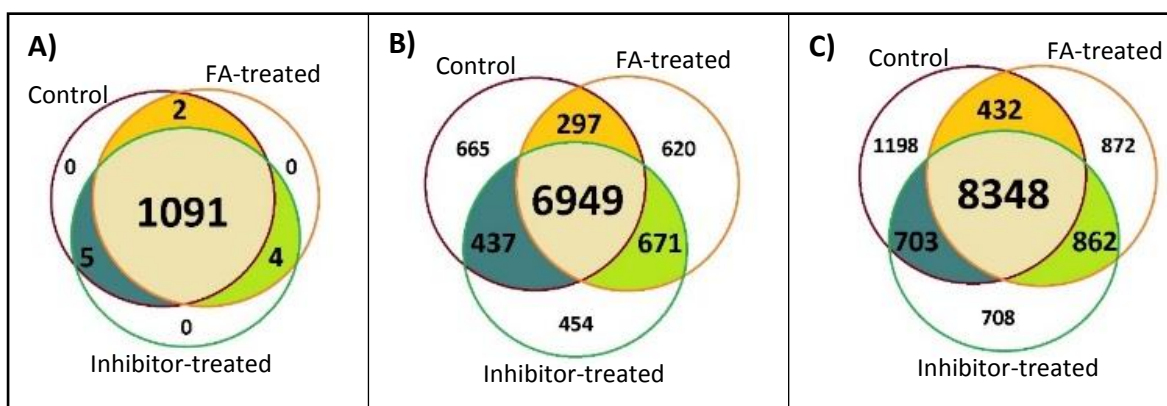


Figure 50: Identifications of proteins, peptides and MS/MS spectra in different treatment groups in Scaffold. A) Proteins identified in each treatment group. B) Total unique peptide counts for all proteins in the dataset. C) The total number of unique spectra detected for all the peptides in the dataset.

spectra identified for individual replicates and that identified for the group as a whole suggested that each individual replicate contained several spectra that were unique to it and may not have been present in the other biological replicates of the same treatment group. It was inferred, therefore, that spectral counts assigned to a protein did not just vary among the three treatment groups but also varied among the four biological replicates within a group. Further, the identification of a protein may be based on the peptides exclusive to a given replicate or a group and not found in other groups or even other replicates within the same group. While the difference in the spectral counts of a protein among the groups is the basis of label-free quantification using spectral counts, the variance among the biological replicates was the likely cause of the lack of significance of the data.

The variance observed in the data was attributed to biological variance between the biological replicates and technical variance between the four LC-MS/MS runs in each biological replicate. The difference among the biological replicates was assessed by comparing the list of proteins identified in each biological replicate to the protein lists of the other three replicates within the same group. To this end, a consensus protein list was created for each biological replicate. The proteins identified in this list were identified based on the collated information from all four LC-MS/MS runs performed for a given biological replicate and thus represents a consensus list of proteins for the biological replicate irrespective of the technical variance among the different LC-MS/MS runs. (Table 21 of Appendix).

A comparison of the collated protein lists showed that at least 80 % of the proteins identified in a replicate were present in the other three biological replicates of the same treated group (Figure 51). Of the remaining proteins, about 10 % were present in three of the four replicates of the control group and about 8.4 % and 8.3 % were present in three of the four replicates of the FA-treated and inhibitor-treated groups, respectively. The CV for the QV assigned to a protein within a group was then calculated in order to better understand the cumulative effect of technical and biological variance. CV is defined as the ratio of the standard deviation within a group to the mean value of the group and provides a measure of how dispersed the data are. The control group had 174 proteins with $CV \leq 10\%$; the FA-treated group and the inhibitor-treated group had 116 and 167 such proteins, respectively (Figure 52). This represents about 10 % of the proteins identified within a group. Of the proteins with $CV \leq 10\%$, only 12 proteins fulfilled the criterion of $CV \leq 10\%$ in all three groups (Table 8 and 9).

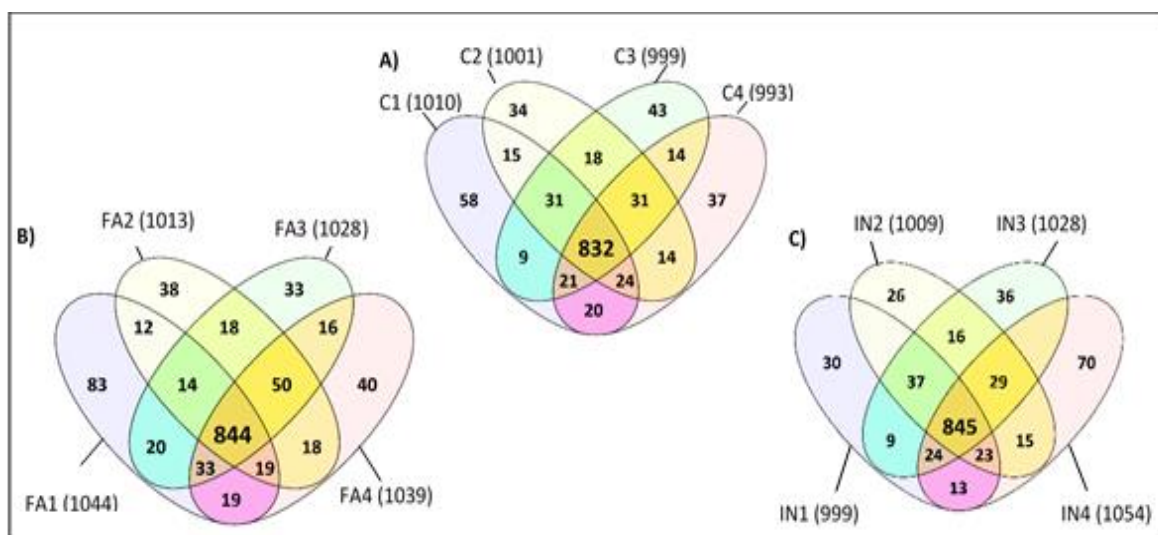


Figure 51: Comparison of proteins identified in the biological replicates of dataset LC3. A) Replicates of control (C); B) Replicates of FA-treated (FA) and C) Replicates of inhibitor -treated (IN).

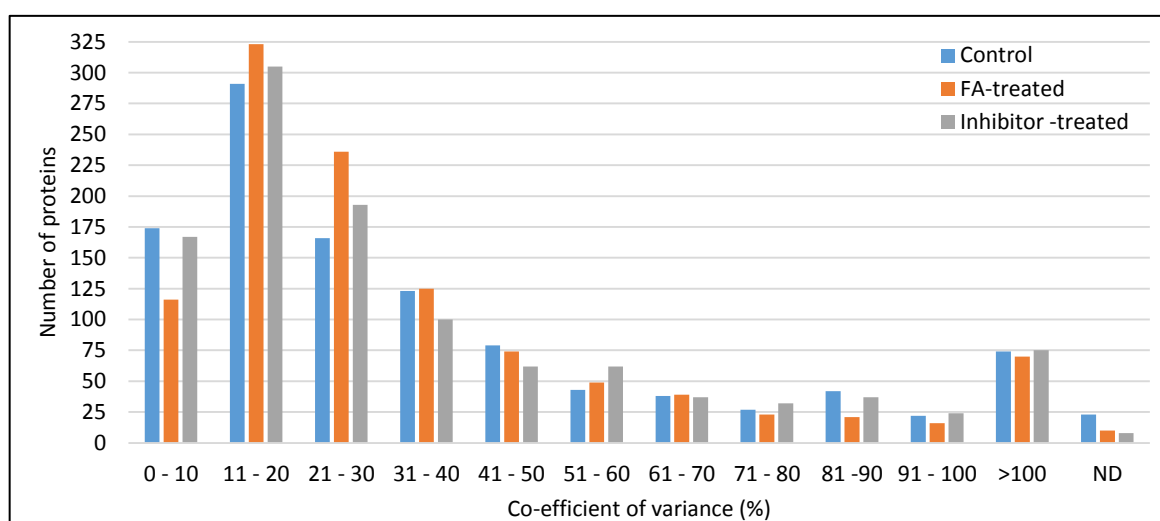


Figure 52: Distribution of proteins with respect to the co-efficient of variance for LC3. The CV of a protein for the four biological replicates of a group was calculated. The number of proteins that lie within a range of CV are indicated.

Table 8: Proteins with co-efficient of variance $\leq 10\%$ in dataset LC3. The proteins listed in this table are those that met the criteria of having a CV $\leq 10\%$. The fold change is the ratio of the mean QVs of the three groups taken two at a time. For example, 'F: C' is the ratio of mean QV in FA-treated (F) group to mean QV in control (C) group. The inhibitor-treated group is represented by 'I'. P values were calculated using a Student's *t*-test. Please refer to Appendix file LC3_Unpooled data.xlsx for QV and exclusive unique peptides of the proteins in this table.

| Protein Name | Gene name | Fold change | | | p value (t-test) | | | Co-efficient of variance (%) | | |
|---|-----------|-------------|-----|-----|------------------|-------|--------|------------------------------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I |
| 14-3-3 protein gamma | YWHAG | 1 | 1 | 1 | 0.62 | 0.62 | 0.87 | 8 | 10 | 4 |
| Alpha-actinin-4 | ACTN4 | 1 | 1 | 1.1 | 0.87 | 0.49 | 0.14 | 4 | 8 | 5 |
| Heat shock 70 kDa protein 1A/1B | HSPA1A | 0.9 | 1.1 | 1 | 0.063 | 0.046 | 0.9 | 7 | 10 | 2 |
| Heat shock protein HSP 90-beta | HSP90AB1 | 0.9 | 1.1 | 1 | 0.28 | 0.17 | 0.55 | 8 | 10 | 4 |
| HLA class I histocompatibility antigen, A-2 alpha chain | HLA-A | 1 | 1.1 | 1.1 | 0.88 | 0.07 | 0.002 | 6 | 9 | 8 |
| Isoform 2 of Eukaryotic translation initiation factor 3 subunit B | EIF3B | 0.9 | 1.1 | 1.1 | 0.34 | 0.12 | 0.45 | 9 | 5 | 9 |
| Isoform 4 of Protein disulfide-isomerase A6 | PDIA6 | 0.9 | 1.1 | 1 | 0.15 | 0.13 | 0.7 | 9 | 10 | 9 |
| Multifunctional protein ADE2 | PAICS | 1.3 | 1 | 1.4 | 0.0011 | 0.65 | 0.0008 | 10 | 5 | 10 |
| Nicotinate-nucleotide pyrophosphorylase [carboxylating] | QPRT | 0.77 | 1.1 | 0.9 | 0.007 | 0.022 | 0.05 | 10 | 9 | 2 |
| Protein disulfide-isomerase A4 | PDIA4 | 0.77 | 1.1 | 1 | 0.021 | 0.024 | 0.48 | 10 | 5 | 6 |
| Ras-related protein Rab-10 | RAB10 | 0.9 | 1.1 | 1 | 0.27 | 0.29 | 0.61 | 6 | 8 | 10 |
| Transitional endoplasmic reticulum ATPase | VCP | 0.77 | 1.2 | 1 | 0.013 | 0.023 | 0.84 | 3 | 10 | 3 |

Table 9: QV and exclusive peptide counts for proteins with co-efficient of variance $\leq 10\%$ in dataset LC3. The proteins listed in this table are those that met the criteria of having a CV $\leq 10\%$ in all three groups. The values listed are QV and exclusive peptide counts for each replicate of the control (C), FA-treated (F) and inhibitor-treated (I) groups.

| Protein Name | Quantitative value | | | | | | | | | | | | Exclusive peptide count | | | | | | | | | | | |
|---|--------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------------------------|----|----|----|----|----|----|----|----|----|----|----|
| | C1 | C2 | C4 | C3 | F1 | F2 | F3 | F4 | I1 | I2 | I3 | I4 | C1 | C2 | C4 | C3 | F1 | F2 | F3 | F4 | I1 | I2 | I3 | I4 |
| 14-3-3 protein gamma | 48.3 | 41.4 | 49.3 | 45.5 | 35.7 | 47.1 | 45.6 | 47.5 | 47.1 | 50.0 | 45.6 | 43.2 | 6 | 5 | 7 | 6 | 5 | 5 | 4 | 5 | 8 | 6 | 6 | 5 |
| Alpha-actinin-4 | 124.2 | 141.6 | 138.8 | 133.9 | 117.1 | 145.8 | 130.5 | 140.3 | 142.4 | 138.0 | 147.7 | 140.1 | 33 | 34 | 38 | 37 | 38 | 37 | 32 | 38 | 35 | 35 | 38 | 37 |
| Eukaryotic translation initiation factor 3 subunit B | 23.2 | 28.3 | 30.5 | 27.8 | 23.0 | 27.7 | 25.8 | 25.3 | 25.9 | 29.5 | 30.3 | 30.0 | 8 | 8 | 8 | 7 | 8 | 6 | 6 | 6 | 8 | 7 | 6 | 7 |
| Heat shock 70 kDa protein 1A/1B | 123.5 | 131.7 | 120.6 | 135.9 | 93.9 | 113.3 | 107.5 | 126.4 | 131.3 | 132.5 | 125.6 | 120.0 | 23 | 23 | 22 | 23 | 21 | 21 | 21 | 26 | 24 | 23 | 23 | 23 |
| Heat shock protein HSP 90-beta | 314.4 | 289.3 | 284.9 | 314.5 | 241.2 | 287.4 | 265.1 | 311.7 | 311.6 | 297.0 | 320.4 | 298.6 | 45 | 48 | 45 | 47 | 47 | 47 | 43 | 46 | 50 | 49 | 49 | 47 |
| HLA class I histocompatibility antigen, A-2 alpha chain | 29.2 | 28.4 | 26.9 | 28.1 | 22.1 | 27.8 | 30.9 | 29.4 | 30.6 | 32.5 | 32.5 | 33.8 | 8 | 7 | 7 | 8 | 7 | 7 | 7 | 7 | 7 | 8 | 7 | 9 |
| Multifunctional protein ADE2 | 45.5 | 39.2 | 39.5 | 39.8 | 48.7 | 58.3 | 56.9 | 53.3 | 51.5 | 56.9 | 61.1 | 60.3 | 9 | 11 | 10 | 10 | 12 | 12 | 10 | 12 | 12 | 12 | 12 | 11 |
| Nicotinate-nucleotide pyrophosphorylase | 36.8 | 32.8 | 32.0 | 30.1 | 21.0 | 27.6 | 24.9 | 25.5 | 30.1 | 30.1 | 29.1 | 26.8 | 8 | 7 | 5 | 8 | 7 | 7 | 6 | 7 | 6 | 8 | 6 | 8 |
| Protein disulfide-isomerase A4 | 358.3 | 334.4 | 351.1 | 335.6 | 210.1 | 309.9 | 289.6 | 253.2 | 342.3 | 304.5 | 303.0 | 284.2 | 37 | 28 | 32 | 37 | 29 | 33 | 27 | 30 | 33 | 31 | 30 | 34 |
| Protein disulfide-isomerase A6 | 117.1 | 112.6 | 114.2 | 132.9 | 88.1 | 102.9 | 126.3 | 104.1 | 116.2 | 117.2 | 135.8 | 118.3 | 18 | 18 | 18 | 18 | 19 | 18 | 17 | 19 | 18 | 18 | 18 | 19 |
| Ras-related protein Rab-10 | 15.6 | 16.3 | 16.4 | 17.8 | 12.9 | 14.9 | 16.5 | 16.9 | 16.0 | 19.5 | 17.3 | 15.5 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 1 | 1 |
| Transitional endoplasmic reticulum ATPase | 152.8 | 165.9 | 173.3 | 157.0 | 109.7 | 140.5 | 141.8 | 150.6 | 170.7 | 159.2 | 164.2 | 149.4 | 32 | 30 | 37 | 33 | 29 | 34 | 29 | 33 | 33 | 31 | 31 | 32 |

The aim of this project was to identify proteomic changes associated with lipid accumulation in HepG2 cells and the possible ameliorative effects of the DGAT inhibitor PF04 on the process. The magnitude of the fold changes observed in this dataset were not statistically significant. Given the experimental design, the variance observed within each group is likely to have a considerable impact on the statistical significance assigned to any change in protein abundance. The four biological replicates representing a treatment group were therefore pooled into one sample in an effort to normalize the variance within a group. Dataset LC4 was generated from the pooled samples of biological replicates as described in the next section.

5.2.2. Dataset LC4

5.2.2.1. Protein identification

Dataset LC4 was generated from the same biological replicates as dataset LC3, with the difference between the two datasets being that the four biological replicates representing a group were pooled after cell lysis. Each pooled sample was analysed using five LC-MS/MS runs. The total number of proteins identified for the three treatment groups using input from all five runs was 970, 1,008 and 1,034 for control, FA-treated and inhibitor-treated samples, respectively (Table 21 in appendix). About 74 % of the proteins identified in this dataset were present in all three samples (Figure 53A). Analysis of individual technical replicates showed that the average number of proteins identified in the technical replicates was 763, 827 and 836 for control, FA-treated and inhibitor-treated samples, respectively (Figure 53B). These proteins were identified in PD with high confidence (FDR < 0.01) and with at least 2 peptides per protein. The average number of peptides identified in a technical replicate for control, FA-treated and inhibitor-treated groups was 5,805, 6,271 and 6,461, respectively. These identifications were based on an average of 16,462, 17,389 and 18,184

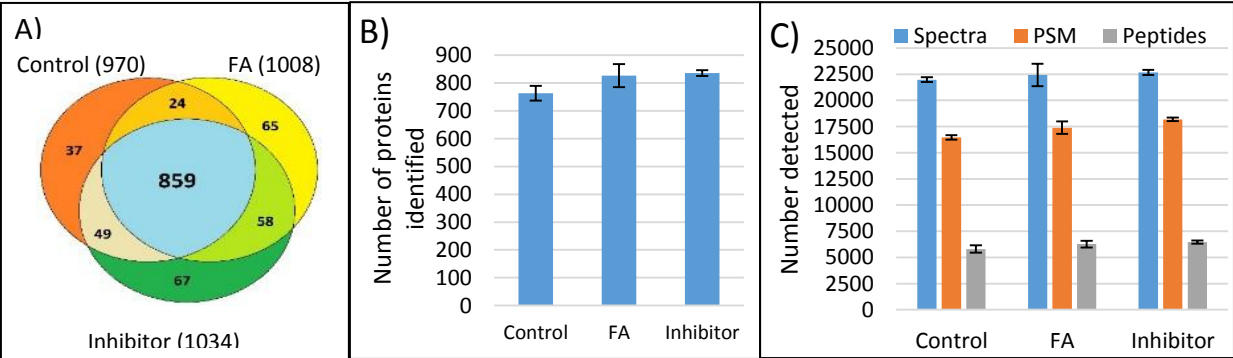


Figure 53: Protein identification for dataset LC4 in Proteome Discoverer. A) Comparison of proteins identified in dataset LC4. **B)** Average number of proteins identified for the technical replicates for the groups indicated. **C)** Average number of spectra, peptide-spectra matches (PSMs) and peptides detected for each group in PD.

peptide-spectra matches determined for the 21,976, 22,436 and 22,684 input spectra detected for control, FA-treated and inhibitor-treated groups, respectively (Figure 53C).

5.2.2.2. Label-free quantification using spectral counts

The files generated in PD for this dataset were organized in Scaffold as they were for dataset LC1 - five LC-MS/MS runs represented one pooled sample and each group was represented by one pooled sample. The number of proteins identified in Scaffold for this dataset was 928 (Table 21 in Appendix). No decoys were identified in the list. Peptide identifications were accepted if they could be established at greater than 99.0 % probability by the Peptide Prophet algorithm, and protein identifications were accepted if they could be established at greater than 99.9 % probability and contained at least 2 identified peptides. About 17 % of the detected spectra in each group could be assigned exclusively to a single peptide. Of the 928 proteins identified in Scaffold, 904 proteins were found in all three groups and 20 proteins were identified in at least two of the three groups (Figure 54A). The number of peptides common to all three groups was 4,839 and this corresponded to 5,399 unique spectra that were common to all three groups (Figure 54B and 54C). Since the dataset contains one pooled sample per treatment group, the QVs of the groups were compared using Fisher's exact test with multiple testing correction by Hochberg-Benjamini method. This test required a comparison of two groups at a time and so three comparisons were made – FA-treated vs. control; inhibitor-treated vs. FA-treated and inhibitor-treated vs. control. At a significance level of 95 %, 28 proteins showed statistically significant changes with $p \leq 1.56 \times 10^{-3}$ for the comparison between control and FA-treated groups, only 1 protein with $p \leq 1.1 \times 10^{-4}$ was considered significant for the comparison between the FA-treated and the inhibitor-treated sample. For a comparison of control and inhibitor-treated groups, five proteins with $p \leq 3.2 \times 10^{-4}$ were considered statistically significant (Table 10)

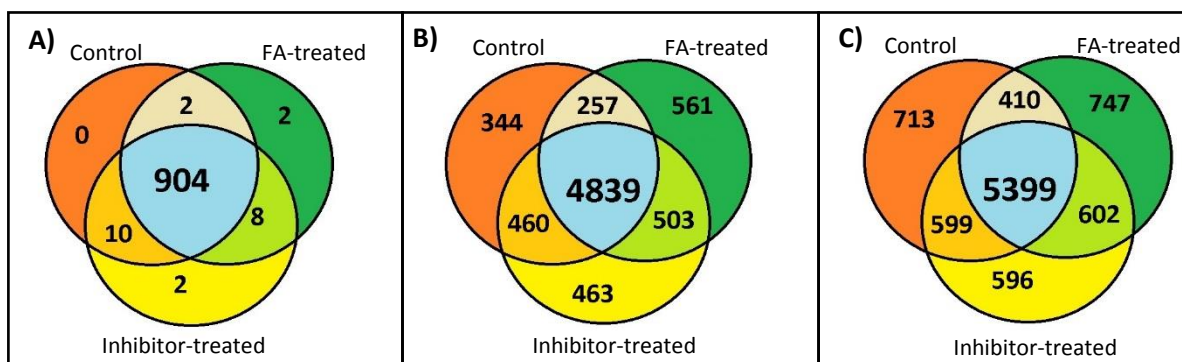


Figure 54: Identifications of proteins, peptides and MS/MS spectra in different treatment groups in Scaffold. A) Proteins identified in each treatment group. B) Total unique peptide counts for all proteins in the dataset. C) The total number of unique spectra detected for all the peptides in the dataset.

Table 10: Proteins with significant changes in abundance in dataset LC4. The table lists proteins with statistically significant changes in abundance as determined by Fisher's exact test. Two groups were compared at a time. The 'fold change' represents the ratio of the QV as stated in the column head; for example, 'F: C' is the ratio of QV in FA-treated group (F) to that in Control group (C). The inhibitor-treated group is represented by 'I'. 'Exclusive unique peptides' indicates the number of peptides that are unique and exclusive to the indicated protein. CV attributed to a protein reflects the technical variance among the LC-MS/MS runs representing each group.

| Protein name | Gene name | Fold change | | | P value | | | CV (%) | | | QV | | | Peptide count | | |
|--|-----------|-------------|------|------|---------|---------|--------|--------|-----|----|--------|--------|--------|---------------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I | C | F | I |
| Proteins significantly changed in FA-treated/Control | | | | | | | | | | | | | | | | |
| U5 small nuclear ribonucleoprotein 200 kDa helicase | SNRNP200 | 4.2 | 0.50 | 2.2 | 0.00012 | 0.02 | 0.056 | 22 | 50 | 25 | 7.34 | 30.90 | 16.15 | 3 | 9 | 5 |
| Splicing factor 3B subunit 3 | SF3B3 | 9.2 | 0.40 | 3.6 | 0.00018 | 0.02 | 0.076 | 11 | 25 | 32 | 2.09 | 19.24 | 7.55 | 1 | 4 | 3 |
| Clathrin heavy chain 1 | CLTC | 1.6 | 0.91 | 1.5 | 0.00026 | 0.23 | 0.0033 | 28 | 14 | 16 | 80.80 | 132.44 | 120.03 | 20 | 26 | 29 |
| Isoform 2 of Reticulocalbin-2 | RCN2 | 0.10 | 3 | 0.40 | 0.00026 | 0.085 | 0.023 | 26 | 64 | 26 | 19.71 | 2.85 | 8.59 | 5 | 1 | 2 |
| 60 kDa heat shock protein, mitochondrial | HSPD1 | 0.77 | 1.1 | 1 | 0.00033 | 0.0048 | 0.19 | 10 | 8 | 7 | 837.77 | 702.31 | 799.34 | 49 | 49 | 46 |
| ATP-dependent RNA helicase A | DHX9 | 2.4 | 0.40 | 0.91 | 0.00038 | 0.00013 | 0.49 | 23 | 60 | 38 | 21.21 | 50.96 | 19.94 | 6 | 14 | 7 |
| Cell division control protein 42 homolog | CDC42 | 4.5 | 0.30 | 1.4 | 0.00048 | 0.0024 | 0.36 | 5 | 6 | 59 | 5.32 | 23.88 | 7.70 | 1 | 2 | 2 |
| Valine--tRNA ligase | VARS | 4.5 | 0.71 | 3.2 | 0.00048 | 0.18 | 0.011 | ND | 18 | 31 | 5.33 | 24.16 | 17.09 | 2 | 7 | 5 |
| Fatty acid synthase | FASN | 1.3 | 0.91 | 1.1 | 0.0005 | 0.019 | 0.11 | 35 | 13 | 6 | 333.19 | 423.77 | 366.34 | 67 | 82 | 73 |
| Nuclease-sensitive element-binding protein 1 | YBX1 | 0.30 | 2.1 | 0.71 | 0.00062 | 0.024 | 0.12 | 140 | 54 | 17 | 34.62 | 11.98 | 24.90 | 9 | 4 | 5 |
| Leucine-rich PPR motif-containing protein, mitochondrial | LRPPRC | 1.4 | 0.77 | 1.2 | 0.00065 | 0.04 | 0.069 | 25 | 13 | 7 | 159.23 | 222.69 | 187.18 | 37 | 49 | 45 |
| ATP-dependent 6-phosphofructokinase, liver type | PFKL | 11 | 0.59 | 6.3 | 0.00072 | 0.12 | 0.028 | 42 | 150 | 70 | 1.22 | 13.83 | 7.66 | 1 | 6 | 3 |
| Multifunctional protein ADE2 | PAICS | 2 | 0.71 | 1.4 | 0.0008 | 0.042 | 0.08 | 8 | 8 | 12 | 34.02 | 66.98 | 47.69 | 6 | 10 | 8 |
| Bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase (cyclizing) | DAK | 2 | 0.71 | 1.3 | 0.00095 | 0.018 | 0.16 | 17 | 8 | 15 | 30.62 | 61.46 | 40.05 | 6 | 15 | 11 |
| ATP synthase subunit beta | ATP5B | 0.77 | 1.1 | 0.77 | 0.0011 | 0.24 | 0.0091 | 9 | 31 | 5 | 278.73 | 210.57 | 225.23 | 25 | 25 | 24 |
| DnaJ homolog subfamily B member 11 | DNAJB11 | 0.08 | 11 | 0.77 | 0.0012 | 0.004 | 0.4 | 68 | 31 | 36 | 12.51 | 0.94 | 10.38 | 4 | 1 | 3 |
| Glycine--tRNA ligase | GARS | 2.3 | 0.59 | 1.4 | 0.0012 | 0.027 | 0.14 | 8 | 7 | 33 | 20.21 | 45.69 | 28.44 | 5 | 10 | 7 |

Table 10 continued...

| Protein name | Gene name | Fold change | | | P value | | | CV (%) | | | QV | | | Peptide count | | |
|---|-----------|-------------|------|------|----------|----------|----------|--------|-----|----|--------|--------|--------|---------------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I | C | F | I |
| Cullin-associated NEDD8-dissociated protein 1 | CAND1 | 13 | 0.50 | 6.5 | 0.0013 | 0.11 | 0.048 | 21 | 6 | 80 | 1.01 | 13.01 | 6.57 | 1 | 3 | 2 |
| Glutamine--fructose-6-phosphate aminotransferase [isomerizing] 1 | GFPT1 | 2.9 | 0.59 | 1.6 | 0.0014 | 0.03 | 0.15 | 220 | 26 | 18 | 10.86 | 30.99 | 17.19 | 5 | 7 | 4 |
| Citrate synthase, mitochondrial | CS | 2.5 | 0.71 | 1.8 | < 0.0001 | 0.053 | 0.015 | 24 | 24 | 28 | 23.22 | 58.64 | 41.92 | 6 | 11 | 10 |
| Isoform 3 of Calumenin | CALU | 0.20 | 2.2 | 0.50 | < 0.0001 | 0.014 | 0.0088 | 220 | 53 | 22 | 45.66 | 10.97 | 24.61 | 8 | 4 | 6 |
| Bifunctional glutamate/proline--tRNA ligase | EPRS | 4.1 | 0.59 | 2.6 | < 0.0001 | 0.05 | 0.017 | 12 | 7 | 26 | 8.19 | 33.32 | 20.98 | 3 | 12 | 7 |
| Protein disulfide-isomerase A3 | PDIA3 | 0.71 | 1.3 | 0.91 | < 0.0001 | 0.0018 | 0.043 | 23 | 23 | 5 | 322.57 | 215.24 | 279.53 | 34 | 32 | 36 |
| Proteins significantly changed in both FA-treated/control and Inhibitor-treated/Control | | | | | | | | | | | | | | | | |
| Polyadenylate-binding protein 4 | PABPC4 | 8.1 | 1.1 | 8.7 | < 0.0001 | 0.41 | < 0.0001 | 220 | 52 | 23 | 4.89 | 39.44 | 42.73 | 1 | 2 | 2 |
| 78 kDa glucose-regulated protein | HSPA5 | 0.71 | 1.1 | 0.77 | < 0.0001 | 0.28 | 0.00022 | 34 | 220 | 10 | 393.96 | 287.60 | 303.06 | 41 | 39 | 40 |
| ADP-ribosylation factor 1 | ARF1 | 13.0 | 1.5 | 20.0 | 0.0013 | 0.14 | < 0.0001 | 43 | 24 | NA | 1.00 | 13.07 | 20.04 | 1 | 5 | 6 |
| Coatomer subunit alpha | COPA | 4.7 | 0.77 | 4.0 | < 0.0001 | 0.25 | < 0.0001 | 24 | 17 | 11 | 8.62 | 40.50 | 34.23 | 2 | 8 | 6 |
| Proteins significantly changed proteins in Inhibitor-treated/Control | | | | | | | | | | | | | | | | |
| Perilipin-2 | PLIN2 | INF | 2.7 | INF | 0.036 | 0.04 | 0.00013 | ND | 68 | 27 | 0 | 4.93 | 13.23 | 0 | 2 | 4 |
| Protein significantly changed in Inhibitor-treated/FA-treated group | | | | | | | | | | | | | | | | |
| Ras-related protein Rab-6A | RAB6A | INF | 0.0 | INF | < 0.0001 | < 0.0001 | 1 | 55 | 28 | ND | 0.00 | 18.82 | 0.00 | 0 | 2 | 0 |

5.2.2.3. Variance associated with data generated from the pooled group of LC4

The protein list generated in PD for each LC-MS/MS run was compared to the other lists for the same pooled sample (Figure 55). The five LC-MS/MS runs of the control sample had 613 identified proteins common among them. This number represents an average of 80 % of the proteins identified in five technical replicates of the control sample. Similarly, the technical replicates of the 673 and 697 proteins were identified in all technical replicates of the FA-treated and inhibitor-treated samples respectively (Table 21 in Appendix). This represented 82 % and 83 % of proteins in the individual protein lists. Variance in the spectral counts assigned to a protein were assessed in Scaffold by loading the technical replicates as biosamples to calculate the CV.

Of the 782 proteins identified for the control group in Scaffold, only 71 proteins showed a CV ≤ 10 %. The number of proteins identified in Scaffold for FA-treated and inhibitor-treated groups was 828 and 847, respectively and of these, 97 and 110 proteins showed a CV ≤ 10 %

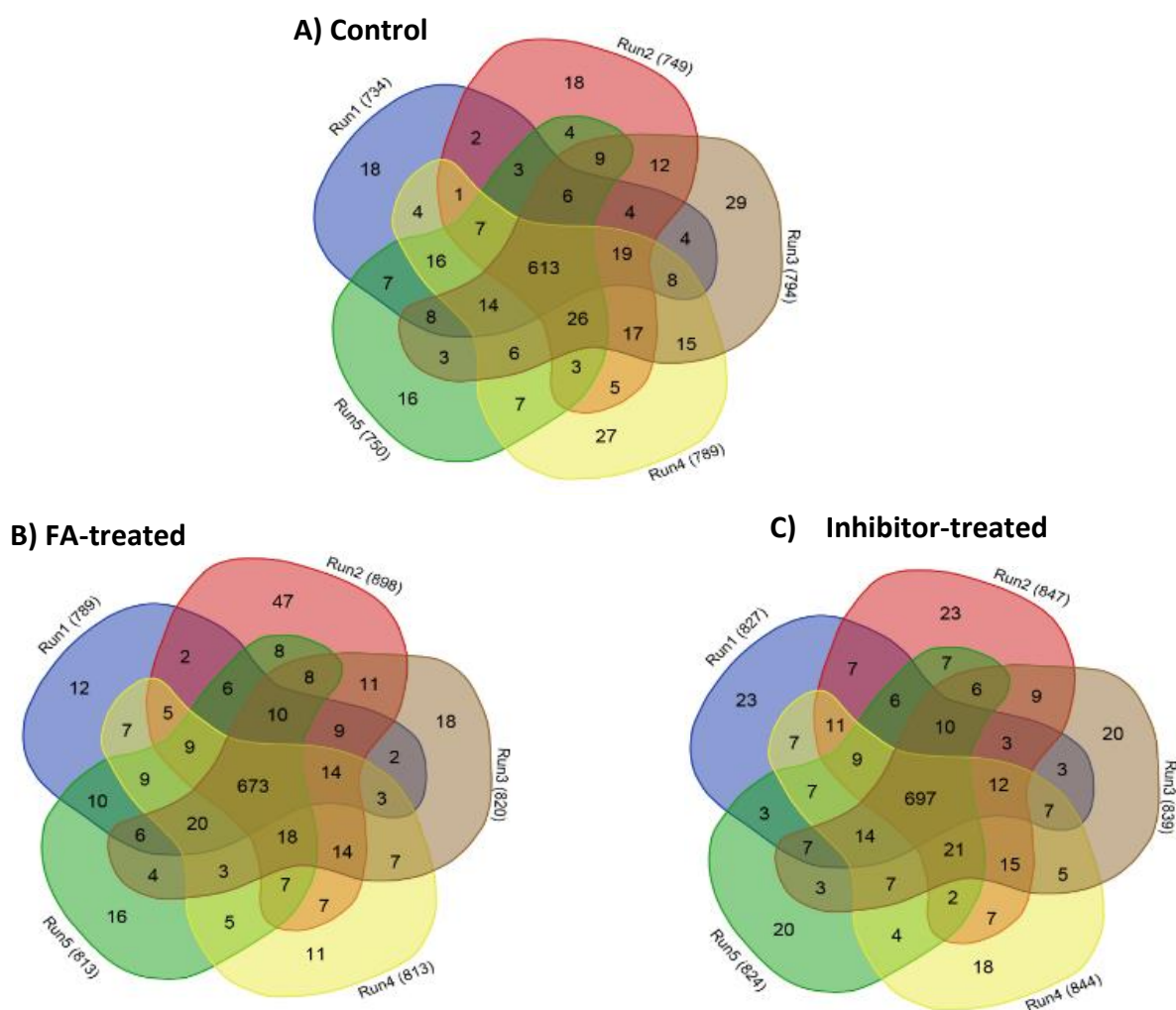


Figure 55: Comparison of proteins identified in each LC-MS/MS run within a group analysed for dataset LC4. The list of proteins identified in PD for each LC-MS/MS run was compared using the Venn diagram tool provided by Bioinformatics & Evolutionary Genomics (<http://bioinformatics.psb.ugent.be/webtools/Venn>)

for the two groups. Most proteins had a CV of more than 10 % (Figure 56). Since the QV assigned to a protein depends on the cumulative data of all five technical replicates, the variance among the technical replicates could have skewed the total spectral count assigned to the protein. This in turn would affect the statistical value assigned to the fold changes of a given protein. Therefore, in addition to the proteins that showed statistically significant changes in QV, all proteins that showed a change in QV with $p \leq 0.05$ before multiple test corrections were put through a functional analysis. References made hereafter to ‘identified proteins’ refers to this sub-set of proteins in dataset LC4.

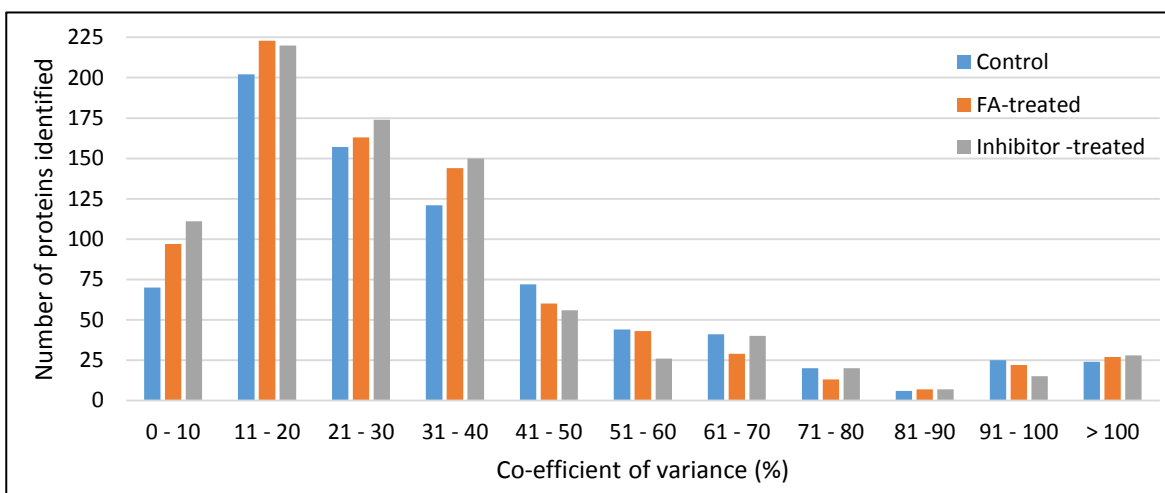


Figure 56: Distribution of proteins with respect to the co-efficient of variance for LC4. The CV of a protein for the five technical replicates of a pooled group was calculated, and the numbers of proteins that lie within a range of CVs is indicated.

5.2.3. Functional analysis of proteins identified in dataset LC4.

The biological relevance of the identified proteins was evaluated through enrichment analysis of protein in GO groups and KEGG pathway databases using the tools hosted by WebGestalt [243]. Of the 928 proteins analysed in dataset LC4, 216 proteins showed a change in abundance with a $p \leq 0.05$ and fold change ≥ 1.3 in at least one comparison (Table 11). These proteins were mapped to their corresponding SwissProt identifiers and uploaded onto WebGestalt to generate the results of the functional analysis described in this section (Table 21 in Appendix).

Table 11: Proteins of dataset LC4 used for functional analysis. Proteins that showed a fold change more than or equal to 1.3 in at least one comparison with $p \leq 0.05$ were selected for functional analysis. The treatment groups are indicated as 'C' for control, 'F' for FA-treated and 'I' for inhibitor-treated. The fold change indicated is the ratio of the QV for the protein in the comparison as indicated. For example, 'F:C' is a ratio of QV of the FA-treated to the control. 'INF' indicates absence of protein in the control sample resulting in the denominator of FC being zero and the ratio being infinite.

| Protein name | Gene name | Fold change | | | P value | | | QV | | | Peptide | | |
|---|-----------|-------------|------|------|----------|--------|----------|-------|-------|-------|---------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I |
| Isoform 2 of Alanine--tRNA ligase, cytoplasmic | AARS | 2.9 | 0.77 | 2.4 | 0.0023 | 0.28 | 0.015 | 9.6 | 27.8 | 22.8 | 4 | 9 | 7 |
| Isoform 2 of ATP-binding cassette sub-family F member 1 | ABCF1 | 7.7 | 1 | 7.5 | 0.024 | 0.56 | 0.028 | 1.0 | 7.8 | 7.6 | 1 | 4 | 3 |
| 3-ketoacyl-CoA thiolase, peroxisomal | ACAA1 | 7.9 | 0.91 | 7.5 | 0.024 | 0.56 | 0.028 | 1.0 | 8.0 | 7.6 | 1 | 4 | 2 |
| Isoform 2 of Very long-chain specific acyl-CoA dehydrogenase, mitochondrial | ACADVL | 2.1 | 0.77 | 1.7 | 0.037 | 0.27 | 0.15 | 10.3 | 21.8 | 17.1 | 3 | 6 | 4 |
| ATP-citrate synthase | ACLY | 1.7 | 0.77 | 1.3 | 0.0037 | 0.12 | 0.073 | 46.9 | 77.4 | 63.0 | 10 | 18 | 15 |
| Isoform Short of Long-chain-fatty-acid--CoA ligase 4 | ACSL4 | 1.9 | 0.71 | 1.2 | 0.017 | 0.072 | 0.29 | 19.2 | 35.9 | 23.7 | 4 | 6 | 6 |
| Isoform 2 of Long-chain-fatty-acid--CoA ligase 5 | ACSL5 | 1.7 | 0.50 | 0.77 | 0.093 | 0.021 | 0.32 | 13.7 | 22.9 | 10.5 | 5 | 7 | 5 |
| Acyl-CoA synthetase short-chain family member 3, mitochondrial | ACSS3 | 3.7 | 0.50 | 1.8 | 0.0032 | 0.038 | 0.2 | 5.3 | 19.9 | 9.4 | 3 | 6 | 5 |
| Actin-related protein 2 | ACTR2 | 2.9 | 0.71 | 1.9 | 0.011 | 0.16 | 0.13 | 6.5 | 19.0 | 12.4 | 2 | 4 | 3 |
| Angiotensinogen | AGT | 0.50 | 3 | 1.5 | 0.044 | 0.0013 | 0.14 | 18.9 | 9.0 | 27.5 | 6 | 4 | 7 |
| Alpha-2-HS-glycoprotein | AHSG | 0.40 | 2.7 | 1.1 | 0.017 | 0.0039 | 0.37 | 23.8 | 9.8 | 26.7 | 6 | 5 | 7 |
| Serum albumin | ALB | 0.77 | 1 | 0.77 | 0.0095 | 0.41 | 0.018 | 281.3 | 227.2 | 232.8 | 38 | 38 | 40 |
| Aldehyde dehydrogenase, mitochondrial | ALDH2 | 1.1 | 0.77 | 0.77 | 0.27 | 0.043 | 0.16 | 80.3 | 89.9 | 67.6 | 15 | 15 | 15 |
| Isoform 2 of Fatty aldehyde dehydrogenase | ALDH3A2 | 5.4 | 0.50 | 2.5 | 0.0088 | 0.1 | 0.18 | 2.2 | 12.1 | 5.6 | 2 | 3 | 2 |
| Protein AMBP | AMBP | 0.50 | 1.9 | 0.91 | 0.017 | 0.037 | 0.4 | 27.2 | 12.8 | 24.7 | 6 | 3 | 5 |
| Isoform B of AP-2 complex subunit alpha-1 | AP2A1 | INF | 0.30 | INF | 0.005 | 0.099 | 0.15 | 0.0 | 8.1 | 2.8 | 0 | 2 | 1 |
| Isoform 2 of AP-2 complex subunit beta | AP2B1 | 1.9 | 1.9 | 3.6 | 0.22 | 0.093 | 0.011 | 4.3 | 8.1 | 15.3 | 2 | 3 | 2 |
| Apolipoprotein B-100 | APOB | INF | 6 | INF | 0.52 | 0.07 | 0.021 | 0.0 | 0.9 | 5.7 | 0 | 1 | 2 |
| Apolipoprotein E | APOE | 0.59 | 1.6 | 1 | 0.017 | 0.014 | 0.53 | 55.6 | 33.5 | 55.3 | 10 | 10 | 10 |
| Beta-2-glycoprotein 1 | APOH | 0.50 | 2 | 0.91 | 0.0034 | 0.0081 | 0.4 | 42.2 | 19.8 | 38.9 | 8 | 6 | 7 |
| ADP-ribosylation factor 1 | ARF1 | 13 | 1.5 | 20 | 0.0013 | 0.14 | < 0.0001 | 1.0 | 13.1 | 20.0 | 1 | 5 | 7 |
| ADP-ribosylation factor 4 | ARF4 | 2.6 | 0.91 | 2.4 | 0.014 | 0.44 | 0.028 | 7.6 | 19.9 | 18.0 | 3 | 3 | 3 |
| ATP synthase subunit alpha, mitochondrial | ATP5A1 | 0.91 | 0.91 | 0.77 | 0.1 | 0.19 | 0.014 | 172.0 | 147.7 | 132.2 | 28 | 25 | 25 |
| ATP synthase subunit beta, mitochondrial | ATP5B | 0.77 | 1.1 | 0.77 | 0.0011 | 0.24 | 0.0091 | 278.7 | 210.6 | 225.2 | 25 | 25 | 24 |
| Isoform 3 of Calcyclin-binding protein | CACYBP | INF | INF | INF | 1 | 0.018 | 0.021 | 0.0 | 0.0 | 5.7 | 0 | 0 | 3 |
| Calreticulin | CALR | 0.77 | 1 | 0.77 | 0.01 | 0.36 | 0.003 | 249.3 | 199.1 | 191.1 | 19 | 21 | 21 |
| Isoform 3 of Calumenin | CALU | 0.20 | 2.2 | 0.50 | < 0.0001 | 0.014 | 0.0088 | 45.7 | 11.0 | 24.6 | 8 | 4 | 6 |
| Cullin-associated NEDD8-dissociated protein 1 | CAND1 | 13 | 0.50 | 6.5 | 0.0013 | 0.11 | 0.048 | 1.0 | 13.0 | 6.6 | 1 | 3 | 2 |

Table 11 continued...

| Protein name | Gene name | Fold change | | | P value | | | QV | | | Peptide | | |
|---|-----------|-------------|------|------|----------|---------|----------|------|-------|------|---------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I |
| Isoform 2 of Calnexin | CANX | 1.5 | 0.77 | 1.2 | 0.0069 | 0.055 | 0.2 | 62.3 | 94.7 | 73.2 | 11 | 14 | 12 |
| Isoform 2 of Cell cycle and apoptosis regulator protein 2 | CCAR2 | 2 | 0.50 | 1.1 | 0.013 | 0.018 | 0.49 | 16.8 | 34.0 | 18.1 | 4 | 9 | 6 |
| Cell division control protein 42 homolog | CDC42 | 4.5 | 0.30 | 1.4 | 0.00048 | 0.0024 | 0.36 | 5.3 | 23.9 | 7.7 | 1 | 2 | 2 |
| Cocaine esterase | CES2 | 6.7 | 0.91 | 6.3 | 0.003 | 0.53 | 0.0039 | 2.1 | 14.0 | 13.1 | 2 | 4 | 3 |
| Clathrin heavy chain 1 | CLTC | 1.6 | 0.91 | 1.5 | 0.00026 | 0.23 | 0.0033 | 80.8 | 132.4 | 120 | 20 | 26 | 29 |
| Clustered mitochondria protein homolog | CLUH | INF | 0.20 | INF | 0.036 | 0.099 | 0.53 | 0.0 | 5.0 | 0.9 | 0 | 2 | 1 |
| Catechol O-methyltransferase | COMT | 0.91 | 1.60 | 1.4 | 0.42 | 0.049 | 0.1 | 24.4 | 22.0 | 35.1 | 7 | 6 | 8 |
| Coatomer subunit alpha | COPA | 4.7 | 0.77 | 4.00 | < 0.0001 | 0.25 | < 0.0001 | 8.6 | 40.5 | 34.2 | 2 | 8 | 6 |
| Coatomer subunit gamma-1 | COPG1 | 2.4 | 0.91 | 2.2 | 0.0019 | 0.42 | 0.0045 | 16.8 | 39.9 | 37.0 | 8 | 11 | 10 |
| Oxygen-dependent coproporphyrinogen-III oxidase, mitochondrial | CPOX | 0.09 | 7.20 | 0.71 | 0.0023 | 0.023 | 0.24 | 11.7 | 1.1 | 7.6 | 5 | 1 | 4 |
| Serine/threonine-protein phosphatase CPPED1 | CPPED1 | 0.40 | 1.2 | 0.50 | 0.05 | 0.44 | 0.1 | 15.9 | 6.9 | 8.5 | 3 | 2 | 3 |
| Citrate synthase, mitochondrial | CS | 2.50 | 0.71 | 1.8 | < 0.0001 | 0.053 | 0.015 | 23.2 | 58.6 | 41.9 | 6 | 11 | 10 |
| Isoform 4 of Exportin-2 | CSE1L | 1.90 | 0.77 | 1.50 | 0.04 | 0.23 | 0.19 | 13.7 | 25.9 | 20.0 | 3 | 6 | 5 |
| Casein kinase II subunit alpha | CSNK2A1 | 2.8 | 0.91 | 2.60 | 0.042 | 0.53 | 0.052 | 5.1 | 14.0 | 13.2 | 3 | 4 | 4 |
| Cystatin-B | CSTB | 0.20 | 5.30 | 1.3 | 0.045 | 0.014 | 0.41 | 8.4 | 2.0 | 10.5 | 2 | 1 | 2 |
| Catenin beta-1 | CTNNB1 | 2.30 | 0.77 | 1.80 | 0.002 | 0.22 | 0.02 | 19.2 | 43.2 | 35.2 | 4 | 9 | 8 |
| Dipeptidyl peptidase 1 | CTSC | 0.40 | 2.70 | 1.1 | 0.085 | 0.04 | 0.45 | 11.7 | 5.0 | 13.4 | 3 | 1 | 3 |
| Cathepsin D | CTSD | 0.71 | 1.2 | 0.77 | 0.043 | 0.27 | 0.15 | 60.1 | 41.0 | 47.6 | 10 | 10 | 11 |
| Isoform A of Protein CutA | CUTA | 0.50 | 1.6 | 0.77 | 0.042 | 0.13 | 0.31 | 23.4 | 12.0 | 19.0 | 3 | 2 | 3 |
| Isoform 2 of NADH-cytochrome b5 reductase 3 | CYB5R3 | 1.90 | 1 | 1.9 | 0.017 | 0.53 | 0.015 | 18.8 | 35.6 | 36.0 | 4 | 5 | 5 |
| Triokinase/FMN cyclase | DAK | 2.00 | 0.71 | 1.30 | 0.00095 | 0.018 | 0.16 | 30.6 | 61.5 | 40.0 | 6 | 15 | 11 |
| Isoform 5 of Acyl-CoA-binding protein | DBI | 0.20 | 3.4 | 0.77 | 0.045 | 0.1 | 0.42 | 8.3 | 1.9 | 6.6 | 3 | 2 | 4 |
| ATP-dependent RNA helicase DDX42 | DDX42 | INF | 0.59 | INF | 0.019 | 0.35 | 0.077 | 0.0 | 6.0 | 3.9 | 0 | 1 | 2 |
| ATP-dependent RNA helicase A | DHX9 | 2.40 | 0.40 | 0.91 | 0.00038 | 0.00013 | 0.49 | 21.2 | 51.0 | 19.9 | 6 | 14 | 7 |
| Dihydrolipoyl dehydrogenase, mitochondrial | DLD | 0.59 | 0.91 | 0.59 | 0.023 | 0.38 | 0.0071 | 55.0 | 34.8 | 31.5 | 12 | 7 | 8 |
| DnaJ homolog subfamily B member 11 | DNAJB11 | 0.08 | 11 | 0.77 | 0.0012 | 0.004 | 0.4 | 12.5 | 0.9 | 10.4 | 4 | 1 | 3 |
| Isoform 2 of Dihydropyrimidinase-related protein 2 | DPYSL2 | 1.10 | 1.6 | 1.70 | 0.45 | 0.014 | 0.0081 | 31.9 | 33.7 | 55.1 | 9 | 8 | 13 |
| Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial | DUT | 0.30 | 2.2 | 0.77 | 0.023 | 0.086 | 0.32 | 14.9 | 5.1 | 11.4 | 4 | 2 | 3 |
| 116 kDa U5 small nuclear ribonucleoprotein component | EFTUD2 | 2.5 | 0.50 | 1.4 | 0.002 | 0.019 | 0.23 | 13.7 | 34.8 | 19.0 | 4 | 7 | 4 |
| Eukaryotic translation initiation factor 2 subunit 3 | EIF2S3 | INF | 1.20 | INF | 0.036 | 0.53 | 0.021 | 0.0 | 4.9 | 5.7 | 0 | 2 | 2 |
| Eukaryotic translation initiation factor 3 subunit A | EIF3A | 2.5 | 0.77 | 2 | 0.011 | 0.27 | 0.056 | 9.6 | 24.0 | 19.0 | 5 | 9 | 7 |

Table 11 continued...

| Protein name | Gene name | Fold change | | | P value | | | QV | | | Peptide | | |
|---|-----------|-------------|------|------|----------|---------|--------|-------|-------|-------|---------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I |
| Eukaryotic translation initiation factor 3 subunit I | EIF3I | 0.71 | 2.20 | 1.6 | 0.36 | 0.034 | 0.11 | 11.2 | 8.1 | 18.2 | 5 | 3 | 7 |
| Eukaryotic initiation factor 4A-III | EIF4A3 | 3.8 | 0.91 | 3.4 | 0.023 | 0.46 | 0.044 | 3.0 | 11.6 | 10.3 | 0 | 2 | 1 |
| Eukaryotic translation initiation factor 4 gamma 1 | EIF4G1 | 2 | 0.71 | 1.3 | 0.042 | 0.13 | 0.31 | 11.6 | 23.0 | 15.3 | 3 | 8 | 4 |
| Echinoderm microtubule-associated protein-like 4 | EML4 | 2 | 0.50 | 0.91 | 0.013 | 0.0044 | 0.44 | 18.9 | 37.2 | 17.2 | 5 | 11 | 6 |
| Bifunctional glutamate/proline--tRNA ligase | EPRS | 4.1 | 0.59 | 2.6 | < 0.0001 | 0.05 | 0.017 | 8.2 | 33.3 | 21.0 | 3 | 12 | 7 |
| Protein FAM3C | FAM3C | INF | 0.20 | INF | 0.036 | 0.099 | 0.53 | 0.0 | 4.9 | 0.9 | 0 | 2 | 1 |
| Phenylalanine--tRNA ligase beta subunit | FARSB | 8.8 | 0.50 | 4.70 | 0.014 | 0.19 | 0.14 | 1.0 | 8.9 | 4.7 | 1 | 3 | 2 |
| Fatty acid synthase | FASN | 1.3 | 0.91 | 1.1 | 0.0005 | 0.019 | 0.11 | 333.2 | 423.8 | 366.3 | 67 | 82 | 73 |
| Squalene synthase | FDFT1 | 2.30 | 0.59 | 1.40 | 0.011 | 0.089 | 0.2 | 11.9 | 27.0 | 17.1 | 4 | 7 | 6 |
| Peptidyl-prolyl cis-trans isomerase FKBP1A | FKBP1A | 0.40 | 3 | 1.10 | 0.056 | 0.017 | 0.39 | 13.2 | 5.0 | 15.2 | 5 | 1 | 4 |
| Filamin-A | FLNA | 1.20 | 1.00 | 1.3 | 0.036 | 0.5 | 0.03 | 119.6 | 149.4 | 150.3 | 29 | 36 | 33 |
| Isoform 15 of Fibronectin | FN1 | 0.59 | 1.80 | 1.10 | 0.013 | 0.00084 | 0.23 | 69.5 | 43.8 | 79.0 | 15 | 12 | 18 |
| Formimidoyltransferase-cyclodeaminase | FTCD | INF | 0.59 | INF | 0.0026 | 0.27 | 0.021 | 0.0 | 9.0 | 5.8 | 0 | 4 | 2 |
| Isoform 3 of Glucose-6-phosphate 1-dehydrogenase | G6PD | 9.00 | 0.50 | 4.7 | 0.0043 | 0.14 | 0.082 | 1.2 | 11.0 | 5.7 | 1 | 6 | 3 |
| Lysosomal alpha-glucosidase | GAA | 0.50 | 1.90 | 0.91 | 0.003 | 0.013 | 0.32 | 41.6 | 18.8 | 36.2 | 9 | 5 | 8 |
| Glycine--tRNA ligase | GARS | 2.30 | 0.59 | 1.4 | 0.0012 | 0.027 | 0.14 | 20.2 | 45.7 | 28.4 | 5 | 10 | 7 |
| Trifunctional purine biosynthetic protein adenosine-3 | GART | 2.30 | 0.59 | 1.30 | 0.044 | 0.1 | 0.4 | 7.4 | 16.7 | 9.5 | 4 | 6 | 4 |
| Translational activator GCN1 | GCN1L1 | 11.00 | 0.30 | 3.70 | 0.0043 | 0.049 | 0.22 | 1.0 | 11.1 | 3.7 | 1 | 4 | 2 |
| Glutamine--fructose-6-phosphate aminotransferase 1 | GFPT1 | 2.90 | 0.59 | 1.6 | 0.0014 | 0.03 | 0.15 | 10.9 | 31.0 | 17.2 | 5 | 7 | 4 |
| Isoform 2 of Golgi apparatus protein 1 | GLG1 | 1.40 | 2.30 | 3.20 | 0.35 | 0.029 | 0.0066 | 6.4 | 8.9 | 20.2 | 3 | 3 | 6 |
| Glutaredoxin-related protein 5, mitochondrial | GLRX5 | 0.10 | 1.8 | 0.20 | 0.016 | 0.52 | 0.038 | 8.7 | 1.1 | 1.9 | 2 | 1 | 1 |
| Isoform 3 of Glypican-3 | GPC3 | 2.9 | 1.00 | 2.9 | 0.028 | 0.55 | 0.024 | 5.2 | 15.1 | 15.2 | 3 | 4 | 5 |
| GrpE protein homolog 1, mitochondrial | GRPEL1 | 0.30 | 2.40 | 0.71 | 0.0057 | 0.059 | 0.2 | 17.6 | 5.1 | 12.5 | 6 | 3 | 5 |
| Isoform 2 of General transcription factor II-I | GTF2I | 4.80 | 1.90 | 9.20 | 0.12 | 0.17 | 0.0094 | 1.0 | 4.8 | 9.3 | 1 | 3 | 7 |
| Histone H1.0 | H1FO | 2.00 | 0.10 | 0.30 | 0.11 | 0.0028 | 0.11 | 6.4 | 13.0 | 1.9 | 2 | 3 | 1 |
| Isoform 4 of Histidine--tRNA ligase, cytoplasmic | HARS | INF | 1.4 | INF | 0.07 | 0.41 | 0.021 | 0.0 | 4.0 | 5.7 | 0 | 2 | 1 |
| Beta-hexosaminidase subunit beta | HEXB | INF | 2.3 | INF | 0.14 | 0.19 | 0.011 | 0.0 | 3.0 | 6.7 | 0 | 1 | 2 |
| Histidine triad nucleotide-binding protein 2, mitochondrial | HINT2 | 0.10 | 1.80 | 0.20 | 0.016 | 0.52 | 0.038 | 8.3 | 1.0 | 1.9 | 2 | 1 | 2 |
| Histone H1.4 | HIST1H1E | 1 | 0.59 | 0.59 | 0.52 | 0.049 | 0.043 | 28.9 | 28.2 | 16.1 | 6 | 6 | 4 |
| Histone H2B type 1-N | HIST1H2BN | 0.50 | 0.77 | 0.40 | 0.013 | 0.31 | 0.0019 | 34.1 | 16.9 | 13.4 | 2 | 2 | 2 |
| Heterogeneous nuclear ribonucleoproteins C1/C2 | HNRNPC | 1.5 | 0.77 | 1.1 | 0.025 | 0.071 | 0.33 | 41.9 | 63.5 | 47.7 | 8 | 10 | 7 |
| Heterogeneous nuclear ribonucleoprotein H2 | HNRNPH2 | 1.7 | 1.2 | 2 | 0.048 | 0.35 | 0.014 | 13.8 | 24.1 | 27.8 | 1 | 2 | 3 |

Table 11 continued...

| Protein name | Gene name | Fold change | | | P value | | | QV | | | Peptide | | |
|--|-----------|-------------|------|------|----------|---------|---------|-------|-------|-------|---------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I |
| Endoplasmic | HSP90B1 | 0.77 | 1.00 | 0.77 | 0.0055 | 0.44 | 0.0088 | 320.3 | 256.6 | 261.8 | 43 | 39 | 38 |
| Heat shock 70 kDa protein 1A/1B | HSPA1A | 0.77 | 1.10 | 0.77 | 0.017 | 0.24 | 0.082 | 164.8 | 127.9 | 139.9 | 25 | 22 | 26 |
| 78 kDa glucose-regulated protein | HSPA5 | 0.71 | 1.10 | 0.77 | < 0.0001 | 0.28 | 0.00022 | 394.0 | 287.6 | 303.1 | 41 | 39 | 40 |
| Heat shock protein beta-1 | HSPB1 | 0.30 | 4.90 | 1.50 | 0.0093 | 0.00016 | 0.15 | 16.9 | 5.0 | 24.5 | 4 | 3 | 4 |
| 60 kDa heat shock protein, mitochondrial | HSPD1 | 0.77 | 1.10 | 1 | 0.00033 | 0.0048 | 0.19 | 837.8 | 702.3 | 799.3 | 49 | 49 | 46 |
| 10 kDa heat shock protein, mitochondrial | HSPE1 | 0.71 | 1.00 | 0.71 | 0.02 | 0.51 | 0.022 | 93.6 | 67.0 | 67.5 | 10 | 10 | 9 |
| Isoform Beta of Heat shock protein 105 kDa | HSPH1 | 1.4 | 0.59 | 0.77 | 0.096 | 0.022 | 0.3 | 25.8 | 37.1 | 20.9 | 6 | 9 | 5 |
| Isoform 3 of Importin-5 | IPO5 | INF | 1.00 | INF | 0.0096 | 0.57 | 0.011 | 0.0 | 7.0 | 6.7 | 0 | 2 | 3 |
| Isoform Mitochondrial of Lysine--tRNA ligase | KARS | 2.20 | 1.20 | 2.60 | 0.037 | 0.38 | 0.012 | 8.4 | 18.8 | 21.9 | 3 | 4 | 4 |
| Importin subunit alpha-1 | KPNA2 | 2.70 | 1 | 2.60 | 0.042 | 0.53 | 0.052 | 5.2 | 14.0 | 13.4 | 2 | 4 | 4 |
| Keratin, type I cytoskeletal 19 | KRT19 | 1.20 | 0.77 | 0.91 | 0.11 | 0.05 | 0.39 | 71.1 | 87.6 | 66.7 | 12 | 11 | 14 |
| Cytosol aminopeptidase | LAP3 | 0.59 | 1.50 | 0.91 | 0.031 | 0.086 | 0.33 | 34.6 | 20.0 | 30.5 | 8 | 7 | 11 |
| Isoform 3 of L-lactate dehydrogenase A chain | LDHA | 1.4 | 0.77 | 1.10 | 0.032 | 0.077 | 0.35 | 72.7 | 98.7 | 79.2 | 16 | 19 | 18 |
| Galectin-1 | LGALS1 | 0.00 | INF | 0.59 | 0.027 | 0.13 | 0.31 | 5.1 | 0.0 | 2.8 | 2 | 0 | 2 |
| Lamin-B2 | LMNB2 | 0.91 | 0.71 | 0.59 | 0.27 | 0.13 | 0.035 | 50.0 | 43.0 | 32.3 | 11 | 8 | 6 |
| Alpha-2-macroglobulin receptor-associated protein | LRPAP1 | 0.20 | 6.40 | 1.1 | 0.053 | 0.04 | 0.58 | 6.4 | 1.1 | 6.8 | 4 | 1 | 5 |
| Leucine-rich PPR motif-containing protein, mitochondrial | LRPPRC | 1.40 | 0.77 | 1.20 | 0.00065 | 0.04 | 0.069 | 159.2 | 222.7 | 187.2 | 37 | 49 | 45 |
| Leucine-rich repeat-containing protein 47 | LRRC47 | 7.30 | 0.59 | 4.60 | 0.014 | 0.27 | 0.082 | 1.2 | 9.0 | 5.7 | 1 | 4 | 2 |
| Mesencephalic astrocyte-derived neurotrophic factor | MANF | 0.59 | 1.70 | 1 | 0.048 | 0.059 | 0.49 | 30.0 | 17.8 | 29.5 | 6 | 4 | 7 |
| Methionine--tRNA ligase, cytoplasmic | MARS | 2.70 | 0.77 | 2.10 | 0.024 | 0.31 | 0.091 | 6.4 | 17.0 | 13.3 | 2 | 5 | 4 |
| Matrin-3 | MATR3 | 0.91 | 0.77 | 0.71 | 0.27 | 0.11 | 0.028 | 82.2 | 72.8 | 57.9 | 18 | 15 | 15 |
| DNA replication licensing factor MCM6 | MCM6 | 4 | 0.40 | 1.6 | 0.0085 | 0.037 | 0.34 | 4.0 | 16.0 | 6.6 | 2 | 6 | 2 |
| Mitochondrial genome maintenance exonuclease 1 | MGME1 | 0.10 | 5.90 | 0.71 | 0.016 | 0.07 | 0.32 | 8.7 | 1.0 | 5.7 | 1 | 1 | 2 |
| Isoform 2 of Mannosyl-oligosaccharide glucosidase | MOGS | 1.80 | 0.91 | 1.60 | 0.023 | 0.33 | 0.071 | 19.0 | 34.9 | 30.4 | 5 | 8 | 7 |
| 39S ribosomal protein L45, mitochondrial | MRPL45 | 0.40 | 1.00 | 0.30 | 0.037 | 0.59 | 0.029 | 13.9 | 5.1 | 4.8 | 2 | 2 | 2 |
| 39S ribosomal protein L49, mitochondrial | MRPL49 | 0.10 | 4.8 | 0.50 | 0.0084 | 0.12 | 0.16 | 9.6 | 1.0 | 4.8 | 3 | 1 | 3 |
| 28S ribosomal protein S35, mitochondrial | MRPS35 | 2.00 | 0.30 | 0.71 | 0.13 | 0.049 | 0.43 | 5.4 | 10.9 | 3.8 | 2 | 2 | 2 |
| Mitochondrial carrier homolog 2 | MTCH2 | 2.10 | 0.50 | 1.00 | 0.037 | 0.03 | 0.57 | 10.7 | 21.9 | 10.4 | 2 | 2 | 2 |
| C-1-tetrahydrofolate synthase, cytoplasmic | MTHFD1 | 1.9 | 0.77 | 1.40 | 0.048 | 0.22 | 0.22 | 12.8 | 23.7 | 18.0 | 4 | 9 | 6 |
| Sialic acid synthase | NANS | 0.50 | 2.30 | 1.20 | 0.069 | 0.017 | 0.34 | 19.3 | 9.9 | 22.7 | 7 | 5 | 7 |

Table 11 continued...

| Protein name | Gene name | Fold change | | | P value | | | QV | | | Peptide | | |
|--|-----------|-------------|------|------|----------|--------|----------|-------|-------|-------|---------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I |
| NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5 | NDUFA5 | 0.30 | 3.60 | 1.20 | 0.047 | 0.012 | 0.38 | 11.4 | 4.0 | 14.3 | 2 | 1 | 2 |
| NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial | NDUFS3 | 0.30 | 1.70 | 0.59 | 0.047 | 0.3 | 0.18 | 11.7 | 3.9 | 6.6 | 5 | 3 | 3 |
| NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial | NDUFS7 | 4.80 | 0.00 | 0.00 | 0.12 | 0.028 | 0.47 | 1.1 | 5.1 | 0.0 | 1 | 2 | 0 |
| Isoform 2 of NADH dehydrogenase [ubiquinone] flavoprotein 1, mitochondrial | NDUFV1 | 0.59 | 2.20 | 1.20 | 0.11 | 0.048 | 0.41 | 14.2 | 7.9 | 17.0 | 3 | 4 | 3 |
| Sialidase-1 | NEU1 | 0.30 | 3.50 | 1.2 | 0.047 | 0.012 | 0.38 | 11.9 | 4.1 | 14.3 | 5 | 2 | 5 |
| Nodal modulator 1 | NOMO1 | 2.40 | 0.50 | 1.10 | 0.0069 | 0.012 | 0.46 | 12.6 | 29.8 | 14.3 | 7 | 8 | 6 |
| Nucleophosmin | NPM1 | 0.77 | 1.00 | 0.71 | 0.037 | 0.45 | 0.022 | 96.5 | 73.0 | 70.2 | 10 | 10 | 9 |
| Vesicle-fusing ATPase | NSF | INF | 0.59 | INF | 0.019 | 0.35 | 0.077 | 0.0 | 6.0 | 3.8 | 0 | 2 | 2 |
| NSFL1 cofactor p47 | NSFL1C | 0.20 | 5.00 | 1.00 | 0.0026 | 0.003 | 0.52 | 15.8 | 3.0 | 15.1 | 5 | 2 | 6 |
| Nucleobindin-1 | NUCB1 | 0.50 | 1.3 | 0.59 | 0.011 | 0.3 | 0.049 | 31.8 | 14.9 | 19.2 | 7 | 4 | 5 |
| Cleavage and polyadenylation specificity factor subunit 5 | NUDT21 | 3.00 | 0.71 | 2.20 | 0.033 | 0.3 | 0.13 | 4.3 | 12.9 | 9.6 | 2 | 2 | 2 |
| Nuclear pore glycoprotein p62 | NUP62 | 0.20 | 0.00 | 0.00 | 0.17 | 0.49 | 0.05 | 4.3 | 1.1 | 0.0 | 2 | 1 | 0 |
| Ornithine aminotransferase, mitochondrial | OAT | 0.50 | 1.70 | 0.91 | 0.017 | 0.05 | 0.34 | 31.7 | 15.9 | 27.6 | 6 | 6 | 6 |
| Protein disulfide-isomerase | P4HB | 0.77 | 1.10 | 0.91 | 0.012 | 0.25 | 0.057 | 380.2 | 319.7 | 336.7 | 44 | 39 | 43 |
| Isoform 2 of Polyadenylate-binding protein 4 | PABPC4 | 8.10 | 1.10 | 8.70 | < 0.0001 | 0.41 | < 0.0001 | 4.9 | 39.4 | 42.7 | 1 | 2 | 2 |
| Multifunctional protein ADE2 | PAICS | 2.00 | 0.71 | 1.4 | 0.0008 | 0.042 | 0.08 | 34.0 | 67.0 | 47.7 | 6 | 10 | 8 |
| Protein deglycase DJ-1 | PARK7 | 0.59 | 1.60 | 1.00 | 0.035 | 0.03 | 0.53 | 44.4 | 27.6 | 44.8 | 8 | 7 | 8 |
| Pyruvate carboxylase, mitochondrial | PC | 1.90 | 1.10 | 2.00 | 0.069 | 0.5 | 0.049 | 10.4 | 19.6 | 20.9 | 4 | 5 | 5 |
| Poly(rC)-binding protein 1 | PCBP1 | 1.10 | 1.20 | 1.3 | 0.37 | 0.091 | 0.041 | 66.5 | 71.7 | 89.5 | 12 | 10 | 12 |
| Prenylcysteine oxidase 1 | PCYOX1 | 3.40 | 0.30 | 0.91 | 0.037 | 0.023 | 0.61 | 3.2 | 11.2 | 2.9 | 2 | 4 | 1 |
| Protein disulfide-isomerase A3 | PDIA3 | 0.71 | 1.30 | 0.91 | < 0.0001 | 0.0018 | 0.043 | 322.6 | 215.2 | 279.5 | 34 | 32 | 36 |
| Protein disulfide-isomerase A4 | PDIA4 | 0.71 | 1.10 | 0.77 | 0.0026 | 0.19 | 0.028 | 201.4 | 147.7 | 163.4 | 33 | 30 | 30 |
| PDZ and LIM domain protein 1 | PDLIM1 | 0.59 | 1.70 | 1.00 | 0.0087 | 0.0068 | 0.53 | 53.7 | 30.9 | 54.0 | 12 | 9 | 11 |
| Pyridoxal-dependent decarboxylase domain-containing protein 1 | PDXDC1 | 7.10 | 0.71 | 4.70 | 0.003 | 0.23 | 0.03 | 2.0 | 14.2 | 9.4 | 2 | 4 | 5 |
| ATP-dependent 6-phosphofructokinase, liver type | PFKL | 11.0 | 0.59 | 6.30 | 0.00072 | 0.12 | 0.028 | 1.2 | 13.8 | 7.7 | 1 | 6 | 3 |
| Profilin-1 | PFN1 | 0.77 | 0.91 | 0.71 | 0.091 | 0.41 | 0.046 | 69.6 | 53.9 | 50.3 | 8 | 9 | 7 |
| Membrane-associated progesterone receptor component 2 | PGRMC2 | 2.30 | 0.10 | 0.30 | 0.13 | 0.016 | 0.27 | 3.5 | 8.0 | 1.0 | 1 | 3 | 1 |
| Perilipin-2 | PLIN2 | INF | 2.70 | INF | 0.036 | 0.04 | 0.00013 | 0.0 | 4.9 | 13.2 | 0 | 2 | 4 |

Table 11 continued...

| Protein name | Gene name | Fold change | | | P value | | | QV | | | Peptide | | |
|---|-----------|-------------|------|------|----------|----------|--------|------|------|------|---------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I |
| Serum paraoxonase/arylesterase 2 | PON2 | 1.00 | 1.60 | 1.60 | 0.51 | 0.054 | 0.044 | 19.7 | 19.7 | 32.4 | 6 | 6 | 7 |
| Peptidyl-prolyl cis-trans isomerase A | PPIA | 0.71 | 1.20 | 0.77 | 0.022 | 0.23 | 0.11 | 89.3 | 63.7 | 73.3 | 10 | 10 | 10 |
| Protein phosphatase 1G | PPM1G | 1.70 | 2.20 | 3.80 | 0.31 | 0.058 | 0.011 | 4.0 | 6.9 | 15.2 | 2 | 2 | 4 |
| DNA-dependent protein kinase catalytic subunit | PRKDC | 2.00 | 0.59 | 1.20 | 0.0052 | 0.031 | 0.26 | 21.4 | 42.5 | 26.6 | 7 | 14 | 7 |
| Protein arginine N-methyltransferase 5 | PRMT5 | 4.90 | 1.30 | 6.1 | 0.025 | 0.38 | 0.0066 | 2.0 | 9.9 | 12.4 | 2 | 3 | 4 |
| Pre-mRNA-processing-splicing factor 8 | PRPF8 | 2.30 | 0.30 | 0.59 | 0.052 | 0.0073 | 0.31 | 6.3 | 14.8 | 3.8 | 1 | 6 | 1 |
| Prosaposin | PSAP | 0.71 | 1.80 | 1.30 | 0.24 | 0.046 | 0.21 | 20.2 | 14.8 | 26.5 | 4 | 3 | 6 |
| Phosphoserine aminotransferase | PSAT1 | 2 | 0.77 | 1.6 | 0.006 | 0.16 | 0.072 | 21.8 | 44.1 | 34.3 | 8 | 9 | 7 |
| Proteasome subunit alpha type-5 | PSMA5 | 0.40 | 2.30 | 1.00 | 0.031 | 0.029 | 0.55 | 20.1 | 8.8 | 20.1 | 5 | 5 | 6 |
| Proteasome subunit beta type-6 | PSMB6 | 0.59 | 1.7 | 1.10 | 0.068 | 0.042 | 0.47 | 31.4 | 19.9 | 33.3 | 5 | 4 | 6 |
| Proteasome subunit beta type-7 | PSMB7 | 0.50 | 2.30 | 1.20 | 0.054 | 0.01 | 0.31 | 21.4 | 11.0 | 25.7 | 6 | 2 | 5 |
| 26S protease regulatory subunit 4 | PSMC1 | 0.59 | 2.00 | 1.20 | 0.089 | 0.023 | 0.32 | 22.9 | 13.8 | 27.6 | 5 | 4 | 5 |
| 26S proteasome non-ATPase regulatory subunit 11 | PSMD11 | 2.30 | 0.77 | 1.90 | 0.037 | 0.38 | 0.09 | 8.4 | 19.0 | 16.1 | 2 | 4 | 3 |
| Glycogen phosphorylase, brain form | PYGB | 7.90 | 0.59 | 4.70 | 0.024 | 0.26 | 0.14 | 1.0 | 8.0 | 4.8 | 1 | 4 | 2 |
| Isoform 2 of Glutamine--tRNA ligase | QARS | 2.10 | 0.77 | 1.80 | 0.019 | 0.33 | 0.064 | 12.9 | 26.7 | 22.7 | 5 | 12 | 9 |
| Isoform 2 of Ras-related protein Rab-6A | RAB6A | INF | 0.00 | INF | < 0.0001 | < 0.0001 | 1 | 0.0 | 18.8 | 0.0 | 0 | 2 | 0 |
| Ran GTPase-activating protein 1 | RANGAP1 | 1.50 | 0.50 | 0.71 | 0.15 | 0.026 | 0.25 | 17.2 | 25.0 | 12.4 | 4 | 9 | 3 |
| Isoform 3 of Histone-binding protein RBBP4 | RBBP4 | 2.20 | 1.20 | 2.70 | 0.062 | 0.36 | 0.019 | 7.1 | 15.9 | 19.2 | 4 | 5 | 6 |
| RNA-binding protein 14 | RBM14 | 0.71 | 0.59 | 0.40 | 0.17 | 0.12 | 0.012 | 27.9 | 19.9 | 12.4 | 5 | 5 | 3 |
| Isoform 2 of Reticulocalbin-2 | RCN2 | 0.10 | 3.00 | 0.40 | 0.00026 | 0.085 | 0.023 | 19.7 | 2.8 | 8.6 | 5 | 1 | 2 |
| Transforming protein RhoA | RHOA | 1.80 | 2.30 | 4.00 | 0.28 | 0.071 | 0.011 | 3.3 | 5.8 | 13.4 | 2 | 2 | 4 |
| 60S ribosomal protein L18 | RPL18 | 0.50 | 1.10 | 0.59 | 0.033 | 0.47 | 0.053 | 25.2 | 13.0 | 14.2 | 5 | 5 | 5 |
| 60S ribosomal protein L28 | RPL28 | 1.50 | 0.40 | 0.59 | 0.17 | 0.029 | 0.24 | 11.7 | 17.8 | 7.5 | 5 | 5 | 4 |
| 60S ribosomal protein L3 | RPL3 | 1.30 | 0.71 | 0.91 | 0.095 | 0.048 | 0.42 | 49.3 | 64.2 | 45.7 | 8 | 14 | 15 |
| 60S ribosomal protein L34 | RPL34 | 2.30 | 0.30 | 0.77 | 0.075 | 0.025 | 0.43 | 6.1 | 14.1 | 4.8 | 2 | 4 | 3 |
| 60S acidic ribosomal protein P2 | RPLP2 | 0.77 | 0.77 | 0.71 | 0.18 | 0.26 | 0.048 | 53.6 | 44.1 | 37.2 | 7 | 7 | 8 |
| 40S ribosomal protein S18 | RPS18 | 0.71 | 1.40 | 0.91 | 0.039 | 0.12 | 0.31 | 47.5 | 30.9 | 42.0 | 8 | 7 | 8 |
| Ubiquitin-40S ribosomal protein S27a | RPS27A | 0.50 | 3.60 | 1.70 | 0.23 | 0.035 | 0.23 | 6.1 | 2.9 | 10.5 | 3 | 2 | 3 |
| 40S ribosomal protein S3 | RPS3 | 1.4 | 0.77 | 1.10 | 0.047 | 0.11 | 0.35 | 39.3 | 56.7 | 44.0 | 10 | 12 | 10 |
| 40S ribosomal protein S5 | RPS5 | 2.00 | 0.71 | 1.40 | 0.011 | 0.092 | 0.19 | 18.0 | 35.8 | 24.8 | 2 | 8 | 3 |
| Ribosomal protein S6 kinase alpha-3 | RPS6KA3 | 12 | 0.20 | 1.8 | 0.0024 | 0.0051 | 0.54 | 1.0 | 11.9 | 1.9 | 1 | 3 | 1 |
| 40S ribosomal protein S7 | RPS7 | 0.71 | 1.2 | 0.77 | 0.037 | 0.25 | 0.15 | 62.4 | 43.7 | 51.3 | 8 | 6 | 7 |

Table 11 continued...

| Protein name | Gene name | Fold change | | | P value | | | QV | | | Peptide | | |
|--|-----------|-------------|------|------|---------|--------|-------|-------|-------|-------|---------|----|----|
| | | F:C | I:F | I:C | F:C | I:F | I:C | C | F | I | C | F | I |
| Ribosomal L1 domain-containing protein 1 | RSL1D1 | 3.20 | 0.10 | 0.40 | 0.11 | 0.03 | 0.46 | 2.1 | 6.9 | 0.9 | 1 | 2 | 1 |
| Isoform SCP2 of Non-specific lipid-transfer protein | SCP2 | 0.59 | 2.10 | 1.30 | 0.18 | 0.028 | 0.22 | 16.9 | 10.9 | 22.7 | 6 | 4 | 8 |
| 45 kDa calcium-binding protein | SDF4 | 0.40 | 1.70 | 0.71 | 0.0061 | 0.12 | 0.11 | 27.6 | 10.9 | 18.0 | 4 | 1 | 2 |
| Succinate dehydrogenase [ubiquinone] flavoprotein subunit, mitochondrial | SDHA | 1.40 | 1.30 | 1.80 | 0.22 | 0.18 | 0.035 | 17.0 | 23.0 | 30.3 | 4 | 4 | 8 |
| Protein transport protein Sec23A | SEC23A | 2.00 | 1.00 | 2 | 0.019 | 0.53 | 0.02 | 14.7 | 30.1 | 29.6 | 5 | 8 | 8 |
| Isoform 10 of Protein transport protein Sec31A | SEC31A | 2.10 | 0.71 | 1.40 | 0.012 | 0.099 | 0.19 | 14.2 | 29.9 | 20.0 | 6 | 8 | 6 |
| Splicing factor 3B subunit 3 | SF3B3 | 9.20 | 0.40 | 3.60 | 0.00018 | 0.02 | 0.076 | 2.1 | 19.2 | 7.5 | 1 | 4 | 3 |
| Calcium-binding mitochondrial carrier protein Aralar2 | SLC25A13 | 2.1 | 0.50 | 1 | 0.0053 | 0.0025 | 0.5 | 18.6 | 39.7 | 18.0 | 5 | 11 | 5 |
| U5 small nuclear ribonucleoprotein 200 kDa helicase | SNRNP200 | 4.20 | 0.50 | 2.20 | 0.00012 | 0.02 | 0.056 | 7.3 | 30.9 | 16.1 | 3 | 9 | 5 |
| Small nuclear ribonucleoprotein E | SNRPE | 0.00 | INF | 1.10 | 0.027 | 0.018 | 0.57 | 5.2 | 0.0 | 5.7 | 1 | 0 | 3 |
| Sepiapterin reductase | SPR | 6.90 | 0.40 | 2.80 | 0.003 | 0.047 | 0.18 | 2.0 | 13.9 | 5.6 | 2 | 2 | 2 |
| Spectrin beta chain, non-erythrocytic 1 | SPTBN1 | 1.50 | 0.91 | 1.30 | 0.0019 | 0.12 | 0.045 | 88.7 | 133.4 | 114.0 | 22 | 40 | 32 |
| Sequestosome-1 | SQSTM1 | INF | INF | INF | 1 | 0.0092 | 0.011 | 0.0 | 0.0 | 6.7 | 0 | 0 | 2 |
| Isoform 2 of Threonine--tRNA ligase, cytoplasmic | TARS | 1.90 | 0.40 | 0.77 | 0.082 | 0.029 | 0.41 | 9.6 | 18.0 | 7.6 | 4 | 4 | 3 |
| Transcription elongation factor B polypeptide 2 | TCEB2 | 0.00 | INF | 3.90 | 0.48 | 0.035 | 0.14 | 1.2 | 0.0 | 4.8 | 1 | 0 | 2 |
| Serotransferrin | TF | 0.71 | 1.30 | 0.91 | 0.0032 | 0.045 | 0.15 | 130.1 | 87.4 | 112.0 | 26 | 24 | 23 |
| Transferrin receptor protein 1 | TFRC | 1.60 | 1.10 | 1.80 | 0.072 | 0.42 | 0.035 | 17.1 | 27.8 | 30.5 | 6 | 8 | 8 |
| Transketolase | TKT | 1.20 | 1.10 | 1.30 | 0.13 | 0.31 | 0.047 | 62.7 | 76.8 | 83.5 | 16 | 20 | 16 |
| Talin-1 | TLN1 | 1.80 | 1.00 | 1.70 | 0.03 | 0.47 | 0.044 | 18.6 | 33.6 | 32.3 | 8 | 11 | 10 |
| Isoform 2 of Tropomyosin alpha-3 chain | TPM3 | 0.59 | 1.40 | 0.77 | 0.012 | 0.07 | 0.23 | 54.4 | 31.8 | 45.7 | 13 | 9 | 13 |
| Tropomyosin alpha-4 chain | TPM4 | 0.59 | 1.10 | 0.71 | 0.031 | 0.34 | 0.086 | 49.2 | 30.8 | 35.1 | 6 | 4 | 5 |
| Translationally-controlled tumor protein | TPT1 | 0.59 | 1.90 | 1.10 | 0.078 | 0.024 | 0.36 | 27.8 | 16.9 | 31.3 | 3 | 4 | 6 |
| Transcription intermediary factor 1-beta | TRIM28 | 0.77 | 1.50 | 1.20 | 0.16 | 0.015 | 0.15 | 55.9 | 44.7 | 68.5 | 14 | 13 | 17 |
| Tubulin alpha-1C chain | TUBA1C | 1.20 | 0.77 | 1.00 | 0.042 | 0.023 | 0.44 | 189.0 | 226.1 | 185.3 | 3 | 3 | 3 |
| Tubulin beta chain | TUBB | 0.77 | 1.10 | 0.91 | 0.016 | 0.19 | 0.1 | 333.4 | 278.3 | 300.4 | 29 | 29 | 29 |
| Tubulin beta-4B chain | TUBB4B | 0.77 | 1.10 | 0.91 | 0.018 | 0.082 | 0.24 | 300.0 | 248.4 | 281.1 | 5 | 4 | 5 |
| Thioredoxin domain-containing protein 12 | TXNDC12 | 0.10 | 7.10 | 0.71 | 0.0084 | 0.04 | 0.32 | 9.6 | 0.9 | 6.7 | 3 | 1 | 3 |

5.2.3.1. Enrichment of identified proteins in GO categories

Enrichment analysis for GO terms was performed on comparisons of two groups at a time and so three comparisons were made – FA-treated vs. control and inhibitor-treated vs. FA-treated and inhibitor-treated vs. control.

5.2.3.1.1. Comparison of FA-treated group to control group

Given that the challenge faced by cells is the influx of exogenous FA, biological processes involved in assimilating the FA into TAG and other lipids or channelling them for breakdown were expected to be represented in the identified proteins. Most proteins identified were classified as participants in 'metabolic processes' (GO: 0008152). Of the 216 proteins, the change in abundance of 157 proteins in the FA-treated sample was more than 30 % of the control (FC ranging from 0.7 to 1.3) with values of $p \leq 0.05$. Of these 157 proteins, 120 showed enrichment in 'cellular metabolic processes' and of these, 30 proteins were enriched in 'carboxylic acid metabolic processes'. Within the GO category of 'biological regulation' (GO: 0065007), 10 identified proteins showed enrichment in 'regulation of nuclease activity'. 'Response to stimuli' (GO: 0048265) is defined as "a change in state or activity of a cell or an organism in terms of movement, secretion, enzyme production, gene expression as a result of stimulus" [298]. Within this group, the identified proteins were found enriched in processes involved in responses to misfolded and unfolded proteins (Table 12A).

The molecular functions attributed to the proteins complemented the biological processes they were classified under (Table 12B). Binding and catalytic activity are the two sub-categories GO is broadly divided into for annotation of molecular function. Binding to unfolded proteins and anion binding were the most common molecular functions attributed to the proteins. These functions are vital for chaperones that facilitate protein folding [299]. Biocatalysis in the cell is dependent on the availability of co-factors such as NAD and FAD, and nucleoside triphosphates like ATP and GTP are important in transfer of functional groups [300]. Enzymes like lactate dehydrogenase A and glucose-6-phosphate dehydrogenase use NAD as hydrogen donor and a proton carrier [289, 301], while chaperones like heat shock proteins are dependent on ATP to carry out their protein folding functions [302]. These were among the proteins enriched in the GO categories of nucleotide and purine binding.

Cellular compartments in GO are broadly grouped into cytoplasm, membrane-enclosed lumen and organelles (Table 12C). Cytoskeletal proteins such as tubulin (beta 4B class IVb and beta class I), tropomyosins and filamin were among the 64 proteins enriched in the GO

sub-group 0005829 representing the cytosol. Membrane-enclosed lumen in GO are defined as the “enclosed volume within a sealed membrane or between two sealed membranes”, and this group “encompasses the volume enclosed by the membranes of a particular organelle, e.g. endoplasmic reticulum lumen, or the space between the two lipid bilayers of a double membrane surrounding an organelle”. Among the 68 proteins enriched in this GO category were proteins localized in the ER lumen such as DnaJ (Hsp40) homolog, calnexin, prolyl 4-hydroxylase (beta polypeptide), 78 kDa glucose-regulated protein and endoplasmin. These are chaperones that play an important role in protein processing in the ER and are actively involved in ensuring correct protein folding (full list of proteins in LC4_pooled.xlsx; refer Table 21 in Appendix).

Table 12: Enrichment of protein identified in LC4 in GO categories for the comparison of FA-treated group to the control group. The table shows the number of reference proteins in the group (C), the number of proteins from LC4 identified in the group (O), the expected number in the category (E) and ratio of enrichment (R) defined as the ratio of the number of proteins identified (O) to the number of protein expected in the category (E). The p value was adjusted by multiple testing correction using Hochberg-Benjamini method. The statistics in this table are for the enrichment analysis performed; the p-value indicates the probability of the enrichment being by chance and not a result of the treatment.

| GO Category | GO ID | Statistics | | | | |
|---|---------|------------|-----|-------|-------|----------------------|
| | | C | O | E | R | P value |
| Biological processes | | | | | | |
| Translation | 0006412 | 838 | 34 | 8.48 | 4.01 | 1.34e ⁻⁹ |
| Small molecule metabolic | 0044281 | 2515 | 62 | 25.44 | 2.44 | 1.34e ⁻⁹ |
| Response to topologically incorrect protein | 0035966 | 139 | 14 | 1.41 | 9.96 | 4.19e ⁻⁸ |
| Response to unfolded protein | 0006986 | 132 | 14 | 1.34 | 10.34 | 2.78e ⁻⁸ |
| Cellular metabolic process | 0044237 | 8612 | 121 | 87.12 | 1.38 | 8.13e ⁻⁷ |
| Regulation of nuclease activity | 0032069 | 72 | 10 | 0.73 | 13.73 | 3.86e ⁻⁷ |
| Organic acid metabolic process | 0006082 | 974 | 32 | 9.85 | 9.85 | 3.67e ⁻⁷ |
| Carboxylic acid metabolic process | 0019752 | 849 | 30 | 8.59 | 3.49 | 2.7e ⁻⁷ |
| Molecular function | | | | | | |
| Anion binding | 0043168 | 2402 | 58 | 23.73 | 2.42 | 8.49e ⁻⁰⁹ |
| RNA binding | 0036094 | 854 | 31 | 8.44 | 3.63 | 3.99e ⁻⁰⁸ |
| Aminoacyl-tRNA ligase activity | 0004812 | 44 | 9 | 0.43 | 20.43 | 1.83e ⁻⁰⁸ |
| Small molecule binding | 0036094 | 2630 | 30 | 25.98 | 2.28 | 1.61e ⁻⁰⁸ |
| Nucleoside phosphate binding | 1901265 | 2437 | 24 | 24.08 | 2.17 | 4.65e ⁻⁰⁷ |
| Unfolded protein binding | 0051082 | 127 | 12 | 1.25 | 9.44 | 1.59e ⁻⁰⁷ |
| Purine nucleotide binding | 0017076 | 1871 | 42 | 18.48 | 2.27 | 4.25e ⁻⁰⁶ |
| Cellular compartment | | | | | | |
| Cytosol | 0005829 | 2372 | 64 | 21.14 | 3.03 | 9.78e ⁻¹⁶ |
| Organelles | 0043226 | 10,651 | 134 | 94.94 | 1.40 | 9.73e ⁻¹¹ |
| Membrane-enclosed lumen | 0031974 | 3385 | 68 | 30.17 | 2.25 | 7.57e ⁻¹¹ |

5.2.3.1.2. Comparison of inhibitor-treated group to FA-treated and control group

The abundance of 90 proteins identified in the inhibitor-treated sample showed changes ≥ 1.3 times that in the FA-treated sample with values of $p \leq 0.05$. Similar to the proteins in the FA-treated group, 79 of the 90 proteins were enriched in the GO category of 'metabolic processes' (Table 13A). Molecular functions of these proteins were annotated with anion-binding and very-low density lipoprotein receptor binding (Table 13B). NADH dehydrogenase (ubiquinone) 1 alpha subcomplex 5, NADH dehydrogenase (ubiquinone) flavoprotein 1 and NADH-coenzyme Q reductase were enriched in the sub-category of NADH dehydrogenase (quinone) activity. Another catalytic molecular function attributed to the proteins in this list was ligase activity exhibited by t-RNA synthetases. The proteasomal subunits alpha type 5, beta type 6 and 7 were found enriched in the category of threonine-type endopeptidase activity (GO ID: 0004298). The enrichment analysis suggested that the proteins identified in this comparison were capable of forming macromolecular complexes, and many of them were localised in the cytosol (Table 13C).

Of the 75 proteins identified in the comparison of the inhibitor-treated group to the control group, 54 proteins were identified in the comparisons described above. The biological processes in which these proteins were found enriched included metabolic processes, vesicle mediated transport and protein transport (Table 14A). Protein binding was the molecular function category with the greatest number of proteins (Table 14B). Proteins were also found enriched in the anion binding and RNA binding categories. Cellular compartment categories the proteins were found enriched in included cytoplasm and membrane-bounded organelle (Table 14C).

Table 13: Enrichment of protein identified in LC4 in GO categories for the comparison of inhibitor-treated group to FA-treated group. The table shows the number of reference proteins in the group (C), the number of proteins from LC4 identified in the group (O), the expected number in the category (E), ratio of enrichment (R) defined as the ratio of the number of proteins identified (O) to the number of protein expected in the category (E). The p value was adjusted by multiple testing correction using Hochberg-Benjamini method. The statistics in this table are for the enrichment analysis performed; the p-value indicates the probability of the enrichment being by chance and not a result of the treatment.

| GO Category | Go ID | Statistics | | | | |
|--|---------|------------|----|-------|-------|----------------------|
| | | C | O | E | R | P value |
| Biological processes | | | | | | |
| Small molecule metabolic process | 0044281 | 2515 | 37 | 14.96 | 2.4 | 5.47e ⁻⁰⁶ |
| Cellular metabolic process | 0044237 | 8612 | 76 | 51.21 | 1.48 | 5.47e ⁻⁰⁶ |
| mRNA metabolic process | 0016071 | 613 | 18 | 3.65 | 4.49 | 5.47e ⁻⁰⁶ |
| Primary metabolic process | 0044238 | 8719 | 71 | 51.85 | 1.37 | 0.0009 |
| Molecular function | | | | | | |
| Anion binding | 0043168 | 2402 | 29 | 13.52 | 2.15 | 0.0046 |
| Very-low-density lipoprotein particle receptor binding | 0070326 | 4 | 2 | 0.02 | 88.86 | 0.0081 |
| Threonine-type peptidase activity | 0070003 | 21 | 3 | 0.12 | 25.39 | 0.0081 |
| Catalytic activity | 0003824 | 5371 | 46 | 30.22 | 1.52 | 0.0108 |
| Cellular compartment | | | | | | |
| Organelle part | 0044422 | 6812 | 67 | 35.55 | 1.88 | 4.60e ⁻¹⁰ |
| Cytoplasm | 0005737 | 9130 | 78 | 47.65 | 1.64 | 3.09e ⁻¹⁰ |
| Macromolecular complex | 0032991 | 3864 | 46 | 20.17 | 2.28 | 5.35e ⁻⁰⁸ |
| Membrane-enclosed lumen | 0031974 | 3385 | 44 | 17.67 | 2.49 | 1.10e ⁻⁰⁸ |

Table 14: Enrichment of protein identified in LC4 in GO categories for the comparison of inhibitor-treated group to control group. The table shows the number of reference proteins in the group (C), the number of proteins from LC4 identified in the group (O), the expected number in the category (E), ratio of enrichment (R) defined as the ratio of the number of proteins identified (O) to the number of protein expected in the category (E). The p value was adjusted by multiple testing correction using Hochberg-Benjamini method. The statistics in this table are for the enrichment analysis performed; the p-value indicates the probability of the enrichment being by chance and not a result of the treatment.

| Go category | GO ID | Statistics | | | | |
|--------------------------------|---------|------------|----|-------|-------|----------------------|
| | | C | O | E | R | P value |
| Biological processes | | | | | | |
| Metabolic process | 0008152 | 9488 | 65 | 46.69 | 1.39 | 6.80e ⁻⁰⁵ |
| Cellular localization | 0051641 | 1977 | 28 | 9.73 | 2.88 | 1.65e ⁻⁰⁵ |
| Macromolecule localization | 0033036 | 1828 | 27 | 9.0 | 3.0 | 1.65e ⁻⁰⁵ |
| Protein transport | 0015031 | 1234 | 21 | 6.07 | 3.46 | 3.80e ⁻⁰⁵ |
| Molecular function | | | | | | |
| Anion binding | 0043168 | 2402 | 28 | 11.63 | 2.41 | 0.0001 |
| Protein binding | 0005515 | 7337 | 54 | 35.52 | 1.52 | 0.0003 |
| Small molecule binding | 0036094 | 2630 | 29 | 12.73 | 2.28 | 0.0002 |
| Structural molecule activity | 0005198 | 621 | 12 | 3.01 | 3.99 | 0.0008 |
| Aminoacyl-tRNA ligase activity | 0004812 | 44 | 5 | 0.21 | 23.47 | 0.0001 |
| Cellular compartment | | | | | | |
| Organelles | 0043226 | 10651 | 69 | 46.22 | 1.49 | 1.74e ⁻⁰⁸ |
| Cytoplasm | 0005737 | 9130 | 67 | 39.62 | 1.69 | 5.02e ⁻¹⁰ |

5.2.3.2. Enrichment of identified proteins in KEGG pathways

With the overview of biological processes and molecular functions provided by the GO analysis, the identified proteins were analysed for enrichment in the KEGG pathways to further understand the pathways affected by these proteins. The KEGG pathway enrichment analysis was performed using all 216 proteins without any pair-wise comparison of treatment groups. This section describes the KEGG pathways the proteins were found enriched in. The top eight pathways are listed in Table 15.

Table 15: Enrichment in KEGG pathways. Pathways in which the identified proteins were found to be enriched. The table shows the number of reference proteins in the group (C), the number of proteins from LC4 identified in the group (O), the expected number in the category (E), ratio of enrichment (R) defined as the ratio of the number of proteins identified (O) to the number of protein expected in the category (E). The p value was adjusted by multiple testing correction using Hochberg-Benjamini method. The statistics in this table are for the enrichment analysis performed; the p-value indicates the probability of the enrichment being by chance and not a result of the treatment.

| Pathway | KEGG ID | Statistics | | | | |
|---|---------|------------|----|------|-------|----------|
| | | C | O | E | R | P value |
| Metabolic pathways | 01100 | 1130 | 47 | 5.58 | 8.42 | 8.86e-28 |
| Protein processing in endoplasmic reticulum | 04141 | 165 | 14 | 0.81 | 17.18 | 4.50e-12 |
| Aminoacyl-tRNA biosynthesis | 00970 | 63 | 10 | 0.31 | 32.14 | 1.70e-11 |
| Ribosome | 03040 | 127 | 10 | 0.45 | 22.01 | 6.37e-10 |
| Spliceosome | 03010 | 92 | 9 | 0.63 | 14.35 | 1.90e-07 |
| RNA transport | 00670 | 18 | 9 | 0.75 | 12.07 | 5.68e-07 |
| Fatty acid metabolism | 03013 | 151 | 6 | 0.21 | 28.25 | 5.68e-07 |

5.2.3.2.1. Metabolic processes

The KEGG pathway hsa01100 is a comprehensive map of various metabolic pathways involved in the synthesis and breakdown of macromolecules. The proteins identified in dataset LC4 play a role in the various pathways of carbohydrate, FA and nucleotide metabolism that constitute this KEGG category (Table 16).

5.2.3.2.1.1 Glucose metabolism

Glucose metabolism is central to almost all metabolic processes in the cell. It is the major substrate for cellular metabolism through glycolysis in which the 6-carbon glucose yields two molecules of 3-carbon pyruvate, ATP and NADH+H⁺ through various intermediates. Liver isoform of phosphofructokinase and triokinase/FMN cyclase in the list of identified proteins play roles in the glycolytic pathway [303]. Both enzymes showed increased abundance in the FA-treated sample compared to the control. The abundances of these two proteins in the

Table 16: List of proteins enriched in KEGG pathway hsa01100 for metabolic processes. The ‘fold change’ represents the ratio of the QV as stated in the column head; for example, ‘F: C’ is the ratio of QV in FA-treated group (F) to that in control group (C). ‘I’ represents the inhibitor-treated group. *INF indicates the FC is infinite since the denominator is zero for proteins not were detected in the control group.

| Protein name | SwissProt ID | Gene Symbol | Fold change | | |
|--|--------------|-------------|-------------|------|------|
| | | | F:C | I:F | I:C |
| Acetyl-coa acyltransferase 1 | P09110 | ACAA1 | 7.9 | 0.91 | 7.5 |
| Acyl-coa dehydrogenase, very long chain | P49748 | ACADVL | 2.1 | 0.77 | 1.7 |
| ATP citrate lyase | P53396 | ACLY | 1.7 | 0.77 | 1.3 |
| Acyl-coa synthetase long-chain family member 4 | O60488 | ACSL4 | 1.9 | 0.71 | 1.2 |
| Acyl-coa synthetase long-chain family member 5 | Q9ULC5 | ACSL5 | 1.7 | 0.50 | 0.77 |
| Acyl-coa synthetase short-chain family member 3 | Q9H6R3 | ACSS3 | 3.7 | 0.50 | 1.8 |
| Aldehyde dehydrogenase 2 family (mitochondrial) | P05091 | ALDH2 | 1.1 | 0.77 | 0.77 |
| Aldehyde dehydrogenase 3 family, member A2 | P51648 | ALDH3A2 | 5.4 | 0.50 | 2.5 |
| ATP synthase subunit alpha | P25705 | ATP5A1 | 0.91 | 0.91 | 0.77 |
| ATP synthase subunit beta, mitochondrial | P06576 | ATP5B | 0.77 | 1.1 | 0.77 |
| Catechol-O-methyltransferase | P21964 | COMT | 0.91 | 1.6 | 1.4 |
| Coproporphyrinogen oxidase | P36551 | CPOX | 0.09 | 7.2 | 0.71 |
| Citrate synthase | O75390 | CS | 2.5 | 0.71 | 1.8 |
| Dihydroxyacetone kinase 2 homolog | Q3LXA3 | DAK | 2 | 0.71 | 1.3 |
| Dihydrolipoamide dehydrogenase | P09622 | DLD | 0.59 | 0.91 | 0.59 |
| Deoxyuridine triphosphatase | P33316 | DUT | 0.30 | 2.2 | 0.77 |
| Glutamyl-prolyl-trna synthetase | P07814 | EPRS | 4.1 | 0.59 | 2.6 |
| Fatty acid synthase | P49327 | FASN | 1.3 | 0.91 | 1.1 |
| Farnesyl-diphosphate farnesyltransferase 1 | P37268 | FDFT1 | 2.3 | 0.59 | 1.4 |
| Formiminotransferase cyclodeaminase | O95954 | FTCD | INF* | 0.59 | INF* |
| Glucose-6-phosphate dehydrogenase | P11413 | G6PD | 9 | 0.50 | 4.7 |
| Glucosidase, alpha; acid | P10253 | GAA | 0.50 | 1.9 | 0.91 |
| Trifunctional purine biosynthetic protein adenosine-3 | P22102 | GART | 2.3 | 0.59 | 1.3 |
| Glutamine--fructose-6-phosphate transaminase 1 | Q06210 | GFPT1 | 2.9 | 0.59 | 1.6 |
| Hexosaminidase B (beta polypeptide) | P07686 | HEXB | INF* | 2.3 | INF* |
| Leucine aminopeptidase 3 | P28838 | LAP3 | 0.59 | 1.5 | 0.91 |
| Lactate dehydrogenase A | P00338 | LDHA | 1.4 | 0.77 | 1.1 |
| Mannosyl-oligosaccharide glucosidase | Q13724 | MOGS | 1.8 | 0.91 | 1.6 |
| C-1-tetrahydrofolate synthase, cytoplasmic | P11586 | MTHFD1 | 1.9 | 0.77 | 1.4 |
| N-acetylneuraminic acid synthase | Q9NR45 | NANS | 0.5 | 2.3 | 1.2 |
| NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5, 13kda | Q16718 | NDUFA5 | 0.3 | 3.6 | 1.2 |
| NADH dehydrogenase [ubiquinone] iron-sulfur protein 3, mitochondrial | O75489 | NDUFS3 | 0.3 | 1.7 | 0.59 |
| NADH dehydrogenase [ubiquinone] iron-sulfur protein 7, mitochondrial | O75251 | NDUFS7 | 4.8 | 0 | 0 |
| NADH dehydrogenase (ubiquinone) flavoprotein 1 | P49821 | NDUFV1 | 0.59 | 2.2 | 1.2 |
| Ornithine aminotransferase | P04181 | OAT | 0.5 | 1.7 | 0.9 |
| Multifunctional protein ADE2 | P22234 | PAICS | 2 | 0.71 | 1.4 |
| Pyruvate carboxylase | P11498 | PC | 1.9 | 1.1 | 2 |
| Phosphofructokinase, liver | P17858 | PFKL | 11 | 0.59 | 6.3 |
| Paraoxonase 2 | Q15165 | PON2 | 1 | 1.6 | 1.6 |
| Phosphoserine aminotransferase 1 | Q9Y617 | PSAT1 | 2 | 0.77 | 1.6 |
| Glutamyl-trna synthetase | P47897 | QARS | 2.1 | 0.77 | 1.8 |
| Sterol carrier protein 2 | P22307 | SCP2 | 0.59 | 2.1 | 1.3 |

Table 16 continued...

| Protein name | SwissProt ID | Gene Symbol | Fold change | | |
|---|--------------|-------------|-------------|-----|------|
| | | | F:C | I:F | I:C |
| Succinate dehydrogenase complex, subunit A, flavoprotein (Fp) | P31040 | SDHA | 1.4 | 1.3 | 1.8 |
| Sepiapterin reductase | P35270 | SPR | 6.9 | 0.4 | 2.8 |
| Transketolase | P29401 | TKT | 1.2 | 1.1 | 1.3 |
| UDP-glucose 6-dehydrogenase | O60701 | UGDH | 0.9 | 1.3 | 1.1 |
| Cytochrome b-c1 complex subunit Rieske, mitochondrial | P47985 | UQCRCF1 | 0.3 | 2.2 | 0.71 |

inhibitor- treated sample was lower than that in the FA-treated sample but greater than that of the control sample. Lactate dehydrogenase A (LDH-A) catalyses the inter-conversion of pyruvate and lactate in the cell; its abundance in FA-treated sample was 1.4 times that of the control sample. Treatment with the inhibitor was able to counter this, and the abundance of LDH-A in the inhibitor-treated sample was comparable to that in the control sample. The pyruvate formed through glycolysis or from lactate is converted into acetyl CoA by the pyruvate dehydrogenase complex. This complex is formed by three enzymes - pyruvate dehydrogenase, dihydrolipoyl transacetylase and dihydrolipoyl dehydrogenase. While peptides for all three enzymes were detected in the samples analysed in dataset LC4, only dihydrolipoyl dehydrogenase was detected with fold changes of $p < 0.05$. The abundance of dihydrolipoyl dehydrogenase in the FA-treated sample was about the same as that in the inhibitor-treated sample, and both these treatments produced a decrease in the abundance of dihydrolipoyl dehydrogenase with respect to its abundance in the control cells.

Acetyl CoA enters the Krebs cycle to produce various 6-carbon and 4-carbon intermediates that undergo decarboxylation to produce reducing equivalents ($\text{NADH} + \text{H}^+$ and FADH_2) and ATP. The mitochondrial citrate synthase and cytosolic ATP-citrate lyase (ACL) both carry out the reversible formation of citrate from oxaloacetate and acetyl CoA. The latter however requires ATP and a divalent ion to carry out the reaction and the reversible reaction i.e. formation of acetyl CoA in the cytosol is favoured over the forward reaction of citrate formation [304]. The abundance of mitochondrial citrate synthase and ACL in the FA-treated sample was 2.5 and 1.7 times that of the control. The abundance of both these proteins in the inhibitor-treated sample was lower than that of the FA-treated sample but higher than that of the control sample.

Glucose-6-phosphate dehydrogenase (G6PD) and glutamine-fructose-6-phosphate aminotransferase 1 (GFAT1) are two proteins that divert the 2-carbon sugars into ancillary

pathways – G6PD catalyses the reversible oxidation of glucose 6-phosphate and diverts it to the pentose phosphate pathway while GFAT1 catalyses the rate-limiting irreversible reaction in which fructose 6-phosphate is isomerized to glucosamine 6-phosphate with the amino group donor – glutamine itself being converted to glutamate [305]. The number of exclusive unique peptides assigned to G6PD for the three groups was 1, 6 and 3 for the control, FA-treated and inhibitor-treated samples, respectively. The corresponding QV for the protein was 1.22, 10.99 and 5.69. The increase in the abundance in FA-treated sample had the lowest p value associated with it ($p = 0.0043$), while the change in abundance in the inhibitor-treated sample with respect to the other two groups had p values of more than 0.05. The glucosamine 6-phosphate formed by GFAT1 is the first metabolite in the hexosamine biosynthesis pathway whose final product is UDP-*N*-acetylglucosamine which is used to bring about glycosylation of proteins [306]. Overexpression of *GFPT1*, the gene that encodes GFAT1 protein, in liver cells has been associated with hyperlipidemia, impaired glucose tolerance [307] and an increased production of hexosamines, a consequence of this over-expression has been linked to insulin resistance in the β -cells [308]. In fact, over-expression of *GFPT1* in HepG2 cells has been shown to induce ER stress and thereby contribute to hepatic steatosis under conditions of hyperglycemia [309]. The abundance of GFAT1 in the FA-treated group was almost 3 times that of control and 1.8 times that of the inhibitor-treated group.

5.2.3.2.1.2. Fatty acid metabolism

Excess acetyl CoA is polymerized into FA by the action of fatty acid synthase (FAS), a multi-functional enzyme that catalyses the addition of the 2-carbon acetyl moiety to a growing chain of FA to form the 16-carbon palmitic acid [310]. The endogenous FA may then be elongated to an 18-carbon FA and undergo desaturation through the action of desaturases located in the ER [311]. NADPH provides the reducing equivalents for the desaturation of these FA, and NADH-cytochrome b_5 reductase (B5R), facilitates the regeneration of NAD through the electron transfer chain to allow the further uptake of H^+ as required [312]. The abundances of FAS and B5R were elevated in both FA-treated and inhibitor-treated groups. While there was no difference in the QV values assigned to the two groups, the QV value assigned to these proteins in the control group was lower than that for the treated groups. This is consistent with the observation that NAFLD patients tend to show elevated levels of *de novo* lipogenesis [39]. Irrespective of whether the FA are exogenous or synthesized *de*

novo in the cell, they are 'activated' by the addition of a CoA moiety for retention in the cell [313].

This activation is carried out by acyl-CoA synthetases like acyl-CoA synthetases 4, acyl-CoA synthetases 5 and acyl-CoA synthetase short-chain family member 3 which are thiokinases that catalyse the ATP-dependent esterification of CoA to the acyl chain. The abundance of all three acyl-CoA synthetases was greater in the FA-treated group than the other two. The QV assigned for the proteins in the inhibitor-treated group was greater than the control group but less than that of the FA-treated group.

The activated FA could have two fates in the cell: they could be assimilated into TAG molecules for storage or be broken down by β -oxidation to acetyl CoA and reducing equivalents ($\text{NADH}+\text{H}^+$ and FADH_2). Long-chain specific acyl-CoA dehydrogenase (VLCAD) and fatty aldehyde dehydrogenase are vital dehydrogenases of β -oxidation, and their abundance in the FA-treated sample was greater than that in the inhibitor-treated or control sample [314, 315]. VLCAD was identified in the three samples on the basis of 3, 6 and 4 exclusive unique peptides for the control, FA-treated and inhibitor-treated samples, respectively, while 2, 3 and 2 exclusive peptides were identified for the three corresponding groups for fatty aldehyde dehydrogenase. The abundance of peroxisomal 3-ketoacyl-CoA thiolase which brings about the thiolytic cleavage of the 3-ketoacyl CoA to produce acetyl CoA was also found to be greater in the FA-treated and inhibitor-treated samples than in the control sample.

While no proteins of the Kennedy pathway could be detected in the three samples, the increase in the number of LDs observed through ORO staining demonstrated sustained if not increased TAG synthesis. Proteins of the Arf1/COPI machinery are necessary for initiating LD formation [316] and some of these proteins were found up-regulated in the FA-treated sample. ADP-ribosylation factor 1 and ADP-ribosylation factor 4 are essential in recruitment of COPI proteins [317], and their abundances in the FA-treated samples were 13 and 2.6 times that of the control, respectively. The abundance of ADP-ribosylation factor 1 and 4 in the inhibitor-treated sample was 20 and 2.4 times that of the control, respectively. The abundance of coatamer subunits [alpha (COPA) and gamma-1 (COPG1)] were 4.7 and 2.4 times that of the control, while its abundance in the inhibitor-treated sample was about the same as that in FA-treated sample. Perilipin 2 (also known as adipophilin) is a LD biomarker localised in the LD membrane with the help of the COPI mechanism [318]. In dataset LC4, no peptide attributed to perilipin 2 was detected in the control by the LC-MS/MS method but

the protein was assigned 2 and 4 peptides for the FA-treated and inhibitor-treated samples, respectively. The absence of the perilipin 2 in the control sample despite the conspicuous LDs seen in ORO stained cells has been discussed further in Section 6.4.1.

5.2.3.2.1.3. Purine synthesis and folate metabolism

The synthesis of purines – adenine and guanine – begins with the phosphorylation of ribose-5-phosphate and culminates in inosine monophosphate through a series of reactions that constitute the *de novo* synthesis pathway which utilizes precursors and co-factors like glycine, tetrahydrofolate derivatives, glutamine, aspartate and ATP. Trifunctional purine biosynthetic protein adenosine-3 and multifunctional protein ADE2 catalyse vital irreversible reactions in this pathway, and the abundance of both proteins was greater in the FA-treated sample than the other two samples. Glycine incorporated by 5'-phosphoribosylglycinamide transformylase (GART) domain of the trifunctional purine biosynthetic protein adenosine-3 contributes C4, C5 and N7 carbons of the purine molecule. GART also catalyses the closure of the imidazole ring after it adds the C2 carbon of the purine ring. N- formyl tetrahydrofolate donates the C2 and C8 carbon of the purine ring. This vital one-carbon donor is formed through the action of cytosolic C1-tetrahydrofolate synthase (C1-THF synthase), a trifunctional enzyme that has CH₂- tetrahydrofolate dehydrogenase, 10-formyltetrahydrofolate synthetase and cyclohydrolase activities (Figure 57). The abundance of C1-THF synthase in the FA-treated sample was almost twice that of the control sample and 1.5 times that of the inhibitor-treated sample. Formimidoyltransferase-cyclodeaminase (FTCD) is a bifunctional enzyme that carries out the final two reactions in the L-histidine degradation pathway. It transfers the N-formimino group of the N-formiminoglutamic acid formed during histidine degradation to tetrahydrofolate (THF) to form 5-formimino-THF and glutamate. The cyclodeaminase domain of FTCD then converts 5-formimino-THF to 5,10-methenyltetrahydrofolate which can be converted to 10 -formyl THF or 5,10- methylene THF

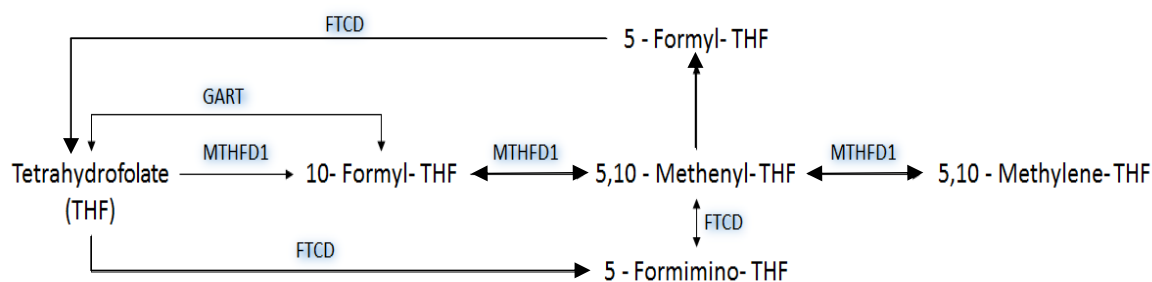


Figure 57: Schematic representation of the role of C1-THF synthase (MTHFD1), Formimidoyltransferase-cyclodeaminase (FTCD) and 5'-phosphoribosylglycinamide transformylase (GART) domain of the trifunctional purine biosynthetic protein adenosine-3 in folate metabolism.

by C1-THF synthase. Although peptides corresponding to this protein were not detected in the control sample, 4 and 2 unique peptides could be attributed to the protein in FA-treated and inhibitor-treated samples, respectively.

5.2.3.2.2. Protein processing and trafficking in the ER

Nearly one-third of the eukaryotic proteins traffic through the ER as part of the secretory pathway [319]. The proteins processed in the ER require removal of signal peptide [320], N-linked glycosylation [321] and disulphide bond formation [322] to attain the right configuration for activity. These covalent modifications are brought about by ER folding factors like the classical chaperones and their co-chaperones and the carbohydrate-binding chaperones. These chaperones are assisted by folding catalysts of the protein disulfide-isomerase (PDI) and proline *cis*–*trans* isomerase families [323] such as peptidyl-prolyl *cis*-*trans* isomerase A and peptidyl-prolyl *cis*-*trans* isomerase FKBP1A. Various components of the chaperone machinery of the ER and proteins like protein transport protein Sec23A and protein transport protein Sec31A that are a part of COPII [324] were found enriched in the KEGG pathway hsa04141 (Table 17). Both transport proteins showed increased abundance in the FA-treated group with respect to the control group. The abundance of protein transport protein Sec23A in the inhibitor-treated group was about the same as that in the FA-treated group. Protein transport protein Sec31A was assigned 6, 8 and 6 exclusive peptides for the control, FA-treated and inhibitor-treated groups, respectively; however, the QV assigned to the protein in the inhibitor-treated group was greater than that of the control group.

The abundance of classical ER chaperones like endoplasmin, 78 kDa glucose-regulated protein (Grp78; also known as BiP) and its co-chaperone DnaJ homolog subfamily B member 11 in the FA-treated sample were reduced in comparison to those in the control sample. The abundances of endoplasmin and Grp78 in the inhibitor-treated sample were about the same as those in the FA-treated sample. The co-chaperone DNAJB11 assigned 4, 1 and 3 peptides for control, FA-treated and inhibitor-treated samples, respectively. Its corresponding QV values were 12.5, 0.94 and 10.4 for the three groups, respectively. The PDI family consists of multifunctional oxidoreductases that catalyse the formation and isomerization of disulphide bonds in nascent proteins [325-327]. The protein disulphide isomerase and its name-bearing family members, protein disulphide-isomerase A3 [also known as ER-resident protein 57 (ERp57)] and protein disulphide-isomerase A4 were also reduced in abundance in FA-treated and inhibitor-treated samples with respect to the control sample.

Table 17: Proteins enriched in KEGG pathway hsa04141 for protein processing in the endoplasmic reticulum.

The proteins with changes in abundance of $p \leq 0.05$ in dataset LC4 were analysed using the KEGG enrichment tool on WebGeStalt. The table lists the proteins identified as enriched in protein processing pathways in the ER. The 'fold change' represents the ratio of the QV as stated in the column head; for example, 'F: C' is the ratio of QV in FA-treated group (F) to that in control group (C). 'I' represents the inhibitor-treated group.

| Protein name | Swissprot ID | Gene name | Fold change | | |
|---|--------------|-----------|-------------|------|------|
| | | | F:C | I:F | I:C |
| Calreticulin | P27797 | CALR | 0.77 | 1 | 0.77 |
| Calnexin | P27824 | CANX | 1.5 | 0.77 | 1.2 |
| DnaJ (hsp40) homolog, subfamily b, member 11 | Q9UBS4 | DNAJB11 | 0.08 | 11 | 0.77 |
| Endoplasmin | P14625 | HSP90B1 | 0.77 | 1 | 0.77 |
| 78 kDa glucose-regulated protein | P11021 | HSPA5 | 0.71 | 1.1 | 0.77 |
| Mannosyl-oligosaccharide glucosidase | Q13724 | MOGS | 1.8 | 0.9 | 1.6 |
| Nsfl1 (p97) cofactor (p47) | Q9UNZ2 | NSFL1C | 0.2 | 5 | 1 |
| Prolyl 4-hydroxylase, beta polypeptide | P07237 | P4HB | 0.77 | 1.1 | 0.9 |
| Protein disulphide isomerase family a, member 3 | P30101 | PDIA3 | 0.71 | 1.3 | 0.9 |
| Protein disulphide isomerase family a, member 4 | P13667 | PDIA4 | 0.71 | 1.1 | 0.77 |
| Sec23 homolog A | Q15436 | SEC23A | 2 | 1 | 2 |
| Sec31 homolog a | O94979 | SEC31A | 2.1 | 0.71 | 1.4 |
| UDP-glucose glycoprotein glucosyltransferase 1 | Q9NYU2 | UGGT1 | 1.4 | 0.9 | 1.2 |

PDI proteins belong to the thioredoxin superfamily, and new members like thioredoxin domain-containing protein 12 (also known as ERp18) are assigned to the PDI family based on their capacity to form disulphide bonds rather than break them [328, 329]. The abundance of ERp18 in the control sample was 10 times that of the FA-treated sample, and the abundance of the protein in the inhibitor-treated sample was about 7 times that of the FA-treated group. PDI and ERp57 are known to interact with calnexin and calreticulin to bring about the correct folding of nascent glycoproteins [330-332]. The abundance of calreticulin in the FA-treated sample was lower than that in the control sample, while the abundance of calnexin in the FA-treated group was 1.6 times that of the control. For the inhibitor-treated sample, the abundance of calreticulin was comparable to that in the FA-treated sample, while the abundance of calnexin was greater than that in the control sample but less than that in the FA-treated sample. Other proteins detected in dataset LC4 and known to be involved in the calnexin binding cycle included glucosidase II, mannosyl-oligosaccharide glucosidase and UDP-glucose: glycoprotein glucosyltransferase 1 (UGT1, Figure 58). There was no appreciable change in the abundances of the glucosidase II subunits, but the abundance of mannosyl-oligosaccharide glucosidase in the FA-treated and inhibitor-treated samples was 1.8 and 1.6 times that of control respectively. The abundance of UGT1 in the FA-treated and inhibitor-treated samples was also greater than that in the control sample. The lipotoxicity of FA, especially saturated FA like palmitic acid, is at least in part effected through the ER where the FA can induce ER stress through increased saturation of the ER

membrane. The increased saturation is known to cause depletion of ER chaperones, and their dislodgment from the ER membrane [333]. This can lead to an imbalance in the $[Ca^{2+}]$ balance across the membrane [334]. The reduced abundance of the classical chaperones along with the reduction in the abundance of three calcium-binding and calcium storage proteins of the ER – reticulocalbin 2 [335], calumenin [336] and 45 kDa calcium-binding protein [115] – in the FA-treated sample with respect to the control sample may be considered a sign of stress in the cells. The abundance of these Ca^{2+} binding proteins in the inhibitor-treated sample was greater than that in the FA-treated sample, these proteins were lower than in the control-treated sample. Overexpression of *VAPB* gene that encodes vesicle-associated membrane protein-associated protein B/C (VAMP-associated protein B/C) has been shown to promote UPR by activation of the IRE1/XBP1 pathway [337] which in turn is capable of inducing an inflammatory response through JNK activation [338]. VAPB also interacts with the mitochondrial protein PTPIP51 to regulate calcium homeostasis [339]. Cell division control protein 42 homolog is a GTPase that has been implicated in FA-induced activation of the JNK pathway [340]. Increased abundance of VAMP-associated protein B/C and cell division control protein 42 homolog supports the notion of cells being under stress. VAMP-associated protein B/C abundance in the FA-treated sample was 10 times that of control with the inhibitor not causing any change in the abundance with respect to the FA-treated samples. The abundance of cell division control protein 42 homolog in the FA-treated sample was 4.5 times that of control and 3.3 times that of the inhibitor-treated sample.

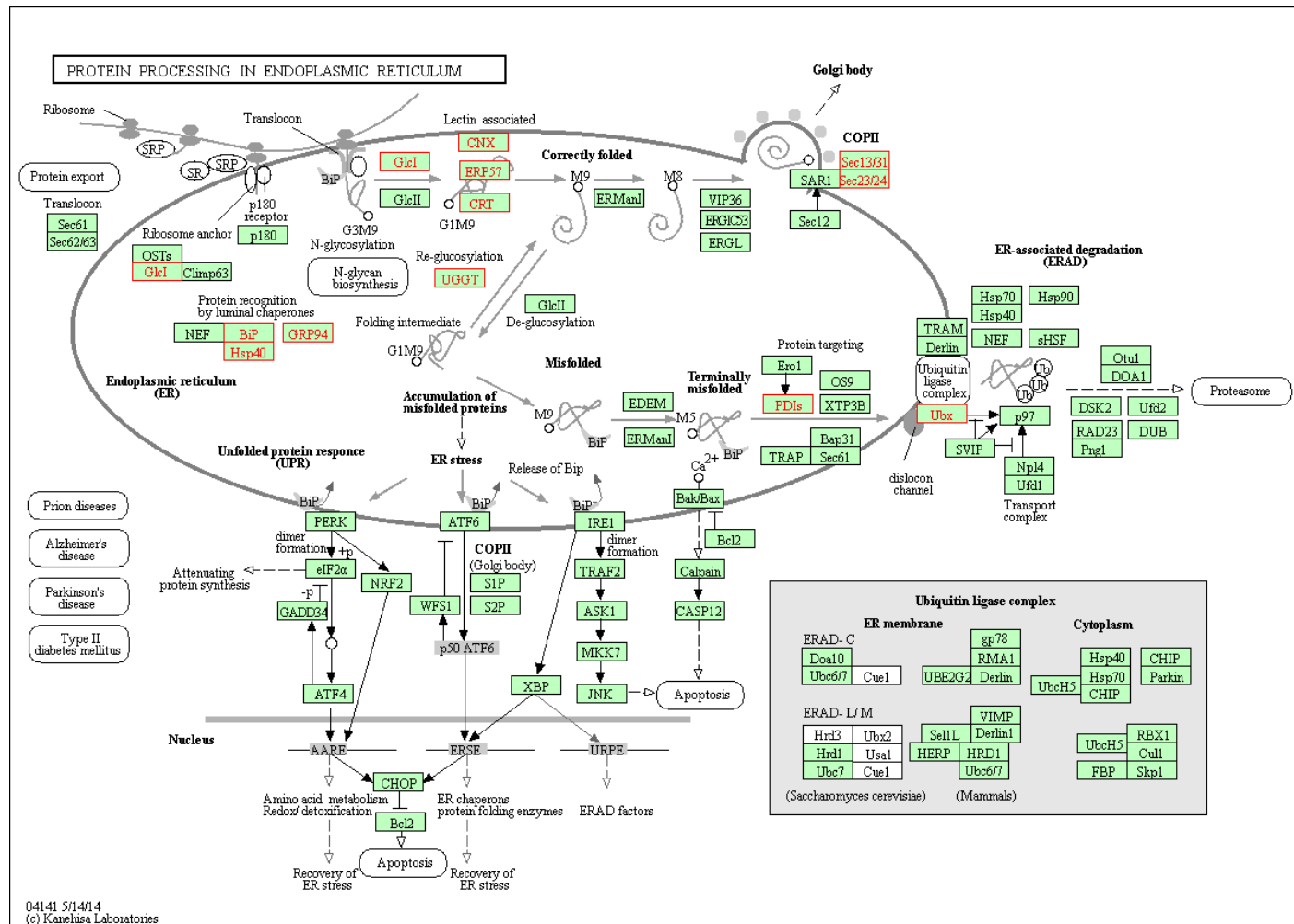


Figure 58: Protein enriched in KEGG pathway for protein processing in endoplasmic reticulum. The schematic map shows proteins identified in dataset LC4 in red. The figure was generated using the enrichment analysis tool on WebGestalt which also maps the proteins on templates available in the KEGG pathway database.

5.2.3.3. Mitochondrial proteins

FA oxidation and the Krebs cycle are important processes in the cell since they produce reducing equivalents (NADH and FADH₂) and a plethora of precursors. The retention of many of these precursors is regulated by transporters located in the mitochondrial membrane [341]. Calcium-binding mitochondrial carrier protein Aralar2 (also known as citrin) is a proton-dependent glutamate-aspartate antiporter belonging to the solute carrier family 25 (aspartate/glutamate carrier) that facilitates the exchange of mitochondrial aspartate for cytosolic glutamate and NADH [342]. Given that citrin deficiency has been associated with NAFLD and NASH [343], the increased abundance of citrin in the FA-treated sample (see Table 11) may be a part of a protective mechanism trying to maintain the cytosolic NADH:NAD ratio [344]. The abundance of this protein in the inhibitor-treated sample was about the same as that of the control.

The potential increase in the supply of electrons to the mitochondria has the capacity to uncouple the electron transport chain [345], and therefore the changes in the abundances of proteins associated with the mitochondrial respiratory complex were considered important. These changes include reduced abundance of cytochrome b-c1 complex subunit Rieske, NADH dehydrogenase [ubiquinone] 1 alpha subcomplex subunit 5, NADH dehydrogenase [ubiquinone] iron-sulphur protein 3, NADH dehydrogenase [ubiquinone] iron-sulphur protein 7, NADH dehydrogenase [ubiquinone] flavoprotein 1 and ATP synthase subunits ATP5A1 and ATP5B in the FA-treated sample.

The abundances of two mitochondrial chaperones – 60 kDa heat shock protein and 10 kDa heat shock protein – were also lower in the FA-treated sample than in the inhibitor-treated and control samples. Another protein whose abundance was reduced several fold in the FA-treated sample with respect to the other two samples was mitochondrial genome maintenance exonuclease 1. This protein is a RecB-type exonuclease localized in the mitochondria and has an important role in the maintenance of the mitochondrial genome [346]. Leucine-rich PPR motif-containing protein (LRP130) is a mitochondrial protein that has been reported to influence the expression of mitochondrial genes [347, 348] and promote oxidation of FA [349]. Over-expression of LRP130 has also been associated with increased oxidative phosphorylation which in turn increases the uptake of FA in primary hepatocytes derived from transgenic mice [350]. This protein was assigned 37, 49 and 47 exclusive peptides for control, FA-treated and inhibitor-treated samples, respectively, and the corresponding QV values suggested that its abundance in the FA-treated sample was 1.4

times that of the control and 1.2 times the abundance of the protein in the inhibitor-treated sample.

5.2.4. Validation of identified proteins

LC-MS/MS-based shotgun proteomics methods are high-throughput screens that rely on the principle of the number of peptides generated being proportional to the amount of protein present in the sample. Limitations to this principle include digestion efficiency during sample preparation and technical limitations in detecting all possible peptides for a given protein. This sets up the need for validation of the results through orthogonal methods such as Western blotting confirm the changes in protein abundance seen in the LC-MS/MS analysis. Targets for validation were selected from datasets LC3 and LC4 since they were generated from the same biological samples (Table 18).

Table 18: Proteins from dataset LC3 and LC4 selected for validation. The 'fold change' shown in this column is the ratio of QV for two groups. 'F' indicates the FA-treated sample, 'C' indicates the control sample and 'I' indicates the inhibitor-treated sample.

| Protein name | Gene name | M. W | Fold change | | | Peptides | | | | | | P Value | | |
|--|-----------|------|-------------|------|------|----------|----------|----------|----------|----------|----------|--------------------|--------------------|--------------------|
| | | | F:C | I:F | I:C | C | | F | | I | | F:C | I:F | I:C |
| Dataset LC3 | | | | | | | | | | | | | | |
| Perilipin-2 | PLIN2 | 48 | INF* | 1.9 | INF* | 0 0 | 0 0 | 2 0 | 0 3 | 3 2 | 2 5 | 0.18 | 0.31 | 0.011 |
| 60 kDa Heat shock protein | HSPD1 | 61 | 0.71 | 1.2 | 0.77 | 48 48 | 47 47 | 48 48 | 48 51 | 50 51 | 48 48 | | | |
| Catalase | CAT | 60 | 0.59 | 1.3 | 0.77 | 13 13 | 13 13 | 13 13 | 13 13 | 13 10 | 12 11 | 0.0042 | 0.034 | 0.018 |
| Glutathione reductase | GSR | 54 | 0.71 | 1 | 0.71 | 12 12 | 8 8 | 5 5 | 5 5 | 4 5 | 8 5 | | | |
| Erlin-1 | ERLIN1 | 39 | 6 | 0.9 | 5.5 | 1 1 | 0 0 | 1 1 | 2 2 | 1 2 | 1 2 | 0.0034 | 0.87 | 0.036 |
| BAG family molecular chaperone 2 | BAG2 | 24 | 0.4 | 2 | 0.77 | 1 1 | 2 2 | 0 0 | 1 1 | 3 1 | 1 0 | | | |
| Dataset LC4 | | | | | | | | | | | | | | |
| Fatty acid synthase | FASN | 273 | 1.3 | 0.9 | 1.1 | 67 | | 82 | | 73 | | 5.0e ⁻⁴ | 0.019 | 0.11 |
| 60 kDa Heat shock protein | HSPD1 | 61 | 0.77 | 1.1 | 1 | 49 | | 49 | | 46 | | 3.3e ⁻⁴ | 4.8e ⁻³ | 0.19 |
| Catalase | CAT | 60 | 0.77 | 1.4 | 1.1 | 15 | | 11 | | 13 | | 0.18 | 0.056 | 0.29 |
| Glutathione reductase | GSR | 54 | 1 | 0.59 | 0.7 | 7 | | 5 | | 5 | | 0.47 | 0.089 | 0.13 |
| Perilipin-2 | PLIN2 | 48 | INF* | 2.7 | INF* | 0 | | 2 | | 4 | | 0.036 | 0.04 | 1.3e ⁻⁴ |
| *Perilipin2 was detected in the FA-treated and inhibitor treated samples of datasets LC3 and LC4; it was not detected in the control group of either datasets. Its fold change is shown as 'INF' since the QV for the protein in the control sample is zero. | | | | | | | | | | | | | | |

5.2.4.1. Optimization of Western blot protocol

Reliable quantification using Western blotting requires accurate quantification of total protein concentration, equal loading and adequate separation of groups using SDS–PAGE [351]. The uniform, if not complete transfer of separated proteins from the gel to the PVDF membrane is critical for quantification-based inferences [352, 353]. And last but not least, validation of the antibodies used and determining the linear range within which the signal detected would be proportional to the amount of protein present needs to be determined [245, 354]. Each step in the method was optimized for validation.

The protein quantification done using a BioRad protein quantification assay was based on Bradford's method. The quantification was confirmed by running a 1D gel with 10 µg of protein to be analysed by Western blotting. The gel was assessed for uniform banding after staining with Coomassie blue. The protein quantification was considered accurate as long as the standard protein curve (0.1 to 0.5 mg/ml BSA) showed a linear correlation with $R^2 > 0.98$ (Figure 59A) and the confirmation gel showed uniform banding (Figure 59B). The uniform transfer of separated proteins was confirmed by staining the post-transfer gel with Coomassie blue G 250.

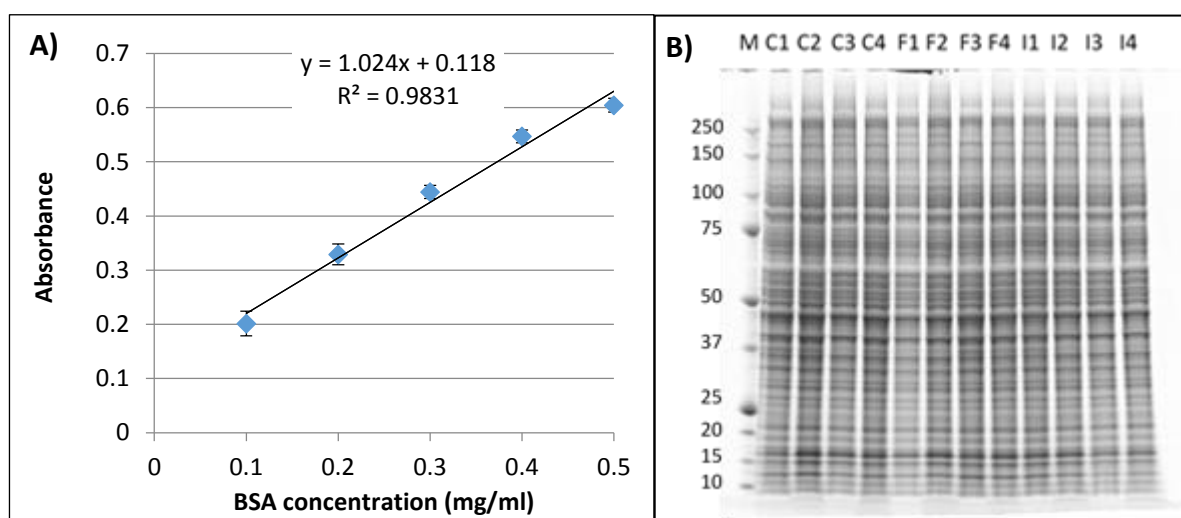


Figure 59: Protein quantification for dataset LC1. A) Standard curve for protein quantification using a BioRad protein assay and BSA concentrations ranging from 0.1 to 0.5 mg/ml. **B)** The biological replicates used to generate dataset LC1 were separated using 1D electrophoresis. The lane marked 'M' shows standard molecular weight markers labeled in kDa. Lanes C1 to C4 represent control samples, lanes F1 to F4 represent FA-treated samples and lanes I1 to I4 represent inhibitor-treated samples.

The primary antibodies tested were optimized individually to determine suitable dilutions and presence of non-specific bands. Based on the banding pattern and the molecular weight of the target protein, the Western blots were multiplexed to maximize efficient use of resources and time (Figure 60A). The amount of protein to be loaded onto the gel was determined by testing a linear range of protein amounts with different antibodies (Figure 60B). This also served as a check for saturation of signal intensity to ensure accurate quantification (Figure 60C). The intensity of the antibody signal was proportional to the increase in protein amounts, and therefore 10 µg of total protein was loaded for detection of most proteins. TRAP1 was detected using 20 µg of total protein.

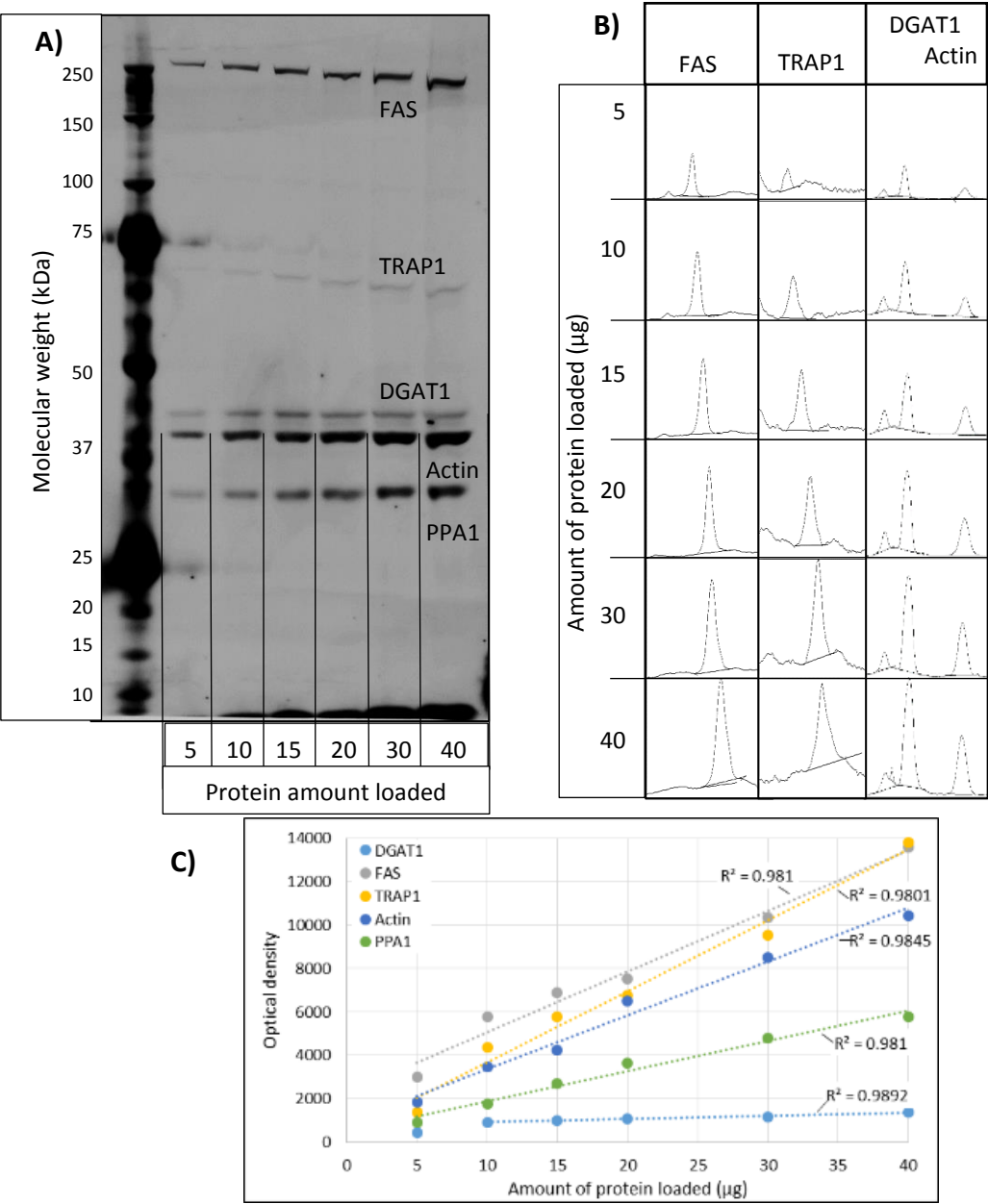


Figure 60: Calibration curves for Western blots. A) Multiplexed Western blot showing FAS, TRAP1, DGAT1, Actin and PPA1. Although DGAT1 was not detected in the datasets, the antibody was validated for assessing the possible effects of DGAT inhibitors. **B)** Quantification of optical density of the bands using ImageJ. **C)** Linear range of quantification for the proteins detected.

5.2.4.2. Target proteins

The primary criterion for selection of a protein for validation was the role it might play in lipid metabolism and lipotoxicity. The second criterion was to select proteins whose fold change could be validated. Another aspect of this validation was to highlight the impact of technical and biological variance on the LC-MS/MS experimental design. Since dataset LC3 and LC4 were generated from the same biological sample, the proteins identified and the changes in abundances associated with these proteins was expected to be similar. Comparison of the proteins identified showed that most proteins identified in dataset LC4 (generated from the pooled sample) were also identified in dataset LC3 generated from individual replicates (Figure 61A). The overlap between the two datasets reduced to a little more than a third when the list was filtered to retain proteins with $p \leq 0.05$ and fold change of more than or equal to 30 % (FC =0.71 to 1.3) (Figure 61B). The targets chosen for Western blot validation included proteins identified in both datasets.

Erlin1 is a cholesterol-binding protein that also serves as a regulator for sterol regulatory element binding proteins (SREBPs), a key transcription factors of genes involved in cholesterol and fatty acid synthesis [355]. Along with Erlin2, it also plays a role in recognizing and targeting of inositol 1,4,5-trisphosphate receptors for ER-associated degradation [356]. The gene encoding this protein is also part of the CPN1-ERLIN1-CHUK locus on chromosome 10 that has been reported to influence liver fat deposition and hepatic inflammation [357]. The abundance of Erlin1 in the FA-treated group was found to be 6 times that of control group with values of $p < 0.05$ assigned to the fold change in dataset LC3. BAG2 is a co-chaperone and nucleotide exchange factor required for Hsp70-mediated folding of proteins [358]. The abundance of this protein in the FA-treated group of dataset LC3 was 0.4 times that of control in dataset LC3. Both Erlin1 and BAG2 were detected in the Western blots as

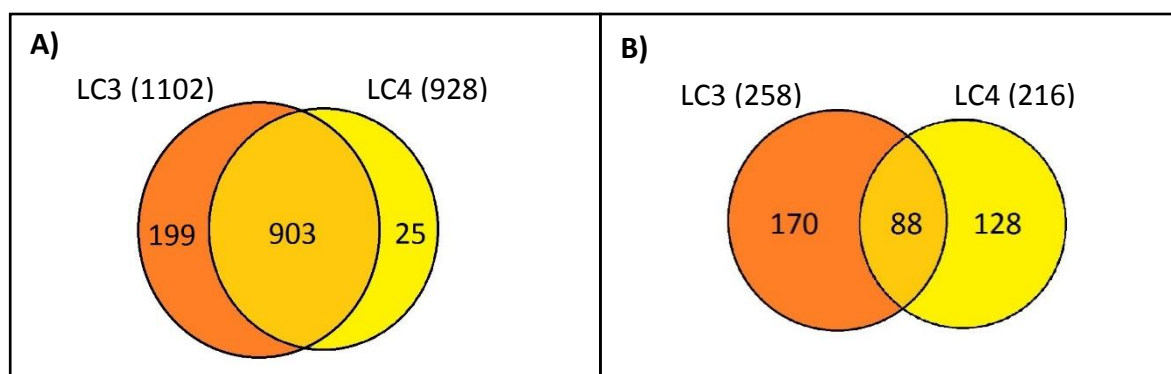


Figure 61: Comparison of proteins identified in datasets LC3 and LC4. A) All proteins identified in the dataset; **B)** Proteins with $p \leq 0.05$ and fold change ≥ 1.3 .

discrete bands, but changes in abundance predicted by the LC-MS/MS data could not be validated (Figure 62A and 62B). This suggested that multiple testing correction applied to the dataset was able to rule out false positives in spite of magnitude of change and the p values assigned to it without the FDR correction.

Similar negative results were obtained for mitochondrial 60 kDa heat shock protein (Hsp60) catalase and glutathione reductase (Figure 62C, 62D and 62E). These proteins were selected as targets for validation because all three contribute towards maintaining a stable redox state in the cell [359]. Catalase was identified in both datasets with its abundance in FA-treated sample reduced to 0.63 and 0.77 times that of control for datasets LC3 and LC4 respectively. The fold change in LC3 was assigned $p < 0.05$ and the CV associated within each group was less than 10 % for the control and inhibitor group. For dataset LC4, the fold change was assigned $p > 0.05$ and the CV for the normalized spectral counts assigned to catalase were 17 % for the technical replicates of the control and FA-treated group and 6 % for the inhibitor-treated group. In both datasets, it had more than 10 peptides assigned to it for all groups tested. The Western blot results support lack of statistical significance assigned to catalase. Hsp60 was also identified in both datasets with $p < 0.05$ and was chosen as a target because it plays an essential role in the folding of newly imported protein [360]. The changes in its abundance seen in the LC-MS/MS analysis could not be validated by the Western blots. The inclusion of glutathione reductase was also meant to serve as a corroboration of the results seen in the analysis of dataset LC1.

Perilipin 2, FAS and DGAT1 were selected as target proteins involved in lipid metabolism. Perilipin 2 serves as a marker for LDs [291]; FAS is central to de novo synthesis of FA and DGAT1 synthesizes TAG which accumulate in the LDs. FAS was detected in all three samples. The LC-MS/MS data of dataset LC3 suggest there was almost no change in abundance, but dataset LC4 indicated that the abundance of FAS in the FA-treated sample was 1.3 times that of the control sample. The p value assigned to FAS in dataset LC3 was $p > 0.05$ for all three comparisons, while it was assigned $p < 0.05$ in dataset LC4. The Western blots supported the lack of change suggested in the data from LC3. The protein could be identified in Western blots as a distinct band with no change in abundance for any of the three categories tested (Figure 62F).

DGAT1 was not detected in the LC-MS/MS data but this protein was included in the Western blots since the analysis included inhibitor-treated samples (Figure 62F). The challenge posed by the treatment was that of an excessive influx of exogenous FA. Given the preferential use

of exogenous FA by DGAT1, it would be vital in the assimilation of these FA into TAG [124]. The abundance of DGAT1, as reported in this thesis, was not affected by the influx of exogenous FA. The effect of the exogenous FA on the intracellular activity however remains to be determined. In vitro assays using microsomes as the enzyme source describe 'overt' and 'latent' DGAT activity contributing to the total DGAT activity [361] - the DGAT activity attributed to the intact microsome is termed 'overt' (cytosol-facing) activity [362] and the increase in DGAT activity after lysis is termed as the 'latent' (ER lumen-facing) activity. The latent activity is thought to be corollary to DGAT activity in the ER lumen [362]. The multi-domain membrane topology of DGAT1 is the rationale behind attributing it with 'latent' DGAT activity [362]. This would allow DGAT1 to initiate LD formation through localization of TAG synthesis on the ER membrane [274] and make it vital to the extent of lipid accumulation seen in HepG2 cells after the FA treatment. Data published by Wurie *et al.* also suggests that the majority of the overt DGAT activity in HepG2 cells can be attributed DGAT1 [363]. While there is evidence to suggest that the overt and latent activity can be affected differently by ontogeny and hypolipidemic drugs [361, 364], the effect of exogenous FA on these activities has not been (to the best of my knowledge) described in literature and future experiments could shed light on this aspect.

The Western blot results showed that irrespective of the magnitude of change and the p-value assigned to it, the changes predicted by the LC-MS/MS data were not supported by Western blot quantification. This was further highlighted by the blots for perilipin 2. No peptides for perilipin 2 were detected in the control groups of either two datasets. It is one of the few proteins that has been identified in the FA-treated samples of all four datasets and has not been detected in the control samples of three of the four datasets described in this thesis. Verifying its absence in the control sample would have served as a satisfactory validation but the protein was identified in all samples of all the three groups by Western blot and the LC-MS/MS data for this protein could not be validated and (Figure 62C). Given its function as a marker for LDs, the abundance of perilipin 2 was further investigated by 2D electrophoresis and immunolabeling.

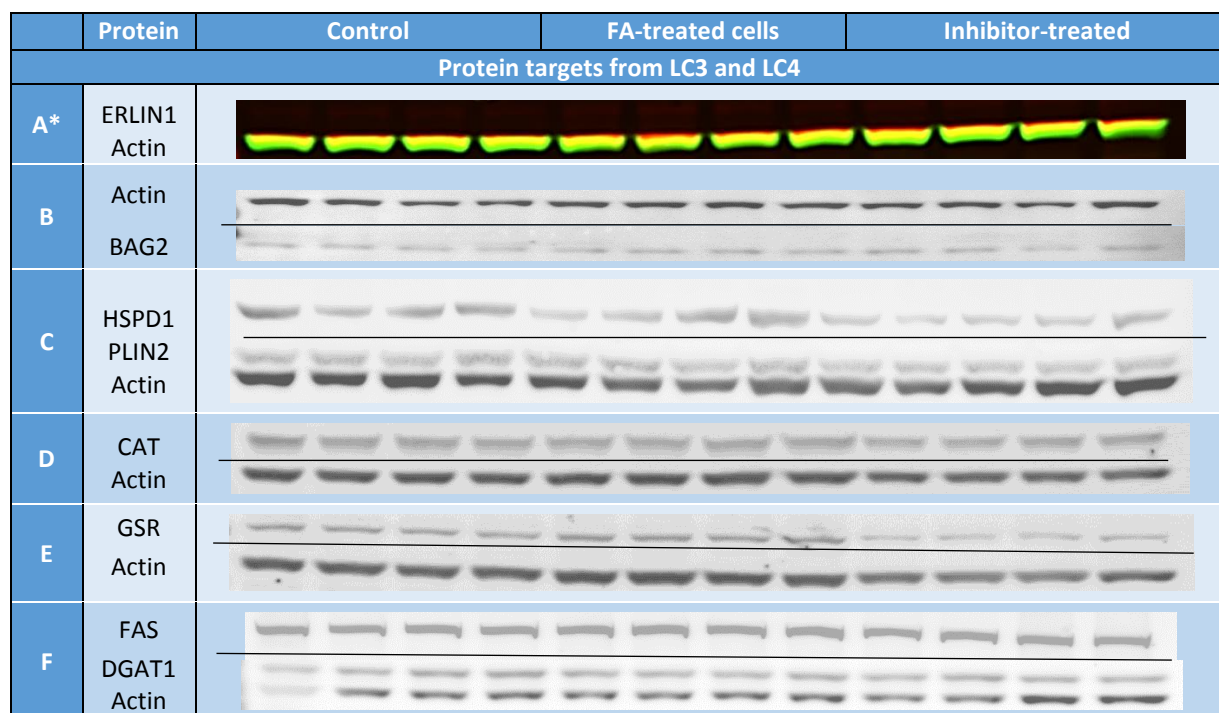


Figure 62: Western blots for validation of target proteins. Four biological replicates of each condition were tested. Actin was used as a loading control. Some images have been cropped for conserving space. The cropping is indicated by a line separating the target proteins. (Uncropped images in Appendix figure 72). * The bands for ERLIN1 and actin tend to overlap in 1D gels due to the small difference between their molecular weights. The two proteins were identified and quantified using secondary antibodies tagged with different fluorophores – Erlin1 was detected in the red channel using anti-rabbit tagged with AlexaFluor 647 while actin was detected in the green channel using anti-mouse tagged with AlexaFluor555.

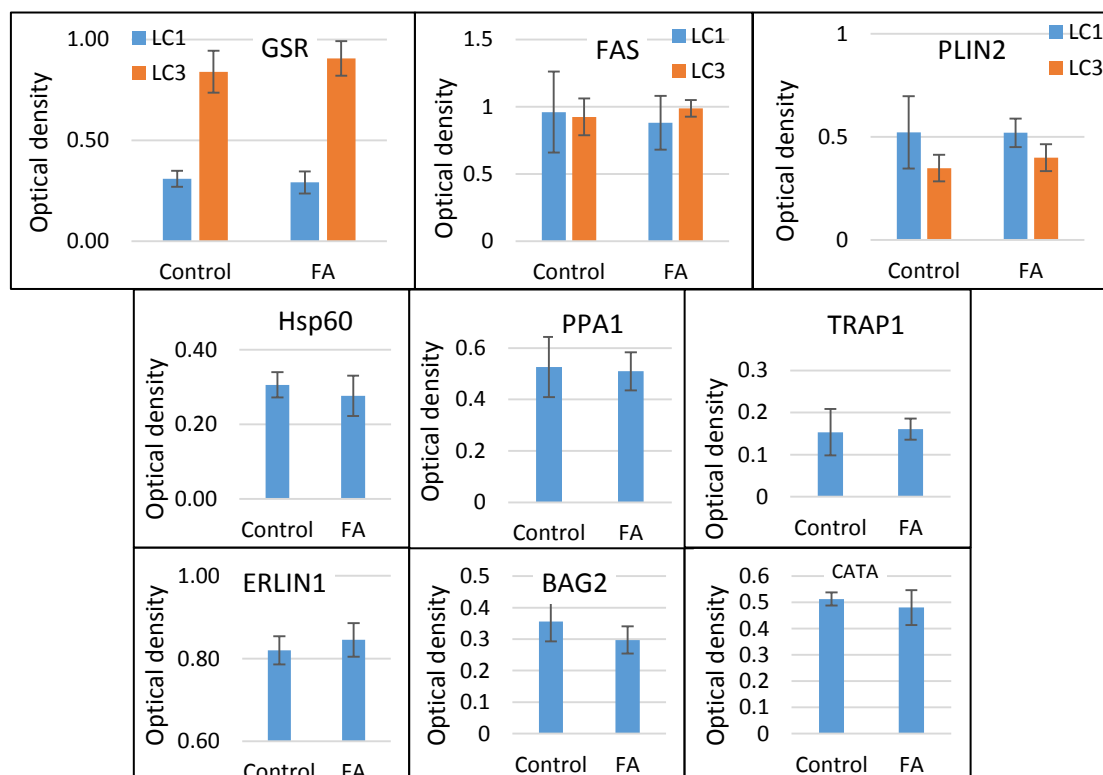


Figure 63: Quantification of Western blot bands. The bands detected in the blots shown in Figure 62 were quantified using ImageJ. The difference between the mean optical densities for protein bands detected in control groups were compared to that of the FA-treated group using a Student's t-test. None of the changes were statistically significant.

5.2.5. Further validation of perilipin 2 by 2D-electrophoresis

Given the disparity in the LC-MS/MS data and the Western blot results, a third method was sought to support the observations made using either of the two methods. Perilipin 2 was selected as a target for this validation because it was exclusively detected in all but control samples by LC-MS/MS data in all datasets but the protein was detected in the control samples tested using Western blots. To rule out the Western blot results as a false positive, the presence of perilipin 2 in control cells was examined using immunohistochemistry. This section describes the results of this validation.

5.2.5.1. Confirmation of antibody specificity for perilipin 2

The antibody used to detect perilipin 2 was raised in rabbit against a synthetic peptide (Abcam product ab90212); this synthetic peptide is described by the manufacturer as having 'sequence similarity to the perilipin family'. The perilipin family consists of 5 proteins, and of these, the amino acid sequences of perilipin 2 (48 kDa) and perilipin 3 (PLIN3; 47 kDa) have a sequence homology of up to 43 % [365]. Given the sequence homology and negligible difference in molecular weight, differentiating between the two using one-dimensional electrophoresis and Western blots would be difficult if the antigen could not discriminate between the two proteins. A protein sample from FA-treated cells was therefore probed with the anti-perilipin antibody (ab78920) after separation by 2D gel electrophoresis to check the specificity of the antibody.

The protein sample from FA-treated cells was separated in the first dimension using isoelectric focusing along a non-linear gradient of pH 3 to 11. The second dimension was separated on the basis of molecular weight. Perilipin 2 tends to have an isoelectric point near neutral pH of 6.72 [366], while PLIN3 has an isoelectric point around pH 5.3 [367, 368] with experimental evidence to suggest it could be as low as pH 4.71 [369, 370]. The reported isoelectric point for actin isolated from HepG2 cells ranges from 5.09 to 5.28 [371, 372]. Since only one spot (spot 2, Figure 64) was detected by the anti-perilipin antibody and its isoelectric point was more basic than that of actin, antibody ab78920 was shown to be capable of discriminating between perilipin 2 and perilipin 3, and spot 2 was inferred to be perilipin 2. This confirmed that the bands detected in the Western blots by antibody ab78920 were that of perilipin 2.

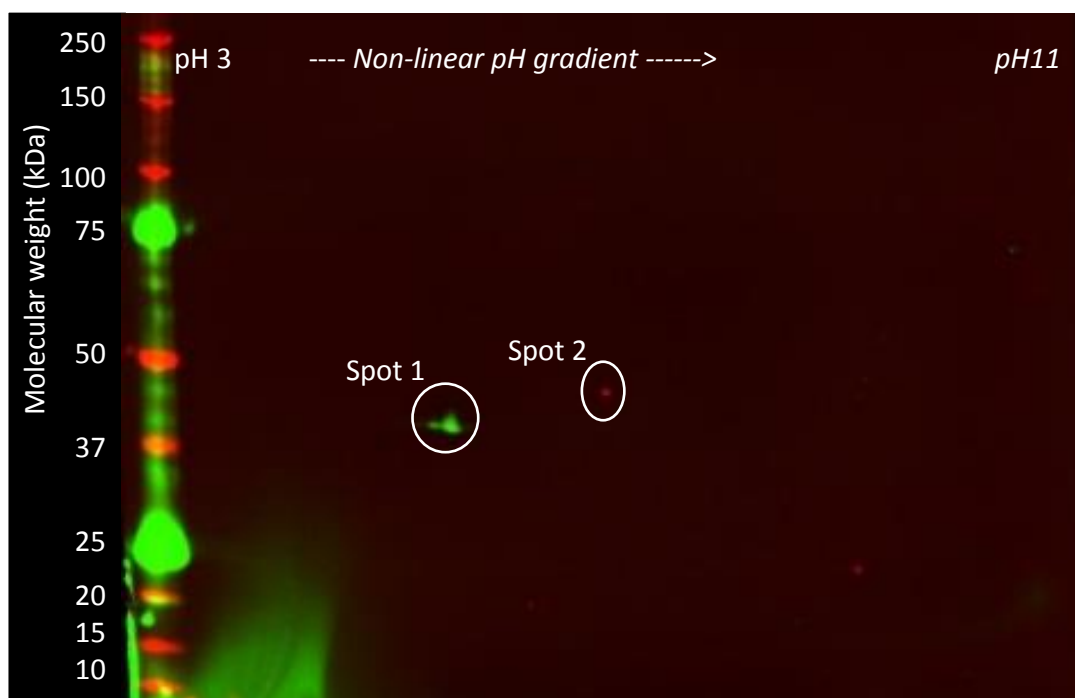


Figure 64: Detection of perilipin 2 after 2D separation. Spot 1 represents actin and spot 2 represents perilipin 2.

5.2.5.2. Immunolabeling of perilipin 2 in the cell

Having confirmed the specificity of ab78920 for perilipin 2, the presence of perilipin 2 in untreated HepG2 cells was conclusively shown by using the same antibody for immunolabeling the protein in fixed cells. The presence of perilipin 2 in both FA-treated and untreated cells agreed with the Western blot results.

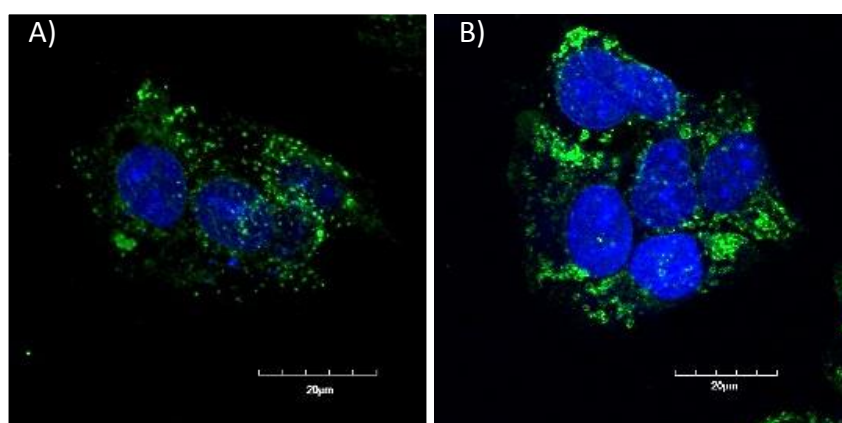


Figure 65: Detection of perilipin 2 in HepG2 cells. The images show collapsed Z stacks of **A)** untreated HepG2 cells and **B)** FA-treated HepG2 cells. The nucleus of the cells was stained with DAPI and appears as blue in the figure; the green spots represent perilipin 2 in the figure.

5.6. Summary

LC-MS/MS data representing individual biological samples were analysed to identify significant proteins. Since no statistically significant changes were observed in dataset LC3, the biological replicates were pooled to generate dataset LC4. Functional analysis of the identified proteins suggested that the proteins were enriched in metabolic pathways and protein processing pathways of the ER. Using Western blots, the presence of many of these proteins in the various groups was confirmed but the changes in abundance suggested by the LC-MS/MS data could not be supported. The presence of perilipin 2 in control cells was detected by both Western blots and immunohistochemistry and so using perilipin 2 as a representative target protein, it was shown that the Western blot results were not false positive results.

CHAPTER 6: DISCUSSION

6.1. Background overview

NAFLD starts with simple steatosis and may progress to NASH and cirrhosis. While there is a view that simple steatosis and steatohepatitis may be two discrete conditions rather than stages of NAFLD [16], the inflammation accompanying steatosis in NASH is generally considered as the cellular response to the lipotoxicity associated with steatosis. Since not every patient with simple steatosis progresses to develop NASH, it was hypothesized that lipid accumulation in steatotic cells would trigger protective cellular mechanisms that would allow the steatotic cells to cope with the problem without aggravating the disease. To test this hypothesis, HepG2 cells were treated with 1 mM FA mixture and the proteomic changes induced by lipid accumulation investigated. The shotgun proteomics approach used in this project was aimed at providing an overview of the proteome and quantifying changes in protein abundance by label-free spectral counting using LC-MS/MS.

6.2. HepG2 cells as a cell-culture model for lipid accumulation

FAs are classified as saturated and unsaturated (monounsaturated and polyunsaturated) based on the nature of the acyl chain as their structure affects their biological functions. A mixture of two FA – oleic acid (C18:1) and palmitic acid (C16:0) - commonly found in the diet [373] and synthesized *de novo* in the cell [374] was tested for effects on HepG2 cells. A combination of treatment time and FA concentration was used to induce lipid accumulation. Stock solutions of 30 mM FA were used to make a 2:1 mixture of oleic acid and palmitic acid as described by Gomez-Lechon *et al.* [230]. In their study, Gomez-Lechon *et al.* used mixtures of FA with different ratios of oleic acid and palmitic acid and observed that palmitate alone led to acute toxicity while a mixture of oleate and palmitate was able to stimulate lipid accumulation seen in steatotic cells with minor toxic and apoptotic effects [230]. Using Nile red-based fluorimetric method, they reported lipid accumulation as high as 6 times of control in HepG2 cells treated with the 1 mM FA mixture of oleate and palmitate (2: 1) for 24 h. In the current study, the extent of lipid accumulation induced by the 1 mM FA mixture of same composition after a 24 h treatment time was about 2.5 fold of control. Although this is less than reported by Gomez-Lechon *et al.*, it is similar to that reported by Yao *et al.*[231]. The lack of standard model has led to FA-induced lipid accumulation being detected and described in different ways. Even when studies use the same method, the units used to express the values may differ (Table 19).

Table 19: Comparison of studies describing lipid accumulation reported in HepG2 cells after 24 h treatment time. The FA used as indicated as PA for palmitic acid and OA for oleic acid.

| Authors | FA used | Method | Lipid accumulation reported as |
|---------------------------|---|---|--|
| Gomez-Lechon et al. [230] | 1 mM OA: PA mixture (2: 1; v/v) | Fluorimetric determination using Nile red | Relative fluorescence intensity of 5-6 fold of control |
| Yao et al. [231] | 1 mM OA: PA mixture (2 volumes: 1 volume) | Colorimetric estimation using ORO at 510 nm | Absorbance of about 0.175 for FA-treated cells and about 0.05 for control cells. |
| | | Fluorimetric determination using Nile red | Median geometric fluorescence intensity of about 300 for FA-treated cells and about 100 for control cells. |
| Cao et al. [68] | 108 μ M PA | Enzymatic assay kit for TAG quantification | Increase of 52.73 ± 7.1 μ g TAG/ mg total protein |
| Kuo et al. [375] | 0.05 – 1 mM PA | ORO-stained cells in micrographs | Images of stained cells |
| Cui et al. [233] | 1 mM OA | Colorimetric estimation using ORO at 405 nm | Optical density of about 0.57 for FA-treated cells. |

The susceptibility of HepG2 cells to the toxic effects of FA also seems to vary between reports. In the current study, HepG2 cells treated with 1 mM FA mixture were able to retain about 76 % (\pm 2.6) of the metabolic activity of the untreated cells as measured by the MTT assay. Gomez-Lechon *et al.* have reported no toxic effect on HepG2 cells at concentrations less than 2 mM FA mixture but they do comment on the HepG2 cells being ‘more sensitive to FA than primary hepatocytes’. The cell viability reported by Yao *et al.* is not affected even after a 36 h treatment with 1 mM FA mixture. Based on my results, treatment of HepG2 cells with 1 mM FA mixture for 6 h was considered to represent a cell that accumulated lipid but was able to retain, in spite of the reported lipotoxicity of FA treatments, more than 90 % of the metabolic activity of the untreated control cells. This treatment was used to test the effect of DGAT inhibitors on lipid accumulation and to analyse of potential proteomic changes.

6.3. Effect of DGAT inhibitors on lipid accumulation

The DGAT enzymes carry out the final step in the formation of TAG through the glycerol-3-phosphate pathway – the two isozymes catalyse the esterification of FA to diacylglycerol to form TAG [376]. Based on this, assays that measure DGAT activity aim to quantify the TAG formed at the end of the reaction period [236]. The effect of DGAT inhibitors is consequently measured on terms of the TAG formed. The current project has made use of the ORO assay that stains and quantifies LDs. This provided an indirect measure of DGAT activity since the extent of lipid accumulation in the form of LDs has been shown to depend on the amount of TAG produced through the action of DGAT enzymes [377, 378]. Since the extended purpose

of using DGAT inhibitors in this project was to mitigate steatosis in hepatocytes, the effect of inhibitor on the extent of lipid accumulation was deemed to be more suitable a parameter than the DGAT activity in the cell. DGAT activity was however measured using an *in vitro* assay that used microsomal preparations as source for DGAT activity.

DGAT inhibitor A92 has been shown to be orally-active and selective for DGAT1 [124, 252]. Animal models have shown dose- dependent reduction in dyslipidaemia by A92 and, administered at 3 mg/kg, A92 can reduce the serum TAG by 53 % in hyperlipidemic hamsters [253]. In cell culture systems, A92 inhibits DGAT1 activity measured by the incorporation of FA into TAG. Velliquette et al. have reported an IC_{50} of 17 μ M for A92 in HEK293H cells treated with [14 C]-glycerol and 0.3 mM oleic acid for 5 h. The amount of [14 C]- TAG was measured by liquid scintillation counting after separation using thin layer chromatography [366]. Tsuda et al. have treated HT-29 colon cancer cells and HepG2 cells with 0.2 μ Ci [14 C]-oleic acid for 30 min and measured the amount of radiolabelled TAG formed using thin layer chromatography [271]. A92 significantly reduced TAG formation at all concentrations tested in their assay. This is supported by the results described by Qi et al. in which A92 produced a dose-dependent reduction in the amount of labelled TAG produced after treating HepG2 cells with 150 μ M [13 C]-oleic acid and 1 μ M A92 for 2 h [124]. Labelled triolein was measured using a LC-MS method [367]. The A92 dependent decrease in TAG accumulation described in these studies could not be duplicated in my experiments. Although there was some initial success in cells sourced from ATCC, no detectable drop in lipid accumulation could be registered in cells sourced from the Malaghan Institute. No morphological differences could be ascertained between the cells from the two stocks and they showed the same extent of lipid accumulation when treated with the FA mix. However, genotyping the cells from the two stocks would be required to rule out any possible genomic changes that might contribute to the lack of effect of A92.

Explaining this rather unexpected lack of effect has been made more difficult by the lack of available information on the mechanisms involved in the cellular uptake of A92. Since no receptor-dependent mechanism has been described in literature, it may be assumed that A92 may be able to enter the cell through passive diffusion on account of it being a small molecule soluble in DMSO [379]. Once inside the cell, the effect of the molecule depends on its intracellular stability. The stability of the molecule would depend on the extent to which the molecule is modified and/or metabolized before it can interact with the intended target – DGAT1. The stability of the molecule after uptake in HepG2 cells has not been described in

literature and needs further investigation. While the effect of the molecule has been demonstrated after oral administration [255], microsomal assays performed as part of this project demonstrated that the molecule could exert an effect *in vitro* without any modifications made by post-ingestion metabolism. The stability of A92 in stored stock solution was confirmed in this project; however, it may also be useful to test the stability of A92 in the culture medium to ensure its availability to the cells throughout the treatment period.

The experimental design of the current project included the use of 100 μ M A92 – a higher concentration of the inhibitor may be required to ensure intracellular concentration of A92 is sufficient to bring about inhibition in the face of treatment with 1 mM FA. There is also an implicit assumption of the uptake of the molecule being greater than any possible expulsion by active or passive mechanisms. These two aspects were addressed by reducing the FA concentration to 0.25 mM and 0.5 mM to possibly improve the kinetics in favour of the inhibitor and cancelling the pre-incubation time before the treatment to ensure least possible time for efflux of the inhibitor from the cells. Neither change could produce a significant effect. The removal of BSA from the culture medium intensified the lipotoxic effects without improving the inhibitor performance.

Besides the considerations described above, the perceived lack of effect may be because of the limited sensitivity offered by the ORO-based assay used to detect the lipid accumulation. To the best of my knowledge, there have no reported studies describing the effect of A92 using the ORO assay. While the use of ORO assay offered the practical advantage of being able to provide physical quantification immediately after visual confirmation of lipid accumulation under the microscope, it is not likely to be as sensitive as measurements of radiolabelled TAG by thin layer chromatography or LC-MS measurements. Moreover, the results of experiments performed in the current project used unlabelled FA and no additional glycerol was added as substrate. Understanding the enzyme kinetics of A92 may help shed light on the case at hand. Reducing the treatment time and concentrations further to those comparable to the radiolabelled assays would not have produced the extent of lipid accumulation required to mimic steatotic hepatocytes and hence were not tested.

The second DGAT1 inhibitor tested – PF04 – was able to limit lipid accumulation in HepG2 cells treated with 1 mM FA. Lipid accumulation in the presence of PF04 was about 20 % less than that observed in the absence of the DGAT1 inhibitor. PF04 has been reported to cause a 59 % drop in plasma TAG in high fat-habituated Sprague Dawley rats when administered

at a concentration of 15 mg/kg [256]. Inhibition of TAG synthesis in HT-29 cells was measured using 0.01 % lauric acid (C12:0) and 50 μ Ci [3 H]-glycerol. Similar to the methods that describe DGAT1 inhibition by A92, the amount of labelled TAG produced after a 6 h incubation was measured and the IC₅₀ was reported to be 19 nM [131]. At 50 μ M and higher concentrations, PF04 was able to reduce the lipid accumulation induced by the condition used for my experiment. Although direct comparison is once again made difficult by the difference in the experimental design and detection methods, the study by Dow *et al.* did express concern over the low passive permeability of PF04 suggesting that the distribution co-efficient might limit the cellular uptake of the PF04 in HT-29 cells. The passive permeability of PF04 in HepG2 cells has not been reported in literature to the best of my knowledge. However, since the effect produced by PF04 at 100 μ M concentration was statistically significant, it was used to treat HepG2 cells and assess the proteomic changes induced on account of the lipid accumulation.

6.4. Proteomic changes in HepG2 cells

Steatosis observed in NAFLD patients, especially obese patients, has been attributed to alterations in adipose tissue and an imbalance between the intrahepatic production and removal of intrahepatic TAG [79]. Patients with NAFLD show increased *de novo* lipogenesis [380] and increased release of FA, especially palmitic acid from adipose tissue into the plasma [79]. This increased flux is on account of increased lipolysis in the adipose tissue and has also been associated with the insulin resistance observed in NAFLD patients [381]. Although the presence of unsaturated FA like oleic acid can attenuate some of the lipotoxic effects of saturated FA such as palmitic acid, both FA lead to lipid accumulation in the liver [230, 231, 382]. In the presence of both saturated and unsaturated FA, the incorporation of the unsaturated FA into TAG is promoted and serves as a protective mechanism against the lipotoxicity induced by saturated FA [53]. This has led to distinction between unsaturated FA being described as promoters of steatosis and saturated FA being described as lipotoxic [383]. Since non-esterified FA delivered to the liver contribute to almost half of the FA incorporated into hepatic TAG [384], a mixture of exogenous FA was used in this study to induce lipid accumulation that would simulate the effects of non-esterified FA on the liver of a NAFLD patient. Although the quantification of the proteins identified in the various datasets could be improved and requires further validation, the changes observed in the dataset are consistent with the NAFLD literature and are discussed in this section.

6.4.1. Lipid metabolism in FA-treated HepG2 cells

About 60 % of the FA assimilated into TAG accumulated in the liver originates from the non-esterified FA: *de novo* synthesis contributes about 26 % and diet contributes about 15 % of the FA [39]. Irrespective of their origin, FA need to be activated into acyl-CoA before they enter a metabolic pathway. This esterification is catalysed by a family of proteins called acyl-CoA synthetases located on the ER and outer mitochondrial membrane; they require ATP, CoA and Mg^{2+} for their activity [385]. Over-expression of the acyl-CoA synthetase 4 (*ACSL4*) gene in the liver samples of obese patients with NASH and NAFLD has also has been reported [386, 387]. In the present proteomics study (dataset LC4), the abundance of *ACSL4* gene product - long-chain acyl-CoA synthetase 4 – along with acyl-CoA synthetase short-chain family member 3 in FA-treated cells was more than that in control cells. Acyl-CoA-binding protein is an intracellular acyl-CoA transporter that is capable sequestering cytosolic acyl-CoA [388]. The reduced abundance of this protein in FA-treated cell, if confirmed through further orthogonal validation, would suggest that regular functioning of proteins like carnitine-palmitoyl-CoA transferase and the ryanodine receptor [389] would also be affected by the acyl CoA esters that would have otherwise been sequestered by the acyl-CoA binding protein [390].

De novo synthesis of acyl-CoA in the cell is performed by fatty acid synthase (FAS) – a multi-functional enzyme composed of two identical polypeptides that catalyse the synthesis of palmitate from acetyl-CoA in the presence of NADPH [391]. The proteomics data in dataset LC4 showed an increased abundance of this protein in FA-treated cells. However, the other datasets and the Western blot results did not show any change. The similar abundance of this protein in FA-treated cells and control cells indicates that *de novo* synthesis could be sustained in spite of exposure to excess amount of exogenous FA. The need to maintain *de novo* synthesis of FA could be explained in the context of a growing number and size of LDs. The final step of TAG synthesis is catalysed by the DGAT enzymes – DGAT1 and DGAT2. The two enzymes are characteristically different in terms of their topology, location and preference of FA [93]. DGAT1 preferentially uses exogenous FA while DGAT2, which has close association with SCD1, tends to preferentially use endogenous FA [124]. The *de novo* synthesis of FA is required for the continued synthesis of TAG by DGAT2. Since only DGAT2 tends to translocate from the ER membrane to the single-layer membrane of the LDs [392], the TAG synthesized by DGAT2 may contribute directly to the growth and maturation of LD as seen in the results of the time-course experiments in chapter 3. The increased abundance

of coatomer proteins and ADP-ribosylation factor 1 in FA-treated cells of dataset LC4 supports the observation of increased number of LD in FA-treated cells [60]. Sustained *de novo* synthesis of FA may be required to provide the substrate pool for DGAT2 and allow for the continued synthesis of TAG and enlargement of LDs.

6.4.2. Increased mitochondrial metabolism

Biological processes like the Krebs cycle and oxidation of FA are localized in the mitochondrion which then converts the reducing power generated in these processes into ATP through oxidative phosphorylation [393]. The HepG2 cells analysed in dataset LC4 showed increased abundance of long-chain acyl-CoA dehydrogenase and 3-ketoacyl-CoA thiolase in the FA-treated cells. These enzymes catalyse the first and last step of a β -oxidation and the reactions – facilitated by co-factors like FAD and Coenzyme A – are irreversible reactions [394]. The dehydrogenase catalyses the formation of a double bond between C2 and C3 of the FA to trans- Δ^2 -enoyl-CoA which in turn is converted to β -ketoacyl CoA through hydration and subsequent oxidation of the double bond [394]. The β -ketoacyl CoA is acted upon by 3-ketoacyl-CoA thiolase to form an acetyl-CoA molecule and an acyl-CoA molecule that is two carbons shorter than the parent FA [394].

Although the proteomics data reported in this project needs further validation and additional experiments to corroborate the increase in β -oxidation, such an increase in β -oxidation has been reported in NAFLD patients [395]. Given that impaired mitochondrial β -oxidation has been shown to cause microvesicular steatosis [396], increased β -oxidation may be one of the coping mechanisms contributing to cell survival. The FA-treated sample in dataset LC4 also showed increased abundance of both citrate synthase and ATP-citrate lyase. These proteins play a role in the removal of acetyl-CoA from the mitochondria. The mitochondrial inner membrane is impermeable to acetyl-CoA and so the acetyl moiety is transferred out of the mitochondria by an indirect shuttle system [397]. Citrate synthase, a mitochondrial enzyme, condenses acetyl-CoA and oxaloacetate to form the 6-carbon citrate which is transported out into the cytoplasm [398]. In the cytoplasm, ATP-citrate lyase cleaves the citrate to re-generate the acetyl-CoA [399]. NAFLD patients have elevated plasma citrate levels [400] and, given that citrate is also the first 6-carbon compound formed in the Krebs cycle, its elevation may explain the increased flux in Krebs cycle reported in another study of NAFLD patients [401]. It is of note that activation of citrate synthase also promotes TAG synthesis by increasing the availability of acetyl-CoA as the starting block of *de novo* lipogenesis [402].

The oxidation of FA also produces NADH and FADH₂ that donate electrons to the respiratory chain and the electrochemical gradient formed is used to drive oxidative phosphorylation. Complex I subunit B13 and NADH-ubiquinone oxidoreductase 30 kDa subunit of Complex I were identified in the proteins of dataset LC4 and although their abundances need to be validated further, the FA-treated sample showed a 2.5 fold decrease in abundance. Confirmation of abundance changes and/or damage to the electron transport chain could be done by assessing the FA-treated HepG2 cells for damage due to oxidative stress and changes in mitochondrial morphology. Given that increased β -oxidation, oxidative stress and mitochondrial structural defects have all been reported in NASH patients showing mitochondrial structural defects [403], the contribution of the metabolite derived from β -oxidation merits further study.

6.4.3. Calcium signalling between mitochondria and the ER

The changes in the mitochondria can have a cascading effect on the ER. The two organelles share a close association through a special ER domain called the mitochondria-associated membrane [404]. In fact, acyl-CoA synthetase isoforms of the rat liver including the rat homolog of long-chain acyl-CoA synthetase 4, have been reported to be localized in the mitochondria-associated membrane [405, 406]. Another key aspect of the relationship between the two organelles is the variety of physiological processes affected by the modulation of calcium signalling between the two organelles [407]. A surge in the cytosolic Ca²⁺ concentration can increase intra-mitochondrial concentration of Ca²⁺ and activate various mitochondrial enzymes involved in oxidative phosphorylation and enzymes of the Krebs cycle such as pyruvate dehydrogenase and α -ketoglutarate dehydrogenase [408]. Further, the mitochondrial aspartate-glutamate carriers Aralar1 and citrin are also stimulated and regulated by the cytosolic Ca²⁺ concentration [342, 409]. Of these two, the abundance of citrin in FA-treated cells of dataset LC4 was 2.2 times that in control cells. It is difficult to determine whether mitochondrial dysfunction is a cause or consequence of NAFLD [410] but evidence in literature suggests a calcium-mediated interplay between the ER and the mitochondria [411].

The micro-environment in the ER lumen is characteristically different from the cytosol. The redox state is more oxidized and it serves as the calcium reservoir for the cell [412]. The ER may have up to 3 mM protein-bound Ca²⁺ concentration with concentration of free Ca²⁺ varying from 1 to 400 μ M [413]. Ca²⁺ ions released by the ER travel no farther than 500 nm before they are sequestered by a calcium binding protein on the ER membrane [414]. This

sequestering of Ca^{2+} creates a transient microdomain with a high concentration of Ca^{2+} around the ER calcium release channels [415]. Mitochondria sense the Ca^{2+} concentration and localize near these domains to facilitate the uptake of Ca^{2+} released by the ER [416-419]. This proximity between the two organelles plays an important role in the transfer of Ca^{2+} between the ER and mitochondria. The concentration of Ca^{2+} in the mitochondria is in turn an important signal for the release of pro-apoptotic factors contained in the mitochondrial intermembrane space [420]. Reducing the Ca^{2+} concentration in the ER has been shown to reduce the transfer of Ca^{2+} to the mitochondria and attenuate the pro-apoptotic signals by reducing mitochondrial swelling and fragmentation [74, 421]. The reduced abundance of Ca^{2+} -binding ER proteins like reticulocalbin-2, calumenin and 45 kDa calcium-binding protein in the FA-treated sample of dataset LC4 may be part of a protective mechanism that tries to reduce further stimulation of the mitochondria by the Ca^{2+} released from the ER lumen.

6.4.4. Protein processing in the ER

Calcium depletion also tends to induce transcription of a family of glucose-regulated protein genes that encode ER chaperones including Grp78 (glucose-related protein of 78 kDa) and endoplasmin [422]. These chaperones are highly-conserved ATPase enzymes of the HSP90 chaperone family vital in the late stage folding of proteins [423]. Grp78 is a calcium-sensitive chaperone in the ER [424, 425] commonly referred to as BiP, the immunoglobulin heavy chain-binding protein [426]. The expression of Grp78 is affected by Ca^{2+} imbalances and it may be involved in 'sensing' the changes in Ca^{2+} concentration in the ER lumen [427]. Like Grp78, endoplasmin also serves as a sensor of Ca^{2+} concentration and binding of Ca^{2+} facilitates the association between the target peptides and the peptide-binding site in the N-terminal domain of the protein [428]. The abundance of both chaperones was reduced in FA-treated sample of dataset LC4. The reduced abundance of Grp78 is of particular importance given its role in the unfolded protein response. In addition to facilitating protein folding and regulating Ca^{2+} stores in the ER lumen, Grp78 also serves as a regulator of UPR [69]. In unstressed cells, Grp78 represses the UPR by forming a stable complex with ER transmembrane stress transducers PKR-like ER kinase (PERK) and inositol-requiring enzyme 1 (IRE1). Misfolded proteins relieve this repression by causing a release of Grp78 from the complex and leads to the activation UPR signalling pathways [338].

The ER-resident, carbohydrate-binding chaperones calnexin and calreticulin along with their co-chaperones of the protein disulphide isomerase (PDI) family [330, 331] were also

identified in dataset LC4. While the abundance of calnexin in the FA-treated sample of dataset LC4 was 1.6 times of control, calreticulin and the PDI chaperones showed reduced abundance in the FA-treated cells. Calreticulin is localized in the ER lumen and, as a calcium binding protein, it serves as buffer and regulator of Ca^{2+} homeostasis [429]. Its primary function is to act with calnexin and PDI co-chaperones to facilitate the proper folding of nascent glycoproteins [430, 431]. The calreticulin-PDI chaperone complex is only likely to be functional when the ER Ca^{2+} store is full [432, 433]. The PDI family members are multifunctional proteins that belong to the thioredoxin superfamily and catalyze the formation, isomerization, or reduction of disulphide bonds in nascent proteins [325, 326]. New members like thioredoxin domain-containing protein 12 (TXNDC12 also known as ERp18) are assigned to the PDI family based on their capacity to form disulphide bonds rather than break them [328, 329].

The proteomic changes discussed in this section describe the multifaceted nature of the changes induced by lipid accumulation in the cell. Targeted studies have been able to highlight the role of individual organelles, proteins and processes involved in the NAFLD and NASH. This study has taken a proteomics approach using FA-treated HepG2 cells and was aimed at being able to detect multiple parallel processes which, irrespective of the order in which they occur, contribute to the aggravation of stress and disease in the cell. The limitation imposed on all the changes described is the small number of protein identified and an even smaller number being quantified with statistical significance. Section 6.5 discusses the shotgun proteomics approach in light of the recent advances made in LC-MS/MS technology.

6.5. Improving proteome coverage

The success of proteomics has been driven by development of HPLC systems and mass spectrometers [434]. These advances have enabled almost complete identification and quantification of the yeast proteome [435]. The coverage reported for mammalian systems ranges from 4,000 – 6,000 protein groups (proteins distinguishable by identified peptides) [436, 437]. An indicator of the comprehensive coverage that has been achieved reported so far is the detection of 166,420 peptides with unique amino-acid sequence from HeLa cells by Nagaraj *et al.* [438]. Using pre-fractionation by gel-filtration and high-resolution LC-MS/MS, Nagaraj *et al.* were able to identify 10,255 different human proteins encoded by 9,207 human genes in HeLa cells [438]. The role of sample preparation, analyte separation and the capabilities of the mass spectrometer are discussed in this section.

6.5.1. Sample preparation and separation of peptides

Tissue samples are invariably heterogeneous and more complex than samples of cultured cell lines. The use of HepG2 cells for this project was intended to reduce the heterogeneity seen in liver tissue and thus simplify the analysis of the hepatocyte proteome. To ensure the comparison between samples was rational, the protein quantification for each dataset was verified by 1D gel electrophoresis and the spectral counts obtained for each protein were normalized with respect to the average total spectral count for the entire dataset. Normalization would ensure that the total spectral count in a dataset would be the same for all the samples compared and that the difference in the spectral counts of a protein could be attributed to the treatment and not a result of variation in detection of spectra.

Different lysis buffers were also assessed to ensure the use of a buffer that would facilitate protein solubilisation. Since the number of proteins identified could not be increased by the use of different lysis buffers, under-sampling on account of simultaneous elution of analytes was considered a fundamental limitation in the LC-MS/MS strategy. Although extensive coverage of the yeast proteome using a 4 h high-resolution one-dimensional LC-MS run has been reported [439], identification 4,500 to 5,000 proteins of a human cell line has required 8 h one-dimensional LC-MS runs with improved column packing [440, 441]. Many studies have also described multi-dimensional separation or pre-fractionation of samples as a strategy to increase proteome coverage [442]. Implementation of such a strategy could improve the number of proteins detected in a shotgun proteomics project such as this one by reducing overlap of analytes.

6.5.2. Effect of instrument sensitivity and peptide-spectra matching

The analysis of the separated peptides, assuming minimal overlap, can be controlled by the various LC-MS/MS settings used and the performance of the mass spectrometer in general. The instrument used for this project was an LTQ Orbitrap XL, an improved LTQ Orbitrap design introduced in 2007. The instrument includes a C-trap to provide better transmission of ions from the LTQ to the Orbitrap analyser and thereby improve sensitivity [443]. Various LC-MS/MS parameters described in chapter 4 were assessed and the final settings used improved the number of proteins identified. An important consideration to be made while comparing the proteome coverage in this project and those reporting several thousand proteins is the instrument used for generating the data. The LTQ Orbitrap technology is an

evolving technology and many studies with better coverage have used next-generation instruments [444, 445]. Instruments like the LTQ-Orbitrap Velos are capable of scan speeds of up to 5 Hz [444] as compared to 1 Hz at a resolution of 60,000 at 400 m/z. Using the 'design of experiments', Andrews *et al.* have shown that optimization of instrument parameters allows for a significant improvement in the number of identifiable protein groups [446]. They were able to identify 570 protein groups in a *S.cerevisiae* sample using a LC run time of 85 min and the LTQ Orbitrap XL™ [446]. Using the LTQ-Orbitrap Velos™, Picotti *et al.* were able to generate a spectral library containing consensus spectra for 100,815 peptides attributable to at least one of the 6,607 sequences in the *S.cerevisiae* proteome [447]. The improvement in the quality and quantity of data generated by the next-generation mass spectrometers is therefore evident. Kalli *et al.* have also demonstrated how the Orbitrap Elite™ deliver higher identification rates than the older instruments by providing better scan speed, sensitivity, dynamic range and resolving power [448].

The number of proteins identified also depends on how many of the detected spectra can be assigned to a known protein by the search algorithms. In this project, an average of 36,500 MS/MS spectra were detected for the three samples compared in dataset LC4 and of these, about 17 % [6,123 (\pm 424)] MS/MS spectra could be assigned to peptides using a tolerance of 10 ppm for parent ion. The low percentage of assigned spectra does leave much to be desired – the unassigned spectra hold the potential to improve the number of proteins identified and contribute to the spectral counts of the peptides already assigned a protein. A case for this is made in the observations related to catalase. This protein was identified in both control and FA-treated samples of dataset LC3 and LC4. In dataset LC3, 13 peptides attributed to this protein were detected in each treatment sample but the cumulative spectral counts of the 13 peptides was deemed significantly different for the two treatment groups. Based on this, it was considered for validation using Western blot. The validation was of particular importance because the statistical significance observed for dataset LC3 was not observed in dataset LC4 in spite of the two originating in the same biological material. The protein was detected in both control and FA-treated samples of Western blot and the validation supported the lack of significance observed in dataset LC4. The extent of unassigned spectra observed in the LC-MS/MS data may provide a possible explanation for the discord between the LC-MS/MS and Western blot data. A similar explanation may also be applied to the results observed for the mitochondrial 60 kDa heat shock protein attributed 49 peptides in control and FA-treated samples of dataset LC4.

One way of improving the number of assigned spectra would be to relax the identification parameters in the Proteome Discoverer workflow. Rabelo *et al.* were able to identify 14,138 peptides which mapped to 4,756 proteins in transfected HepG2 using PatternLab for Proteomics software v 3.2 [449]. They performed peptide sequence matching using the Comet algorithm with a precursor mass tolerance of 40 ppm [449]. Of the 81 proteins reported as being differentially abundant, the abundance of one protein -Myristoylated Alanine Rich Protein Kinase C Substrate – was verified using RT-qPCR [450]. The use of relaxed identification parameters and validation using qPCR may be considered for future experiments since the data reported by Rabelo *et al.* was generated using a similar MS instrument (LTQ-Orbitrap XL™). It may also be worth noting that the data reported by Rabelo *et al.* was generated using the MS instrument at a resolution of 60,000 [449]; experiments reported in this thesis were performed at a resolution of 30,000.

The process of peptide-spectra matching used in this project depended on a virtual database of peptides generated through *in-silico* tryptic digestion of proteins. The spectra generated in the LC-MS/MS experiments was then compared to this virtual database and assigned to a peptide based on the predicted *in-silico* digestion. The spectra may be left unassigned to a peptide if the peptide it originates from is absent in database of predicted peptides. This handicap and possible reason for the large number of unassigned spectra may be addressed by the comprehensive and exhaustive approach described by Picotti *et al.* in which they generated a spectral library for synthetic peptides generated for all known ORFs of the yeast genome [447]. This comprehensive peptide and spectral library was then used to match the spectra generated in discovery-led experiments to possible peptides. Such an approach could help identify the peptides to which the unassigned spectra belong.

In addition to the unassigned spectra that could adversely affect the spectral counts assigned to peptide, the spectral counts are also affected by the proteins inference problem faced while assigning spectra to peptides of very closely related proteins. Protein inference is a computational and statistical process used to determine the sequence in which detected peptides ought to be aligned to identify a protein; closely related proteins tend to have homologous sequences which would give rise to peptides that could be assigned to either or both proteins by the search algorithms and this is commonly known as the protein inference problem [451]. The abundance of perilipin 2 in the four datasets highlights this problem. Perilipin 2 shares sequence homology with proteins of the perilipin family and peptides common to multiple proteins within the same family can skew the quantification of one or

more protein. Perilipin 2 was detected in FA-treated samples in all four datasets. It was absent in three of the four LC-MS/MS datasets but its presence in control sample detected only in dataset LC2. Western blots and immunohistochemistry showed that perilipin 2 was present in FA-treated as well as untreated HepG2 cells. The protein inference problem may be the underlying reason for a protein, physically present in the cell, being reported as absent in untreated cell samples.

6.5.3. Statistical analysis

The aim of this project was to identify proteins whose abundance had changed because of lipid accumulation and/or presence of DGAT inhibitors. In statistical terms, this meant the testing of a null hypothesis that the change in protein abundance was a random event not caused by lipid accumulation and/or the presence of DGAT inhibitors. Since each protein identified would have to be tested for this hypothesis, this meant simultaneous testing of a large number of hypotheses. Depending on the experimental design, either Fisher's exact test or *t*-test was used to determine the statistical significance of the difference in the quantitative value reported for a protein in the three treatment groups tested. Fisher's exact test is based on the analysis of contingency tables where sample sizes are small. The test calculates the significance of the deviation from the null hypothesis with an exact method and does not rely on approximations that depend on the sample size. In Scaffold, this test is performed using a model discussed by Zhang *et al.*[452]. The same report also compared the false positive rate for the Fisher's exact test and *t*-test and showed that false positive rates for all tests analyzed were low when the differences in spiked protein concentrations was large (5 to 10-fold). This included data from a single replicate for the Fisher's exact test. The false positive rate for a 2-fold change was more than 10 % for the Fisher's test; the false positive rate however, reduces with an increase in the number of replicates used to generate the pooled data. Using pooled data from three replicates, they were able to reduce the false positive rate to 3 %. The choice of statistical test, based on the analysis performed by Zhang *et al.*, would depend on the number of replicates used and on whether they were pooled. For a study with three or more unpooled replicates, *t*-test generated the least number of false positives for 2-fold change detection and for a study of less than three replicates or a pooled sample, Fisher's exact test was suggested to work well as long as the pooled data included sufficient number of replicates.

Multiple testing corrections imposed a severe limitation on the interpretation of my results. It reduced the number of statistically significant protein changes identified in pooled

samples and ruled out any change as statistically significant for unpooled samples. These limitations may be discussed in the light of comments shared by Pascovici *et al.* who have described “multiple testing correction as a useful but blunt tools in the context of low power, medium-scale experiments” [453]. They suggest that factors such as the total number of proteins identified, the extent of fold change, cost-restrictions on the number of replicates used and undermining of true effects reduce the value of multiple testing corrections in a proteomics experiment [453]. The concept of multiple testing correction has been borrowed from other high-throughput areas like microarray data but the scale of information available through proteomics is not comparable. For example, Affymetrix human whole-genome microarrays can generate up to 47,000 transcripts [454] and variants while proteomics data generated using the current generation of mass spectrometers without pre-fractionation can have up to 5,000 proteins [441]. As suggested by Pascovici *et al.*, only a small percent of proteins are likely to be differentially expressed and applying strict multiple testing corrections further could reduce the number such that further analyses may become difficult. They suggest that for experiments with 3 or 4 biological replicates, the effect size i.e. fold change needs to be at least four fold for the multiple testing corrections to eliminate false discoveries without discarding true effects [453]. Using a lysate of human cell mixed with lysate of differentially expressed yeast proteins with known fold changes, Wu *et al.* have shown that in experiments with a small number of replicates ($n=3$) and small fold changes, multiple testing corrections enforce low FDR but also discard true positives [455]. This problem is exacerbated when the number of true positives in the dataset is low to begin with.

In spite of these drawbacks, the simulations run by Pascovici *et al.* confirm that in the absence of multiple testing corrections at a $p \leq 0.05$, the FDR is considerably higher. The approach suggested by them in this scenario includes applying a fold-change cut-off that would reduce the FDR % on account of the effect size. In this project – proteins showing change in abundance with $p \leq 0.05$ were further short-listed only those proteins that showed a fold change of 1.3. Western blot targets were chosen from this list for further validation.

6.6. Summary of the proteomics changes observed

Targeted study of changes associated with NAFLD have provided insights into the various molecular mechanisms that contribute to the disease and its progression. The aim of my project was to identify proteomic changes that represent these mechanisms in the FA-treated HepG2 cell model and highlight the interplay among the mechanisms. The results

suggest that HepG2 cells exposed to excess FA manage to minimize the potential lipotoxicity of these molecules by incorporating them into TAG or breaking them down to acetyl-CoA. The evidence for TAG production is seen in the large number of LDs accumulating in the cytoplasm of the FA-treated cells. Increased abundance of enzymes contributing to β -oxidation suggests the production of acetyl-CoA. The interplay between these two processes affects cellular homeostasis and the integrity of the organelles where these enzymes are located.

Breakdown of FA through β -oxidation takes place in the mitochondrial matrix where the Krebs cycle takes place. The acetyl-CoA formed as the end product of β -oxidation can be incorporated into the Krebs cycle to form 6-carbon citrate and reducing equivalents. Citrate synthase, ATP-citrate lyase and citrin are part of the citrate-malate shuttle that facilitates the removal of acetyl-CoA from the mitochondrial matrix [456]. The increased abundance of these proteins seen in dataset LC4 suggests that acetyl-CoA may be transported out of the matrix and into the cytosol for production of FA through *de novo* synthesis. The unchanged abundance of fatty acid synthase in the proteomic study suggests that HepG2 cells may continue to synthesize endogenous FA despite the presence of excess FA in the medium. A targeted study of FAS activity in the presence of exogenous FA would be required to support this inference. These endogenous FA are the preferred substrate for DGAT2 – the acyltransferase that esterifies the FA to a diacylglycerol to form TAG that accumulates in LDs. The importance of DGAT2 in steatosis has been confirmed by Choi *et al.* who have shown that suppression of the *DGAT2* gene in a rat model protected it against diet-induced steatosis [378]. Localization of DGAT2 on the membrane of LDs may help the maturation and growth of LDs and this may in turn facilitate the preferential incorporation of endogenous FA into the LDs [59, 457]. DGAT1 – the isoform that is localized on the ER and uses exogenous FA as its substrate – may use these exogenous FA to initiate LD formation through the budding process in the ER. The FA in the ER may have a direct effect on the integrity of the ER membrane. Saturated FA not esterified into TAG may be incorporated into phospholipids which, if used as membrane precursors, could reduce the fluidity of the ER membrane [333, 458]. The possible consequences of this altered membrane composition include depletion of calcium and activation of the unfolded protein response independent of signals from the ER lumen [459, 460]. Reduced abundance in FA-treated cells of Grp78 and member of the Hsp70 family of chaperones, points to this trend [461]. In addition to Grp78,

other Ca²⁺-dependent chaperones like endoplasmic reticulum chaperones were also identified in dataset LC4.

The mitochondria contain a secondary store of Ca^{2+} in the event that the cytosol concentration of Ca^{2+} increases. A release of Ca^{2+} from the ER could lead to increased uptake of Ca^{2+} by the mitochondria and contribute to the mitochondrial deformity observed in NASH patients. Experiments with primary rat hepatocytes have shown that suppression of Ca^{2+} levels can protect the hepatocytes from the lipotoxicity induced by the redistribution of Ca^{2+} between the ER and mitochondria as a result of saturated FA [462]. The capacity of the mitochondria to manage this increased Ca^{2+} concentration may alter the balance between the pro- and anti-apoptotic signals contained within the mitochondria. Another factor contributing to mitochondrial stress may be the excess availability of electrons donated by NADH and FADH_2 produced through β -oxidation and the Krebs cycle. Uncoupling of the electron transfer chain and thus oxidative phosphorylation has been proposed as a protective mechanism to increase proton leakage out of the mitochondria and limit the production of ROS that accompanies increased oxidative phosphorylation. Failure of such a protective mechanism could lead to increased ROS production and lipid peroxidation characteristic of the second hit in the progression of NAFLD.

6.7. Conclusions

Treatment of HepG2 cells with different concentrations of FA mixture produced a dose-dependent increase in lipid accumulation and a concomitant modest decrease in cell proliferation. A 6 h treatment with 1 mM FA mixture induced a 40 % increase in lipid accumulation compared to control cells. The changes that accompany lipid accumulation include an increase in the number of LDs and altered abundance of proteins that affect the integrity and smooth functioning of the ER and the mitochondria. The growth of LDs may depend on sustained *de novo* synthesis of endogenous FA in spite of the availability of exogenous FA. Reduced abundance of Ca²⁺-binding proteins in the ER suggests that maintenance of Ca²⁺ concentration may determine the extent to which ER stress persists and propagates to the mitochondria.

6.8. Future work

Given the health burden associated with NAFLD, there has been a sustained interest in the molecular mechanisms involved in the disease. Blocking DGAT activity to prevent TAG formation is an approach that is of particular interest to this project. Key features of the

enzymes including their substrate preferences and the non-redundancy in their function have been described in literature, and the design of small-molecule inhibitors to block these enzymes has not only led to clinical trials of the drugs but also helped to show the importance of these enzymes in lipid metabolism. These small molecule inhibitors, however, were not very effective in the HepG2 cell culture system used in the current study. The current study differs from the reports describing beneficial effects of small-molecule inhibitors in experimental design as has been discussed and in this light, it may be beneficial to use an alternative to the inhibitors, such as the use of anti-sense oligonucleotides to suppress DGAT gene expression for future works that include treatment of cells with concentrations of FA in the millimolar range. An understanding of enzyme kinetics and the effect of DGAT inhibitors on the kinetics would serve academic purpose and help in characterizing the overt and latent DGAT activity observed in microsomal systems.

The proteomics analysis of FA-treated cells suggested that the molecular mechanisms in the ER and mitochondria are critical for the protection or aggravation of the disease state. The disturbances in the ER described in this project included altered abundances of chaperones which would affect protein processing in the ER and possibly compromise the ER membrane. A compromised ER membrane in turn would affect Ca^{2+} signalling between the ER and the mitochondria. These associations could be verified by studies that include:

- Identification and validation of proteins involved in the UPR. Kennedy *et al.* have described two approaches to measuring UPR signalling. The first approach uses qPCR to detect and quantify transcripts of UPR-induced genes, and the second approach uses Western blot-based analysis to study the UPR-induced proteins [463]. Given the limited success of Western blots in this project, selected reaction monitoring may also be considered for the qualitative and quantitative assessment of UPR-induced proteins. Among the different measures of ER stress discussed by Osowski and Urano, UPR activation could be measured through Western blot-based analysis of IRE1 α phosphorylation or qPCR-based detection of spliced XBP1 mRNA [464]. XBP1 is a transcription factor that acts down-stream of IRE1 α in the UPR cascade. PDI and Grp78 identified in this project have also been suggested as down-stream markers of UPR.
- Ca^{2+} signalling between the ER and the mitochondrion has been well-studied. Deniaud *et al.* suggested that sustained accumulation of Ca^{2+} in the mitochondria on account of ER stress can lead to apoptosis [465]. A rapid increase in Ca^{2+} did not produce a similar response. Their study described the events in HeLa and colon carcinoma (HCT1160) cells

in which ER-stress was induced by thapsigargin and tunicamycin. Assuming that the changes in ER chaperones seen in the present study indicate possible ER-stress, a correlation between FA-induced ER stress and altered Ca^{2+} signalling as described by Deniaud *et al.* could further define the progression from NAFLD to NASH.

- Oxidative stress and an altered redox environment have been cited as a component of the ‘second-hit’ that aggravates steatosis to NASH [466]. The uncoupling of oxidative phosphorylation and the electron transport chain has been suggested to serve as a protective mechanism to reduce ROS production and mitigate the oxidative stress. Further characterization of the redox environment in the cell would help in understanding the contribution of oxidative stress to the aggravation of disease. The characterization could include direct measurement of ROS and lipid peroxidation products. It could also include validation of the changes observed in dataset LC1 that indicated an increase in abundance of glutathione reductase and glutaredoxin 3.
- Mega-mitochondria with crystalline inclusions are a common feature observed in liver biopsies of NASH patients [78]. Integrating the information available on Ca^{2+} signalling in the ER and ROS levels in the cell with an assessment of mitochondrial morphology would help clarify how accumulation of LDs in the cytoplasm of a hepatocyte proceeds to inflammation, cell death or progressive liver damage [37]. This information would help identify potential new therapeutic approaches to control NAFLD and prevent its conversion to NASH, cirrhosis and end-stage liver failure.
- Sequential window acquisition of all theoretical mass spectra (SWATH), a technique recently developed by Huang *et al.*[467], couples data-independent acquisition with peptide spectral library match in manner similar to selected or multiple reaction monitoring methods. The method relies on a preliminary discovery run to map spectra to fragment ions and thus generate a library which is used to identify and accurately quantify target proteins. Although this technique has only been implemented in the next-generation ABSciex MS instruments, it holds the potential to conclusively validate the LC-MS/MS data generated in my experiments. Proteomics information and spectral data generated through label-free shotgun experiments described in this thesis have the potential of being the reference ion spectral libraries.

REFERENCES

- [1] James, O. F., Day, C. P., Non-alcoholic steatohepatitis (NASH): a disease of emerging identity and importance. *J Hepatol* 1998, 29, 495-501.
- [2] Bellentani, S., Marino, M., Epidemiology and natural history of non-alcoholic fatty liver disease (NAFLD). *Ann Hepatol* 2009, 8 Suppl 1, S4-8.
- [3] Ludwig, J., Viggiano, T. R., McGill, D. B., Oh, B. J., Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clinic proceedings. Mayo Clinic* 1980, 55, 434-438.
- [4] WHO Media Centre: Obesity and overweight- Fact sheet N°311. 2015.
- [5] Loomba, R., Sanyal, A. J., The global NAFLD epidemic. *Nat Rev Gastroenterol Hepatol* 2013, 10, 686-690.
- [6] Adams, L. A., Angulo, P., Lindor, K. D., Nonalcoholic fatty liver disease. *CMAJ* 2005, 172, 899-905.
- [7] Despres, J. P., Lemieux, I., Bergeron, J., Pibarot, P., *et al.*, Abdominal obesity and the metabolic syndrome: contribution to global cardiometabolic risk. *Arterioscler Thromb Vasc Biol* 2008, 28, 1039-1049.
- [8] Calle, E. E., Kaaks, R., Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms. *Nat Rev Cancer* 2004, 4, 579-591.
- [9] Alexander, J., Torbenson, M., Wu, T.-T., Yeh, M. M., Non-alcoholic fatty liver disease contributes to hepatocellular carcinoma in non-cirrhotic liver: A clinical and pathological Study. *Journal of gastroenterology and hepatology* 2013.
- [10] Vernon, G., Baranova, A., Younossi, Z. M., Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment Pharmacol Ther* 2011, 34, 274-285.
- [11] Farrell, G. C., Wong, V. W.-S., Chitturi, S., NAFLD in Asia[mdash]as common and important as in the West. *Nat Rev Gastroenterol Hepatol* 2013, *advance online publication*.
- [12] Annual Update of Key Results 2014/15: New Zealand Health Survey. Wellington: Ministry of Health. 2015.
- [13] Younossi, Z. M., Henry, L., Economic and Quality-of-Life Implications of Non-Alcoholic Fatty Liver Disease. *Pharmacoeconomics* 2015, 33, 1245-1253.
- [14] Cave, M., Deaciuc, I., Mendez, C., Song, Z., *et al.*, Nonalcoholic fatty liver disease: predisposing factors and the role of nutrition. *J Nutr Biochem* 2007, 18, 184-195.
- [15] Henao-Mejia, J., Elinav, E., Jin, C., Hao, L., *et al.*, Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity. *Nature* 2012, 482, 179-185.
- [16] Yilmaz, Y., Review article: is non-alcoholic fatty liver disease a spectrum, or are steatosis and non-alcoholic steatohepatitis distinct conditions? *Aliment Pharmacol Ther* 2012, 36, 815-823.
- [17] Chalasani, N., Younossi, Z., Lavine, J. E., Diehl, A. M., *et al.*, The diagnosis and management of non-alcoholic fatty liver disease: Practice guideline by the American Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology* 2012, 55, 2005-2023.

- [18] Fassio, E., Álvarez, E., Domínguez, N., Landeira, G., Longo, C., Natural history of nonalcoholic steatohepatitis: a longitudinal study of repeat liver biopsies. *Hepatology* 2004, 40, 820-826.
- [19] Teli, M. R., James, O. F., Burt, A. D., Bennett, M. K., Day, C. P., The natural history of nonalcoholic fatty liver: a follow-up study. *Hepatology* 1995, 22, 1714-1719.
- [20] Wong, V. W., Wong, G. L., Choi, P. C., Chan, A. W., *et al.*, Disease progression of non-alcoholic fatty liver disease: a prospective study with paired liver biopsies at 3 years. *Gut* 2010, 59, 969-974.
- [21] Ekstedt, M., Franzen, L. E., Mathiesen, U. L., Thorelius, L., *et al.*, Long-term follow-up of patients with NAFLD and elevated liver enzymes. *Hepatology* 2006, 44, 865-873.
- [22] Caldwell, S. H., Oelsner, D. H., Iezzoni, J. C., Hespenheide, E. E., *et al.*, Cryptogenic cirrhosis: clinical characterization and risk factors for underlying disease. *Hepatology* 1999, 29, 664-669.
- [23] Poonawala, A., Nair, S. P., Thuluvath, P. J., Prevalence of obesity and diabetes in patients with cryptogenic cirrhosis: a case-control study. *Hepatology* 2000, 32, 689-692.
- [24] Farrell, G. C., Larter, C. Z., Nonalcoholic fatty liver disease: from steatosis to cirrhosis. *Hepatology* 2006, 43, S99-S112.
- [25] Tiniakos, D. G., Vos, M. B., Brunt, E. M., Nonalcoholic fatty liver disease: pathology and pathogenesis. *Annu Rev Pathol* 2010, 5, 145-171.
- [26] Machado, M. V., Cortez-Pinto, H., Non-invasive diagnosis of non-alcoholic fatty liver disease. A critical appraisal. *J Hepatol* 2013, 58, 1007-1019.
- [27] Emanuele, E., Is biopsy always necessary? Toward a clinico - laboratory approach for diagnosing nonalcoholic steatohepatitis in obesity. *Hepatology* 2008, 48, 2086-2087.
- [28] Kleiner, D. E., Brunt, E. M., Van Natta, M., Behling, C., *et al.*, Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 2005, 41, 1313-1321.
- [29] Douali, N., Abdenmour, M., Zucker, J. D., Jaulent, M. C., Formalization of clinical practice guidelines: nonalcoholic steatohepatitis diagnosis model-related personalized medicine. *EJBI* 2014, 10, 6-10.
- [30] Sorrentino, P., Tarantino, G., Conca, P., Perrella, A., *et al.*, Silent non-alcoholic fatty liver disease-a clinical-histological study. *J Hepatol* 2004, 41, 751-757.
- [31] Brunt, E. M., Kleiner, D. E., Wilson, L. A., Belt, P., Neuschwander-Tetri, B. A., Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. *Hepatology* 2011, 53, 810-820.
- [32] Schuppan, D., Schattenberg, J. M., Non-alcoholic steatohepatitis: pathogenesis and novel therapeutic approaches. *J Gastroenterol Hepatol* 2013, 28 Suppl 1, 68-76.
- [33] Dongiovanni, P., Anstee, Q. M., Valenti, L., Genetic predisposition in NAFLD and NASH: impact on severity of liver disease and response to treatment. *Curr Pharm Des* 2013, 19, 5219-5238.
- [34] Day, C. P., Genes or environment to determine alcoholic liver disease and non-alcoholic fatty liver disease. *Liver Int* 2006, 26, 1021-1028.
- [35] McCullough, A. J., Pathophysiology of nonalcoholic steatohepatitis. *J Clin Gastroenterol* 2006, 40 Suppl 1, S17-29.

- [36] Day, C. P., Pathogenesis of steatohepatitis. *Best Pract Res Clin Gastroenterol* 2002, 16, 663-678.
- [37] Tilg, H., Moschen, A. R., Evolution of inflammation in nonalcoholic fatty liver disease: the multiple parallel hits hypothesis. *Hepatology* 2010, 52, 1836-1846.
- [38] Diraison, F., Moulin, P., Beylot, M., Contribution of hepatic de novo lipogenesis and reesterification of plasma non esterified fatty acids to plasma triglyceride synthesis during non-alcoholic fatty liver disease. *Diabetes Metab* 2003, 29, 478-485.
- [39] Donnelly, K. L., Smith, C. I., Schwarzenberg, S. J., Jessurun, J., *et al.*, Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *J Clin Invest* 2005, 115, 1343-1351.
- [40] Fabbrini, E., Sullivan, S., Klein, S., Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications. *Hepatology* 2010, 51, 679-689.
- [41] Nolan, C. J., Larter, C. Z., Lipotoxicity: why do saturated fatty acids cause and monounsaturates protect against it? *J Gastroenterol Hepatol* 2009, 24, 703-706.
- [42] Yki-Jarvinen, H., Nutritional modulation of non-alcoholic fatty liver disease and insulin resistance. *Nutrients* 2015, 7, 9127-9138.
- [43] Dhibi, M., Brahmi, F., Mnari, A., Houas, Z., *et al.*, The intake of high fat diet with different trans fatty acid levels differentially induces oxidative stress and non alcoholic fatty liver disease (NAFLD) in rats. *Nutr Metab (Lond)* 2011, 8, 65.
- [44] Araya, J., Rodrigo, R., Videla, L. A., Thielemann, L., *et al.*, Increase in long-chain polyunsaturated fatty acid n - 6/n - 3 ratio in relation to hepatic steatosis in patients with non-alcoholic fatty liver disease. *Clin Sci (Lond)* 2004, 106, 635-643.
- [45] Allard, J. P., Aghdassi, E., Mohammed, S., Raman, M., *et al.*, Nutritional assessment and hepatic fatty acid composition in non-alcoholic fatty liver disease (NAFLD): a cross-sectional study. *J Hepatol* 2008, 48, 300-307.
- [46] Silbernagel, G., Machann, J., Unmuth, S., Schick, F., *et al.*, Effects of 4-week very-high-fructose/glucose diets on insulin sensitivity, visceral fat and intrahepatic lipids: an exploratory trial. *Br J Nutr* 2011, 106, 79-86.
- [47] Softic, S., Cohen, D. E., Kahn, C. R., Role of dietary fructose and hepatic de novo lipogenesis in fatty liver disease. *Dig Dis Sci* 2016, 61, 1282-1293.
- [48] Santoro, N., Kursawe, R., D'Adamo, E., Dykas, D. J., *et al.*, A common variant in the patatin-like phospholipase 3 gene (PNPLA3) is associated with fatty liver disease in obese children and adolescents. *Hepatology* 2010, 52, 1281-1290.
- [49] Santoro, N., Amato, A., Grandone, A., Brienza, C., *et al.*, Predicting metabolic syndrome in obese children and adolescents: look, measure and ask. *Obes Facts* 2013, 6, 48-56.
- [50] Jou, J., Choi, S. S., Diehl, A. M., Mechanisms of disease progression in nonalcoholic fatty liver disease. *Semin Liver Dis* 2008, 28, 370-379.
- [51] Wree, A., Broderick, L., Canbay, A., Hoffman, H. M., Feldstein, A. E., From NAFLD to NASH to cirrhosis-new insights into disease mechanisms. *Nat Rev Gastroenterol Hepatol* 2013, 10, 627-636.
- [52] Leamy, A. K., Egnatchik, R. A., Young, J. D., Molecular mechanisms and the role of saturated fatty acids in the progression of non-alcoholic fatty liver disease. *Prog Lipid Res* 2012.

- [53] Listenberger, L. L., Han, X., Lewis, S. E., Cases, S., *et al.*, Triglyceride accumulation protects against fatty acid-induced lipotoxicity. *Proc Natl Acad Sci U S A* 2003, *100*, 3077-3082.
- [54] Murphy, D. J., Vance, J., Mechanisms of lipid-body formation. *Trends Biochem Sci* 1999, *24*, 109-115.
- [55] Murphy, S., Martin, S., Parton, R. G., Quantitative analysis of lipid droplet fusion: inefficient steady state fusion but rapid stimulation by chemical fusogens. *PLoS One* 2010, *5*, e15030.
- [56] Xu, N., Zhang, S. O., Cole, R. A., McKinney, S. A., *et al.*, The FATP1-DGAT2 complex facilitates lipid droplet expansion at the ER-lipid droplet interface. *J Cell Biol* 2012, *198*, 895-911.
- [57] Singh, R., Cuervo, A. M., Lipophagy: connecting autophagy and lipid metabolism. *Int J Cell Biol* 2012, *2012*, 282041.
- [58] Rambold, A. S., Cohen, S., Lippincott-Schwartz, J., Fatty acid trafficking in starved cells: regulation by lipid droplet lipolysis, autophagy, and mitochondrial fusion dynamics. *Dev Cell* 2015, *32*, 678-692.
- [59] Wilfling, F., Wang, H., Haas, J. T., Krahmer, N., *et al.*, Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocalizing from the ER to lipid droplets. *Dev Cell* 2013, *24*, 384-399.
- [60] Wilfling, F., Thiam, A. R., Olarte, M. J., Wang, J., *et al.*, Arf1/COPI machinery acts directly on lipid droplets and enables their connection to the ER for protein targeting. *Elife* 2014, *3*, e01607.
- [61] Guo, Y., Walther, T. C., Rao, M., Stuurman, N., *et al.*, Functional genomic screen reveals genes involved in lipid-droplet formation and utilization. *Nature* 2008, *453*, 657-661.
- [62] Farese, R. V., Jr., Walther, T. C., Lipid droplets finally get a little R-E-S-P-E-C-T. *Cell* 2009, *139*, 855-860.
- [63] Greenberg, A. S., Coleman, R. A., Kraemer, F. B., McManaman, J. L., *et al.*, The role of lipid droplets in metabolic disease in rodents and humans. *J Clin Invest* 2011, *121*, 2102-2110.
- [64] Basseri, S., Austin, R. C., Endoplasmic reticulum stress and lipid metabolism: mechanisms and therapeutic potential. *Biochem Res Int* 2012, *2012*, 841362.
- [65] Kammoun HL, C. H., Hainault I, Luquet S, Magnan C, Koike T, Ferré P, Fofelle F., GRP78 expression inhibits insulin and ER stress-induced SREBP-1c activation and reduces hepatic steatosis in mice. *J Clin Invest.* 2009, *119*, 1201-1215.
- [66] Lei, X., Zhang, S., Barbour, S. E., Bohrer, A., *et al.*, Spontaneous development of endoplasmic reticulum stress that can lead to diabetes mellitus is associated with higher calcium-independent phospholipase A2 expression: a role for regulation by SREBP-1. *J Biol Chem* 2010, *285*, 6693-6705.
- [67] Velazquez, A. P., Graef, M., Autophagy regulation depends on ER homeostasis controlled by lipid droplets. *Autophagy* 2016, *0*.
- [68] Cao, J., Dai, D. L., Yao, L., Yu, H. H., *et al.*, Saturated fatty acid induction of endoplasmic reticulum stress and apoptosis in human liver cells via the PERK/ATF4/CHOP signaling pathway. *Mol Cell Biochem* 2012, *364*, 115-129.

- [69] Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P., Ron, D., Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. *Nat Cell Biol* 2000, 2, 326-332.
- [70] Shen, J., Chen, X., Hendershot, L., Prywes, R., ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of Golgi localization signals. *Dev Cell* 2002, 3, 99-111.
- [71] Macer, D. R., Koch, G. L., Identification of a set of calcium-binding proteins in reticuloplasm, the luminal content of the endoplasmic reticulum. *J Cell Sci* 1988, 91 (Pt 1), 61-70.
- [72] Michalak, M., Robert Parker, J. M., Opas, M., Ca²⁺ signaling and calcium binding chaperones of the endoplasmic reticulum. *Cell Calcium* 2002, 32, 269-278.
- [73] Scorrano, L., Oakes, S. A., Opferman, J. T., Cheng, E. H., *et al.*, BAX and BAK regulation of endoplasmic reticulum Ca²⁺: a control point for apoptosis. *Science* 2003, 300, 135-139.
- [74] Pinton, P., Giorgi, C., Siviero, R., Zecchini, E., Rizzuto, R., Calcium and apoptosis: ER-mitochondria Ca²⁺ transfer in the control of apoptosis. *Oncogene* 2008, 27, 6407-6418.
- [75] Puri, P., Mirshahi, F., Cheung, O., Natarajan, R., *et al.*, Activation and dysregulation of the unfolded protein response in nonalcoholic fatty liver disease. *Gastroenterology* 2008, 134, 568-576.
- [76] Schattenberg, J. M., Singh, R., Wang, Y., Lefkowitz, J. H., *et al.*, JNK1 but not JNK2 promotes the development of steatohepatitis in mice. *Hepatology* 2006, 43, 163-172.
- [77] Begrich, K., Igoudjil, A., Pessayre, D., Fromenty, B., Mitochondrial dysfunction in NASH: causes, consequences and possible means to prevent it. *Mitochondrion* 2006, 6, 1-28.
- [78] Caldwell, S. H., Swerdlow, R. H., Khan, E. M., Iezzoni, J. C., *et al.*, Mitochondrial abnormalities in non-alcoholic steatohepatitis. *J Hepatol* 1999, 31, 430-434.
- [79] Fabbrini, E., Mohammed, B. S., Magkos, F., Korenblat, K. M., *et al.*, Alterations in adipose tissue and hepatic lipid kinetics in obese men and women with nonalcoholic fatty liver disease. *Gastroenterology* 2008, 134, 424-431.
- [80] Figueira, T. R., Barros, M. H., Camargo, A. A., Castilho, R. F., *et al.*, Mitochondria as a source of reactive oxygen and nitrogen species: from molecular mechanisms to human health. *Antioxid Redox Signal* 2013, 18, 2029-2074.
- [81] Forkink, M., Smeitink, J. A., Brock, R., Willems, P. H., Koopman, W. J., Detection and manipulation of mitochondrial reactive oxygen species in mammalian cells. *Biochim Biophys Acta* 2010, 1797, 1034-1044.
- [82] Niki, E., Do antioxidants impair signaling by reactive oxygen species and lipid oxidation products? *FEBS letters* 2012, 586, 3767-3770.
- [83] Browning, J. D., Horton, J. D., Molecular mediators of hepatic steatosis and liver injury. *J Clin Invest* 2004, 114, 147-152.
- [84] Malaguarnera, M., Di Rosa, M., Nicoletti, F., Malaguarnera, L., Molecular mechanisms involved in NAFLD progression. *J Mol Med (Berl)* 2009, 87, 679-695.
- [85] Robbins, A. L., Savage, D. B., The genetics of lipid storage and human lipodystrophies. *Trends Mol Med* 2015, 21, 433-438.

- [86] Shaikh, S. R., Kinnun, J. J., Leng, X., Williams, J. A., Wassall, S. R., How polyunsaturated fatty acids modify molecular organization in membranes: insight from NMR studies of model systems. *Biochim Biophys Acta* 2015, 1848, 211-219.
- [87] Hoglinger, D., Nadler, A., Schultz, C., Caged lipids as tools for investigating cellular signaling. *Biochim Biophys Acta* 2014, 1841, 1085-1096.
- [88] Arner, P., Bernard, S., Salehpour, M., Possnert, G., *et al.*, Dynamics of human adipose lipid turnover in health and metabolic disease. *Nature* 2011, 478, 110-113.
- [89] Chavez, J. A., Summers, S. A., Lipid oversupply, selective insulin resistance, and lipotoxicity: molecular mechanisms. *Biochim Biophys Acta* 2010, 1801, 252-265.
- [90] Weiss, S. B., Kennedy, E. P., Kiyasu, J. Y., The enzymatic synthesis of triglycerides. *J Biol Chem* 1960, 235, 40-44.
- [91] Coleman, R. A., Haynes, E. B., Hepatic monoacylglycerol acyltransferase. Characterization of an activity associated with the suckling period in rats. *J Biol Chem* 1984, 259, 8934-8938.
- [92] Coleman, R. A., Lee, D. P., Enzymes of triacylglycerol synthesis and their regulation. *Progress in lipid research* 2004, 43, 134-176.
- [93] Yen, C. L., Stone, S. J., Koliwad, S., Harris, C., Farese, R. V., Jr., Thematic review series: glycerolipids. DGAT enzymes and triacylglycerol biosynthesis. *J Lipid Res* 2008, 49, 2283-2301.
- [94] Turchetto-Zolet, A. C., Maraschin, F. S., de Moraes, G. L., Cagliari, A., *et al.*, Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. *BMC Evol Biol* 2011, 11, 263.
- [95] Thompson, J. D., Higgins, D. G., Gibson, T. J., CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994, 22, 4673-4680.
- [96] Felsenstein, J., PHYLIP-phylogeny inference package (version 3.2). *Cladistics* 1989, 5, 164-166.
- [97] Cases, S., Smith, S. J., Zheng, Y. W., Myers, H. M., *et al.*, Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis. *Proc Natl Acad Sci U S A* 1998, 95, 13018-13023.
- [98] Flicek, P., Amode, M. R., Barrell, D., Beal, K., *et al.*, Ensembl 2012. *Nucleic Acids Res* 2012, 40, D84-90.
- [99] Turchetto-Zolet, A. C., Maraschin, F. S., de Moraes, G. L., Cagliari, A., *et al.*, Evolutionary view of acyl-CoA diacylglycerol acyltransferase (DGAT), a key enzyme in neutral lipid biosynthesis. *BMC evolutionary biology* 2011, 11, 263.
- [100] Cases, S., Stone, S. J., Zhou, P., Yen, E., *et al.*, Cloning of DGAT2, a second mammalian diacylglycerol acyltransferase, and related family members. *J Biol Chem* 2001, 276, 38870-38876.
- [101] Yen, C. L., Stone, S. J., Cases, S., Zhou, P., Farese, R. V., Jr., Identification of a gene encoding MGAT1, a monoacylglycerol acyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* 2002, 99, 8512-8517.
- [102] Yen, C. L., Farese, R. V., Jr., MGAT2, a monoacylglycerol acyltransferase expressed in the small intestine. *The Journal of biological chemistry* 2003, 278, 18532-18537.

- [103] Turkish, A. R., Henneberry, A. L., Cromley, D., Padamsee, M., *et al.*, Identification of two novel human acyl-CoA wax alcohol acyltransferases: members of the diacylglycerol acyltransferase 2 (DGAT2) gene superfamily. *The Journal of biological chemistry* 2005, *280*, 14755-14764.
- [104] Winter, A., van Ekeveld, M., Bininda-Emonds, O. R., Habermann, F. A., Fries, R., Genomic organization of the DGAT2/MOGAT gene family in cattle (*Bos taurus*) and other mammals. *Cytogenet Genome Res* 2003, *102*, 42-47.
- [105] Cheng, D., Nelson, T. C., Chen, J., Walker, S. G., *et al.*, Identification of acyl coenzyme A:monoacylglycerol acyltransferase 3, an intestinal specific enzyme implicated in dietary fat absorption. *J Biol Chem* 2003, *278*, 13611-13614.
- [106] Yue, Y. G., Chen, Y. Q., Zhang, Y. Y., Wang, H., *et al.*, The Acyl CoenzymeA:Monoacylglycerol Acyltransferase 3 (MGAT3) Gene is a Pseudogene in Mice but Encodes a Functional Enzyme in Rats. *Lipids* 2011, *46*, 513-520.
- [107] Cao, J., Cheng, L., Shi, Y., Catalytic properties of MGAT3, a putative triacylglycerol synthase. *J Lipid Res* 2007, *48*, 583-591.
- [108] Kantartzis, K., Machicao, F., Machann, J., Schick, F., *et al.*, The DGAT2 gene is a candidate for the dissociation between fatty liver and insulin resistance in humans. *Clin Sci (Lond)* 2009, *116*, 531-537.
- [109] Ludwig, E. H., Mahley, R. W., Palaoglu, E., Ozbayrakci, S., *et al.*, DGAT1 promoter polymorphism associated with alterations in body mass index, high density lipoprotein levels and blood pressure in Turkish women. *Clin Genet* 2002, *62*, 68-73.
- [110] Coudreau, S. K., Tounian, P., Bonhomme, G., Froguel, P., *et al.*, Role of the DGAT gene C79T single-nucleotide polymorphism in French obese subjects. *Obes Res* 2003, *11*, 1163-1167.
- [111] Haas, J. T., Winter, H. S., Lim, E., Kirby, A., *et al.*, DGAT1 mutation is linked to a congenital diarrheal disorder. *J Clin Invest* 2012, *122*, 4680-4684.
- [112] Lehnert, K., Ward, H., Berry, S. D., Ankersmit-Udy, A., *et al.*, Phenotypic population screen identifies a new mutation in bovine DGAT1 responsible for unsaturated milk fat. *Sci Rep* 2015, *5*, 8484.
- [113] Cao, H., Structure-function analysis of diacylglycerol acyltransferase sequences from 70 organisms. *BMC Res Notes* 2011, *4*, 249.
- [114] Sonnhammer, E. L., von Heijne, G., Krogh, A., A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* 1998, *6*, 175-182.
- [115] UniProt Consortium. UniProt: a hub for protein information. *Nucleic Acids Res* 2014, *gku989*.
- [116] McFie, P. J., Stone, S. L., Banman, S. L., Stone, S. J., Topological orientation of acyl-CoA:diacylglycerol acyltransferase-1 (DGAT1) and identification of a putative active site histidine and the role of the n terminus in dimer/tetramer formation. *J Biol Chem* 2010, *285*, 37377-37387.
- [117] Stone, S. J., Levin, M. C., Farese, R. V., Jr., Membrane topology and identification of key functional amino acid residues of murine acyl-CoA:diacylglycerol acyltransferase-2. *J Biol Chem* 2006, *281*, 40273-40282.

- [118] Kuerschner, L., Moessinger, C., Thiele, C., Imaging of lipid biosynthesis: how a neutral lipid enters lipid droplets. *Traffic* 2008, *9*, 338-352.
- [119] Jin, Y., McFie, P. J., Banman, S. L., Brandt, C., Stone, S. J., Diacylglycerol acyltransferase-2 (DGAT2) and monoacylglycerol acyltransferase-2 (MGAT2) interact to promote triacylglycerol synthesis. *J Biol Chem* 2014, *289*, 28237-28248.
- [120] Man, W. C., Miyazaki, M., Chu, K., Ntambi, J., Colocalization of SCD1 and DGAT2: implying preference for endogenous monounsaturated fatty acids in triglyceride synthesis. *J Lipid Res* 2006, *47*, 1928-1939.
- [121] Miyazaki, M., Kim, Y. C., Ntambi, J. M., A lipogenic diet in mice with a disruption of the stearoyl-CoA desaturase 1 gene reveals a stringent requirement of endogenous monounsaturated fatty acids for triglyceride synthesis. *J Lipid Res* 2001, *42*, 1018-1024.
- [122] Shockey, J. M., Gidda, S. K., Chapital, D. C., Kuan, J. C., *et al.*, Tung tree DGAT1 and DGAT2 have nonredundant functions in triacylglycerol biosynthesis and are localized to different subdomains of the endoplasmic reticulum. *Plant Cell* 2006, *18*, 2294-2313.
- [123] Villanueva, C. J., Monetti, M., Shih, M., Zhou, P., *et al.*, Specific role for acyl CoA:Diacylglycerol acyltransferase 1 (Dgat1) in hepatic steatosis due to exogenous fatty acids. *Hepatology* 2009, *50*, 434-442.
- [124] Qi, J., Lang, W., Geisler, J. G., Wang, P., *et al.*, The use of stable isotope-labeled glycerol and oleic acid to differentiate the hepatic functions of DGAT1 and -2. *J Lipid Res* 2012, *53*, 1106-1116.
- [125] Chen, H. C., Enhancing energy and glucose metabolism by disrupting triglyceride synthesis: Lessons from mice lacking DGAT1. *Nutr Metab (Lond)* 2006, *3*, 10.
- [126] Cao, J., Zhou, Y., Peng, H., Huang, X., *et al.*, Targeting Acyl-CoA:diacylglycerol acyltransferase 1 (DGAT1) with small molecule inhibitors for the treatment of metabolic diseases. *J Biol Chem* 2011, *286*, 41838-41851.
- [127] Smith, S. J., Cases, S., Jensen, D. R., Chen, H. C., *et al.*, Obesity resistance and multiple mechanisms of triglyceride synthesis in mice lacking Dgat. *Nat Genet* 2000, *25*.
- [128] Ohshiro, T., Tomoda, H., Acyltransferase inhibitors: a patent review (2010-present). *Expert Opin Ther Pat* 2015, *25*, 145-158.
- [129] Ables, G. P., Yang, K. J., Vogel, S., Hernandez-Ono, A., *et al.*, Intestinal DGAT1 deficiency reduces postprandial triglyceride and retinyl ester excursions by inhibiting chylomicron secretion and delaying gastric emptying. *J Lipid Res* 2012, *53*, 2364-2379.
- [130] Lin, H. V., Chen, D., Shen, Z., Zhu, L., *et al.*, Diacylglycerol acyltransferase-1 (DGAT1) inhibition perturbs postprandial gut hormone release. *PLoS One* 2013, *8*, e54480.
- [131] Dow, R. L., Li, J. C., Pence, M. P., Gibbs, E. M., *et al.*, Discovery of PF-04620110, a potent, selective, and orally bioavailable inhibitor of DGAT-1. *ACS Med Chem Lett* 2011, *2*, 407-412.
- [132] Denison, H., Nilsson, C., Kujacic, M., Lofgren, L., *et al.*, Proof of mechanism for the DGAT1 inhibitor AZD7687: results from a first-time-in-human single-dose study. *Diabetes Obes Metab* 2013, *15*, 136-143.
- [133] Denison, H., Nilsson, C., Lofgren, L., Himmelmann, A., *et al.*, Diacylglycerol acyltransferase 1 inhibition with AZD7687 alters lipid handling and hormone secretion in the gut with intolerable side effects: a randomized clinical trial. *Diabetes Obes Metab* 2014, *16*, 334-343.

- [134] Meyers, C., Gaudet, D., Tremblay, K., Amer, A., *et al.*, The DGAT1 inhibitor LCQ908 decreases triglyceride levels in patients with the familial chylomicronemia syndrome. *J Clin Lipidol* 2012, 6, 266-267.
- [135] Stone, S. J., Myers, H. M., Watkins, S. M., Brown, B. E., *et al.*, Lipopenia and skin barrier abnormalities in DGAT2-deficient mice. *J Biol Chem* 2004, 279, 11767-11776.
- [136] Wakimoto, K., Chiba, H., Michibata, H., Seishima, M., *et al.*, A novel diacylglycerol acyltransferase (DGAT2) is decreased in human psoriatic skin and increased in diabetic mice. *Biochem Biophys Res Commun* 2003, 310, 296-302.
- [137] Yu, X. X., Murray, S. F., Pandey, S. K., Booten, S. L., *et al.*, Antisense oligonucleotide reduction of DGAT2 expression improves hepatic steatosis and hyperlipidemia in obese mice. *Hepatology* 2005, 42, 362-371.
- [138] Wurie, H. R., Buckett, L., Zammit, V. A., Diacylglycerol acyltransferase 2 acts upstream of diacylglycerol acyltransferase 1 and utilizes nascent diglycerides and de novo synthesized fatty acids in HepG2 cells. *FEBS J* 2012, 279, 3033-3047.
- [139] Kim, M. O., Lee, S. U., Lee, H. J., Choi, K., *et al.*, Identification and validation of a selective small molecule inhibitor targeting the diacylglycerol acyltransferase 2 activity. *Biol Pharm Bull* 2013, 36, 1167-1173.
- [140] Imbriglio, J. E., Shen, D. M., Liang, R., Marby, K., *et al.*, Discovery and Pharmacology of a Novel Class of Diacylglycerol Acyltransferase 2 Inhibitors. *J Med Chem* 2015, 58, 9345-9353.
- [141] Eissing, L., Scherer, T., Tödter, K., Knippschild, U., *et al.*, De novo lipogenesis in human fat and liver is linked to ChREBP- β and metabolic health. *Nat Commun* 2013, 4, 1528.
- [142] Docke, S., Lock, J. F., Birkenfeld, A. L., Hoppe, S., *et al.*, Elevated hepatic chemerin mRNA expression in human non-alcoholic fatty liver disease. *Eur J Endocrinol* 2013, 169, 547-557.
- [143] Murphy, S. K., Yang, H. N., Moylan, C. A., Pang, H., *et al.*, Relationship Between Methylome and Transcriptome in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* 2013, 145, 1076-1087.
- [144] Moylan, C. A., Pang, H., Dellinger, A., Suzuki, A., *et al.*, Hepatic gene expression profiles differentiate presymptomatic patients with mild versus severe nonalcoholic fatty liver disease. *Hepatology* 2014, 59, 471-482.
- [145] Rockey, D. C., Caldwell, S. H., Goodman, Z. D., Nelson, R. C., Smith, A. D., Liver biopsy. *Hepatology* 2009, 49, 1017-1044.
- [146] Wieckowska, A., Feldstein, A. E., *Seminars in liver disease* 2008, p. 386.
- [147] Havel, R. J., Kane, J. P., Balasse, E. O., Segel, N., Basso, L. V., Splanchnic metabolism of free fatty acids and production of triglycerides of very low density lipoproteins in normotriglyceridemic and hypertriglyceridemic humans. *J Clin Invest* 1970, 49, 2017-2035.
- [148] Westerbacka, J., Kotronen, A., Fielding, B. A., Wahren, J., *et al.*, Splanchnic balance of free fatty acids, endocannabinoids, and lipids in subjects with nonalcoholic fatty liver disease. *Gastroenterology* 2010, 139, 1961-1971 e1961.
- [149] Iozzo, P., Bucci, M., Roivainen, A., Nagren, K., *et al.*, Fatty Acid Metabolism in the Liver, Measured by Positron Emission Tomography, Is Increased in Obese Individuals. *Gastroenterology* 2010, 139, 846-U203.

- [150] Hebbard, L., George, J., Animal models of nonalcoholic fatty liver disease. *Nat Rev Gastroenterol Hepatol* 2011, 8, 35-44.
- [151] Schattenberg, J. M., Galle, P. R., Animal Models of Non-Alcoholic Steatohepatitis: Of Mice and Man. *Digestive Diseases* 2010, 28, 247-254.
- [152] Larter, C. Z., Yeh, M. M., Animal models of NASH: getting both pathology and metabolic context right. *J Gastroenterol Hepatol* 2008, 23, 1635-1648.
- [153] Machado, M. L. V. M., Diehl, A. M., in: Chalasani, N., Szabo, S. (Eds.), *Alcoholic and Non-Alcoholic Fatty Liver Disease*, Springer International Publishing. 2016, pp. 121-145.
- [154] Oschry, Y., Eisenberg, S., Rat plasma lipoproteins: re-evaluation of a lipoprotein system in an animal devoid of cholesteryl ester transfer activity. *J Lipid Res* 1982, 23, 1099-1106.
- [155] Warden, C. H., Fidler, J. S., Comparisons of diets used in animal models of high-fat feeding. *Cell Metab* 2008, 7, 277.
- [156] Zimmerman, H. J., *Hepatotoxicity: the adverse effects of drugs and other chemicals on the liver*, Lippincott Williams & Wilkins 1999.
- [157] Spicakova, A., Anzenbacher, P., Liskova, B., Kuca, K., et al., Evaluation of possible inhibition of human liver drug metabolizing cytochromes P450 by two new acetylcholinesterase oxime-type reactivators. *Food Chem Toxicol* 2016, 88, 100-104.
- [158] Wrighton, S. A., Stevens, J. C., The human hepatic cytochromes P450 involved in drug metabolism. *Crit Rev Toxicol* 1992, 22, 1-21.
- [159] Rouiller, C., *The liver: morphology, biochemistry, physiology*, Academic Press 2013.
- [160] Postic, C., Dentin, R., Girard, J., Role of the liver in the control of carbohydrate and lipid homeostasis. *Diabetes Metab* 2004, 30, 398-408.
- [161] Doege, H., Baillie, R. A., Ortegon, A. M., Tsang, B., et al., Targeted deletion of FATP5 reveals multiple functions in liver metabolism: alterations in hepatic lipid homeostasis. *Gastroenterology* 2006, 130, 1245-1258.
- [162] Cohen, D. E., Fisher, E. A., *Seminars in liver disease*, NIH Public Access 2013, p. 380.
- [163] Stefan, N., Haring, H. U., The role of hepatokines in metabolism. *Nat Rev Endocrinol* 2013, 9, 144-152.
- [164] Madrigal-Matute, J., Cuervo, A. M., Regulation of liver metabolism by autophagy. *Gastroenterology* 2016, 150, 328-339.
- [165] Emanuelli, B., Vienberg, S. G., Smyth, G., Cheng, C., et al., Interplay between FGF21 and insulin action in the liver regulates metabolism. *J Clin Invest* 2014, 124, 515-527.
- [166] Kmiec, Z., Cooperation of liver cells in health and disease. *Adv Anat Embryol Cell Biol* 2001, 161, III-XIII, 1-151.
- [167] Bhogal, R. H., Hodson, J., Bartlett, D. C., Weston, C. J., et al., Isolation of primary human hepatocytes from normal and diseased liver tissue: a one hundred liver experience. *PLoS One* 2011, 6, e18222.
- [168] Rodriguez-Antona, C., Donato, M. T., Boobis, A., Edwards, R. J., et al., Cytochrome P450 expression in human hepatocytes and hepatoma cell lines: molecular mechanisms that determine lower expression in cultured cells. *Xenobiotica* 2002, 32, 505-520.

- [169] Katsura, N., Ikai, I., Mitaka, T., Shiotani, T., *et al.*, Long-term culture of primary human hepatocytes with preservation of proliferative capacity and differentiated functions. *J Surg Res* 2002, 106, 115-123.
- [170] Adams, R. M., Wang, M., Crane, A. M., Brown, B., *et al.*, Effective cryopreservation and long-term storage of primary human hepatocytes with recovery of viability, differentiation, and replicative potential. *Cell Transplant* 1995, 4, 579-586.
- [171] Bouma, M. E., Rogier, E., Verthier, N., Labarre, C., Feldmann, G., Further cellular investigation of the human hepatoblastoma-derived cell line HepG2: morphology and immunocytochemical studies of hepatic-secreted proteins. *In Vitro Cell Dev Biol* 1989, 25, 267-275.
- [172] Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., *et al.*, From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y)* 1996, 14, 61-65.
- [173] Gstaiger, M., Aebersold, R., Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nat Rev Genet* 2009, 10, 617-627.
- [174] Karp, N. A., Lilley, K. S., Design and analysis issues in quantitative proteomics studies. *Proteomics* 2007, 7 Suppl 1, 42-50.
- [175] Marouga, R., David, S., Hawkins, E., The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal Bioanal Chem* 2005, 382, 669-678.
- [176] Tannu, N. S., Hemby, S. E., Two-dimensional fluorescence difference gel electrophoresis for comparative proteomics profiling. *Nat Protoc* 2006, 1, 1732-1742.
- [177] Wijte, D., de Jong, A. L., Mol, M. A., van Baar, B. L., Heck, A. J., ProteomIQ blue, a potent post-stain for the visualization and subsequent mass spectrometry based identification of fluorescent stained proteins on 2D-gels. *J Proteome Res* 2006, 5, 2033-2038.
- [178] Gundry, R. L., White, M. Y., Murray, C. I., Kane, L. A., *et al.*, Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow. *Curr Protoc Mol Biol* 2009, Chapter 10, Unit10 25.
- [179] Canas, B., Pineiro, C., Calvo, E., Lopez-Ferrer, D., Gallardo, J. M., Trends in sample preparation for classical and second generation proteomics. *J Chromatogr A* 2007, 1153, 235-258.
- [180] Giansanti, P., Tsiatsiani, L., Low, T. Y., Heck, A. J., Six alternative proteases for mass spectrometry-based proteomics beyond trypsin. *Nat Protoc* 2016, 11, 993-1006.
- [181] Shi, Y., Xiang, R., Horvath, C., Wilkins, J. A., The role of liquid chromatography in proteomics. *J Chromatogr A* 2004, 1053, 27-36.
- [182] Sander, L. C., Wise, S. A., Synthesis and Characterization of Polymeric-C18 Stationary Phases for Liquid-Chromatography. *Analytical Chemistry* 1984, 56, 504-510.
- [183] Garcia, M. C., The effect of the mobile phase additives on sensitivity in the analysis of peptides and proteins by high-performance liquid chromatography-electrospray mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, 825, 111-123.
- [184] Yates, J. R., Ruse, C. I., Nakorchevsky, A., Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng* 2009, 11, 49-79.
- [185] Loo, J. A., Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom Rev* 1997, 16, 1-23.

- [186] Hardman, M., Makarov, A. A., Interfacing the orbitrap mass analyzer to an electrospray ion source. *Anal Chem* 2003, 75, 1699-1705.
- [187] Hu, Q., Noll, R. J., Li, H., Makarov, A., *et al.*, The Orbitrap: a new mass spectrometer. *J Mass Spectrom* 2005, 40, 430-443.
- [188] Yamashita, M., Fenn, J. B., Electrospray ion-source - Another variation on the free-jet theme. *J Phys Chem* 1984, 88, 4451-4459.
- [189] Wilm, M., Principles of electrospray ionization. *Mol Cell Proteomics* 2011, 10, M111009407.
- [190] Konermann, L., Ahadi, E., Rodriguez, A. D., Vahidi, S., Unraveling the mechanism of electrospray ionization. *Anal Chem* 2013, 85, 2-9.
- [191] Makarov, A., Denisov, E., Lange, O., Horning, S., Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. *J Am Soc Mass Spectrom* 2006, 17, 977-982.
- [192] Frese, C. K., Altelaar, A. F., Hennrich, M. L., Nolting, D., *et al.*, Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *J Proteome Res* 2011, 10, 2377-2388.
- [193] Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* 2004, 101, 9528-9533.
- [194] Paizs, B., Suhai, S., Fragmentation pathways of protonated peptides. *Mass Spectrom Rev* 2005, 24, 508-548.
- [195] Qian, W. J., Jacobs, J. M., Liu, T., Camp, D. G., 2nd, Smith, R. D., Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol Cell Proteomics* 2006, 5, 1727-1744.
- [196] Boutilier, K., Ross, M., Podtelejnikov, A. V., Orsi, C., *et al.*, Comparison of different search engines using validated MS/MS test datasets. *Anal Chim Acta* 2005, 534, 11-20.
- [197] Eng, J. K., McCormack, A. L., Yates, J. R., An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 1994, 5, 976-989.
- [198] Tabb, D. L., The SEQUEST family tree. *J Am Soc Mass Spectrom* 2015, 26, 1814-1819.
- [199] Perkins, D. N., Pappin, D. J. C., Creasy, D. M., Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 1999, 20, 3551-3567.
- [200] Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., *et al.*, Open mass spectrometry search algorithm. *J Proteome Res* 2004, 3, 958-964.
- [201] Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., *et al.*, Andromeda: a peptide search engine integrated into the MaxQuant environment. *J Proteome Res* 2011, 10, 1794-1805.
- [202] Craig, R., Beavis, R. C., TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* 2004, 20, 1466-1467.
- [203] Hoopmann, M. R., Moritz, R. L., Current algorithmic solutions for peptide-based proteomics data generation and identification. *Curr Opin Biotechnol* 2013, 24, 31-38.

- [204] Ding, Y., Choi, H., Nesvizhskii, A. I., Adaptive discriminant function analysis and reranking of MS/MS database search results for improved peptide identification in shotgun proteomics. *J Proteome Res* 2008, 7, 4878-4889.
- [205] Keller, A., Nesvizhskii, A. I., Kolker, E., Aebersold, R., Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 2002, 74, 5383-5392.
- [206] Kall, L., Canterbury, J. D., Weston, J., Noble, W. S., MacCoss, M. J., Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nat Methods* 2007, 4, 923-925.
- [207] Spivak, M., Weston, J., Bottou, L., Kall, L., Noble, W. S., Improvements to the percolator algorithm for Peptide identification from shotgun proteomics data sets. *J Proteome Res* 2009, 8, 3737-3745.
- [208] Elias, J. E., Gygi, S. P., Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* 2007, 4.
- [209] Elias, J. E., Gygi, S. P., Target-decoy search strategy for mass spectrometry-based proteomics. *Methods Mol Biol* 2010, 604, 55-71.
- [210] Searle, B. C., Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* 2010, 10, 1265-1269.
- [211] Cox, J., Mann, M., MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 2008, 26, 1367-1372.
- [212] Colaert, N., Barsnes, H., Vaudel, M., Helsens, K., *et al.*, Thermo-msf-parser: an open source Java library to parse and visualize Thermo Proteome Discoverer msf files. *J Proteome Res* 2011, 10, 3840-3843.
- [213] Wasinger, V. C., Zeng, M., Yau, Y., Current status and advances in quantitative proteomic mass spectrometry. *Int J Proteomics* 2013, 2013, 180605.
- [214] Beynon, R. J., Pratt, J. M., Metabolic labeling of proteins for proteomics. *Mol. Cell. Proteomics* 2005, 4, 857-872.
- [215] Ong, S.-E., Mann, M., A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protocols* 2007, 1, 2650-2660.
- [216] Krijgsveld, J., Ketting, R. F., Mahmoudi, T., Johansen, J., *et al.*, Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics. *Nat. Biotechnol.* 2003, 21, 927-931.
- [217] Qiu, Y., Sousa, E. A., Hewick, R. M., Wang, J. H., Acid-labile isotope-coded extractants: a class of reagents for quantitative mass spectrometric analysis of complex protein mixtures. *Anal. Chem.* 2002, 74, 4969-4979.
- [218] Wiese, S., Reidegeld, K. A., Meyer, H. E., Warscheid, B., Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. *Proteomics* 2007, 7, 340-350.
- [219] Meegalla, R. L., Billheimer, J. T., Cheng, D., Concerted elevation of acyl-coenzyme A : diacylglycerol acyltransferase (DGAT) activity through independent stimulation of mRNA expression of DGAT1 and DGAT2 by carbohydrate and insulin. *Biochemical and Biophysical Research Communications* 2002, 298, 317-323.

- [220] Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., *et al.*, Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 2011, *11*, 535-553.
- [221] Lundgren, D. H., Hwang, S. I., Wu, L., Han, D. K., Role of spectral counting in quantitative proteomics. *Expert Rev Proteomics* 2010, *7*, 39-53.
- [222] Chelius, D., Bondarenko, P. V., Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res* 2002, *1*, 317-323.
- [223] Megger, D. A., Bracht, T., Meyer, H. E., Sitek, B., Label-free quantification in clinical proteomics. *Biochim Biophys Acta* 2013, *1834*, 1581-1590.
- [224] Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., Kuster, B., Quantitative mass spectrometry in proteomics: a critical review. *Anal Bioanal Chem* 2007, *389*, 1017-1031.
- [225] Lessard, J. L., Two monoclonal antibodies to actin: one muscle selective and one generally reactive. *Cell Motil Cytoskeleton* 1988, *10*, 349-362.
- [226] Javitt, N. B., Hep-G2 Cells as a resource for metabolic studies - lipoprotein, cholesterol, and bile-Acids. *Faseb Journal* 1990, *4*, 161-168.
- [227] Greene, A. E., Silver, R. K., Krug, M., Coriell, L. L., Preservation of cell cultures by freezing in liquid nitrogen vapor. *Proc Soc Exp Biol Med* 1964, *116*, 462-467.
- [228] Yokoyama, W. M., Thompson, M. L., Ehrhardt, R. O., Cryopreservation and thawing of cells. *Curr Protoc Immunol* 2012, *Appendix 3*, 3G.
- [229] Spector, A. A., Structure and lipid binding properties of serum albumin. *Methods in enzymology* 1986, *128*, 320-339.
- [230] Gomez-Lechon, M. J., Donato, M. T., Martinez-Romero, A., Jimenez, N., *et al.*, A human hepatocellular in vitro model to investigate steatosis. *Chem Biol Interact* 2007, *165*, 106-116.
- [231] Yao, H. R., Liu, J., Plumeri, D., Cao, Y. B., *et al.*, Lipotoxicity in HepG2 cells triggered by free fatty acids. *Am J Transl Res* 2011, *3*, 284-291.
- [232] Kinkel, A. D., Fernyhough, M. E., Helterline, D. L., Vierck, J. L., *et al.*, Oil red-O stains non-adipogenic cells: a precautionary note. *Cytotechnology* 2004, *46*, 49-56.
- [233] Cui, W., Chen, S. L., Hu, K. Q., Quantification and mechanisms of oleic acid-induced steatosis in HepG2 cells. *Am J Transl Res* 2010, *2*, 95-104.
- [234] Mosmann, T., Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J Immunol Methods* 1983, *65*, 55-63.
- [235] Fotakis, G., Timbrell, J. A., In vitro cytotoxicity assays: comparison of LDH, neutral red, MTT and protein assay in hepatoma cell lines following exposure to cadmium chloride. *Toxicol Lett* 2006, *160*, 171-177.
- [236] McFie, P. J., Stone, S. J., A fluorescent assay to quantitatively measure in vitro acyl CoA:diacylglycerol acyltransferase activity. *J Lipid Res* 2011, *52*, 1760-1764.
- [237] Winter, D., Steen, H., Optimization of cell lysis and protein digestion protocols for the analysis of HeLa S3 cells by LC-MS/MS. *Proteomics* 2011, *11*, 4726-4730.
- [238] Geiger, T., Wehner, A., Schaab, C., Cox, J., Mann, M., Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol Cell Proteomics* 2012, *11*, M111 014050.

- [239] Rabilloud, T., Adessi, C., Giraudel, A., Lunardi, J., Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 1997, 18, 307-316.
- [240] Mast, S., Peng, L., Jordan, T. W., Flint, H., *et al.*, Proteomic analysis of membrane preparations from developing *Pinus radiata* compression wood. *Tree Physiol* 2010, 30, 1456-1468.
- [241] Trudgian, D. C., Fischer, R., Guo, X., Kessler, B. M., Mirzaei, H., GOAT--a simple LC-MS/MS gradient optimization tool. *Proteomics* 2014, 14, 1467-1471.
- [242] Kall, L., Storey, J. D., MacCoss, M. J., Noble, W. S., Posterior error probabilities and false discovery rates: two sides of the same coin. *J Proteome Res* 2008, 7, 40-44.
- [243] Wang, J., Duncan, D., Shi, Z., Zhang, B., WEB-based GENE SeT Analysis Toolkit (WebGestalt): update 2013. *Nucleic Acids Res* 2013, 41, W77-83.
- [244] Bradford, M. M., A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 1976, 72, 248-254.
- [245] Taylor, S. C., Posch, A., The design of a quantitative western blot experiment. *Biomed Res Int* 2014, 2014, 361590.
- [246] Wang, D., Gao, L., Proteomic analysis of neural differentiation of mouse embryonic stem cells. *Proteomics* 2005, 5, 4414-4426.
- [247] Schneider, C. A., Rasband, W. S., Eliceiri, K. W., NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 2012, 9, 671-675.
- [248] Ferreira, T., Rasband, W. S., ImageJ User Guide — IJ 1.46. imagej.nih.gov/ij/docs/guide/ 2010–2012.
- [249] Ohsaki, Y., Maeda, T., Fujimoto, T., Fixation and permeabilization protocol is critical for the immunolabeling of lipid droplet proteins. *Histochem Cell Biol* 2005, 124, 445-452.
- [250] Varela, L. M., Lopez, S., Ortega-Gomez, A., Bermudez, B., *et al.*, Postprandial triglyceride-rich lipoproteins regulate perilipin-2 and perilipin-3 lipid-droplet-associated proteins in macrophages. *J Nutr Biochem* 2015, 26, 327-336.
- [251] Donato, M. T., Martinez-Romero, A., Jimenez, N., Negro, A., *et al.*, Cytometric analysis for drug-induced steatosis in HepG2 cells. *Chem Biol Interact* 2009, 181, 417-423.
- [252] Javitt, N. B., HepG2 cells as a resource for metabolic studies - Lipoprotein, cholesterol, and bile acids. *Faseb J* 1990, 4, 161-168.
- [253] Wei, Y., Wang, D., Topczewski, F., Pagliassotti, M. J., Saturated fatty acids induce endoplasmic reticulum stress and apoptosis independently of ceramide in liver cells. *Am J Physiol Endocrinol Metab* 2006, 291, E275-281.
- [254] Zhao, G., Souers, A. J., Voorbach, M., Falls, H. D., *et al.*, Validation of diacylglycerolacyltransferase I as a novel target for the treatment of obesity and dyslipidemia using a potent and selective small molecule inhibitor. *J Med Chem.* 2008, 51, 380-383.
- [255] King, A. J., Segreti, J. A., Larson, K. J., Souers, A. J., *et al.*, Diacylglycerol acyltransferase 1 inhibition lowers serum triglycerides in the Zucker fatty rat and the hyperlipidemic hamster. *J Pharmacol Exp Ther* 2009, 330, 526-531.

- [256] Enayetallah, A. E., Ziemek, D., Leininger, M. T., Randhawa, R., *et al.*, Modeling the mechanism of action of a DGAT1 inhibitor using a causal reasoning platform. *PLoS One* 2011, 6, e27009.
- [257] Zhang, Y., Ren, J., Role of cardiac steatosis and lipotoxicity in obesity cardiomyopathy. *Hypertension* 2011, 57, 148-150.
- [258] Futatsugi, K., Kung, D. W., Orr, S. T., Cabral, S., *et al.*, Discovery and optimization of imidazopyridine-based inhibitors of diacylglycerol acyltransferase 2 (DGAT2). *J Med Chem* 2015, 58, 7173-7185.
- [259] Evans, J. F., Islam, S., Urade, Y., Eguchi, N., Ragolia, L., The lipocalin-type prostaglandin D2 synthase knockout mouse model of insulin resistance and obesity demonstrates early hypothalamic-pituitary-adrenal axis hyperactivity. *J Endocrinol* 2013, 216, 169-180.
- [260] Bogdanovic, E., Kraus, N., Patsouris, D., Diao, L., *et al.*, Endoplasmic reticulum stress in adipose tissue augments lipolysis. *J Cell Mol Med* 2015, 19, 82-91.
- [261] McMillian, M. K., Grant, E. R., Zhong, Z., Parker, J. B., *et al.*, Nile Red binding to HepG2 cells: an improved assay for in vitro studies of hepatosteatosis. *In Vitro Mol Toxicol* 2001, 14, 177-190.
- [262] Lin, C. L., Huang, H. C., Lin, J. K., Theaflavins attenuate hepatic lipid accumulation through activating AMPK in human HepG2 cells. *J Lipid Res* 2007, 48, 2334-2343.
- [263] Golini, C. M., Williams, B. W., Foresman, J. B., Further solvatochromic, thermochromic, and theoretical studies on Nile Red. *J Fluoresc* 1998, 8, 395-404.
- [264] Greenspan, P., Fowler, S. D., Spectrofluorometric studies of the lipid probe, Nile red. *J Lipid Res* 1985, 26, 781-789.
- [265] Greenspan, P., Mayer, E. P., Fowler, S. D., Nile red: a selective fluorescent stain for intracellular lipid droplets. *J Cell Biol* 1985, 100, 965-973.
- [266] Ramirez-Zacarias, J. L., Castro-Munozledo, F., Kuri-Harcuch, W., Quantitation of adipose conversion and triglycerides by staining intracytoplasmic lipids with Oil red O. *Histochemistry* 1992, 97, 493-497.
- [267] Mehlem, A., Hagberg, C. E., Muhl, L., Eriksson, U., Falkevall, A., Imaging of neutral lipids by oil red O for analyzing the metabolic status in health and disease. *Nat Protoc* 2013, 8, 1149-1154.
- [268] Sidir, I., Gulseven Sidir, Y., Estimation of ground and excited state dipole moments of Oil Red O by solvatochromic shift methods. *Spectrochim Acta A Mol Biomol Spectrosc* 2015, 135, 560-567.
- [269] Garbarino, J., Sturley, S. L., Saturated with fat: new perspectives on lipotoxicity. *Curr Opin Clin Nutr Metab Care* 2009, 12, 110-116.
- [270] Malhi, H., Bronk, S. F., Werneburg, N. W., Gores, G. J., Free fatty acids induce JNK-dependent hepatocyte lipoapoptosis. *J Biol Chem* 2006, 281, 12093-12101.
- [271] Zhang, Y., Xue, R., Zhang, Z., Yang, X., Shi, H., Palmitic and linoleic acids induce ER stress and apoptosis in hepatoma cells. *Lipids Health Dis* 2012, 11, 1.
- [272] King, A. J., Segreti, J. A., Larson, K. J., Souers, A. J., *et al.*, In vivo efficacy of acyl CoA: diacylglycerol acyltransferase (DGAT) 1 inhibition in rodent models of postprandial hyperlipidemia. *Eur J Pharmacol* 2010, 637, 155-161.

- [273] Tsuda, N., Kumadaki, S., Higashi, C., Ozawa, M., *et al.*, Intestine-targeted DGAT1 inhibition improves obesity and insulin resistance without skin aberrations in mice. *PLoS One* 2014, 9, e112027.
- [274] Li, C., Li, L., Lian, J., Watts, R., *et al.*, Roles of acyl-CoA:diacylglycerol acyltransferases 1 and 2 in triacylglycerol synthesis and secretion in primary hepatocytes. *Arterioscler Thromb Vasc Biol* 2015, 35, 1080-1091.
- [275] ClinicalTrials.gov, U.S. National Institutes of Health.
- [276] Wisniewski, J. R., Zougman, A., Nagaraj, N., Mann, M., Universal sample preparation method for proteome analysis. *Nat Methods* 2009, 6, 359-362.
- [277] Oliveros, J. C., VENNY. An interactive tool for comparing lists with Venn Diagrams. <http://bioinfogp.cnb.csic.es/tools/venny/index.html> 2007.
- [278] Xu, P., Duong, D. M., Peng, J., Systematical optimization of reverse-phase chromatography for shotgun proteomics. *J Proteome Res* 2009, 8, 3944-3950.
- [279] Xie, F., Smith, R. D., Shen, Y., Advanced proteomic liquid chromatography. *J Chromatogr A* 2012, 1261, 78-90.
- [280] Horvath, C. G., Lipsky, S. R., Peak capacity in chromatography. *Anal Chem* 1967, 39, 1893.
- [281] Neue, U. D., Peak capacity in unidimensional chromatography. *J Chromatogr A* 2008, 1184, 107-130.
- [282] Detergents for cell lysis and protein extraction. *Thermo Fisher Scientific Inc.* 2016, TR0019.2
- [283] Alonso, A., Urbaneja, M. A., Goni, F. M., Carmona, F. G., *et al.*, Kinetic studies on the interaction of phosphatidylcholine liposomes with Triton X-100. *Biochim Biophys Acta* 1987, 902, 237-246.
- [284] Hernandez-Borrell, J., Pons, M., Juarez, J. C., Estelrich, J., The action of Triton X-100 and sodium dodecyl sulphate on lipid layers. Effect on monolayers and liposomes. *J Microencapsul* 1990, 7, 255-259.
- [285] Lichtenberg, D., Ahyayauch, H., Goni, F. M., The mechanism of detergent solubilization of lipid bilayers. *Biophys J* 2013, 105, 289-299.
- [286] Koruk, M., Taysi, S., Savas, M. C., Yilmaz, O., *et al.*, Oxidative stress and enzymatic antioxidant status in patients with nonalcoholic steatohepatitis. *Ann Clin Lab Sci* 2004, 34, 57-62.
- [287] Pham, K., Pal, R., Qu, Y., Liu, X., *et al.*, Nuclear glutaredoxin 3 is critical for protection against oxidative stress-induced cell death. *Free Radic Biol Med* 2015, 85, 197-206.
- [288] Kajander, T., Kellosalo, J., Goldman, A., Inorganic pyrophosphatases: one substrate, three mechanisms. *FEBS Lett* 2013, 587, 1863-1869.
- [289] Augoff, K., Hryniewicz-Jankowska, A., Tabola, R., Lactate dehydrogenase 5: an old friend and a new hope in the war on cancer. *Cancer Lett* 2015, 358, 1-7.
- [290] Arner, E. S., Holmgren, A., Physiological functions of thioredoxin and thioredoxin reductase. *Eur. J. Biochem.* 2000, 267, 6102-6109.

- [291] Straub, B. K., Gyoengyoesi, B., Koenig, M., Hashani, M., *et al.*, Adipophilin/perilipin-2 as a lipid droplet-specific marker for metabolically active cells and diseases associated with metabolic dysregulation. *Histopathology* 2013, 62, 617-631.
- [292] Kaiser, K., Stelzer, G., Meisterernst, M., The coactivator p15 (PC4) initiates transcriptional activation during TFIIA-TFIID-promoter complex formation. *EMBO J* 1995, 14, 3520-3527.
- [293] Baumeister, W., Walz, J., Zuhl, F., Seemuller, E., The proteasome: paradigm of a self-compartmentalizing protease. *Cell* 1998, 92, 367-380.
- [294] Liu, D., Bienkowska, J., Petosa, C., Collier, R. J., *et al.*, Crystal structure of the zeta isoform of the 14-3-3 protein. *Nature* 1995, 376, 191-194.
- [295] Huang da, W., Sherman, B. T., Lempicki, R. A., Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009, 37, 1-13.
- [296] Khatrri, P., Sirota, M., Butte, A. J., Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS Comput Biol* 2012, 8, e1002375.
- [297] Tipney, H., Hunter, L., An introduction to effective use of enrichment analysis software. *Hum Genomics* 2010, 4, 202.
- [298] Mi, H., Poudel, S., Muruganujan, A., Casagrande, J. T., Thomas, P. D., PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Res* 2016, 44, D336-342.
- [299] Hartl, F. U., Martin, J., Neupert, W., Protein folding in the cell: the role of molecular chaperones Hsp70 and Hsp60. *Annu Rev Biophys Biomol Struct* 1992, 21, 293-322.
- [300] Zhao, H., van der Donk, W. A., Regeneration of cofactors for use in biocatalysis. *Curr Opin Biotechnol* 2003, 14, 583-589.
- [301] Stanton, R. C., Glucose-6-phosphate dehydrogenase, NADPH, and cell survival. *IUBMB Life* 2012, 64, 362-369.
- [302] Lee, G. J., Vierling, E., A small heat shock protein cooperates with heat shock protein 70 systems to reactivate a heat-denatured protein. *Plant Physiol* 2000, 122, 189-198.
- [303] Rodrigues, J. R., Couto, A., Cabezas, A., Pinto, R. M., *et al.*, Bifunctional homodimeric trikinase/FMN cyclase: contribution of protein domains to the activities of the human enzyme and molecular dynamics simulation of domain movements. *J Biol Chem* 2014, 289, 10620-10636.
- [304] Sullivan, A. C., Singh, M., Srere, P. A., Glusker, J. P., Reactivity and inhibitor potential of hydroxycitrate isomers with citrate synthase, citrate lyase, and ATP citrate lyase. *J Biol Chem* 1977, 252, 7583-7590.
- [305] Broschat, K. O., Gorka, C., Page, J. D., Martin-Berger, C. L., *et al.*, Kinetic characterization of human glutamine-fructose-6-phosphate amidotransferase I: potent feedback inhibition by glucosamine 6-phosphate. *J Biol Chem* 2002, 277, 14764-14770.
- [306] Wells, L., Hart, G. W., O-GlcNAc turns twenty: functional implications for post-translational modification of nuclear and cytosolic proteins with a sugar. *FEBS Lett* 2003, 546, 154-158.
- [307] Veerababu, G., Tang, J., Hoffman, R. T., Daniels, M. C., *et al.*, Overexpression of glutamine: fructose-6-phosphate amidotransferase in the liver of transgenic mice results in

- enhanced glycogen storage, hyperlipidemia, obesity, and impaired glucose tolerance. *Diabetes* 2000, 49, 2070-2078.
- [308] Tang, J., Neidigh, J. L., Cooksey, R. C., McClain, D. A., Transgenic mice with increased hexosamine flux specifically targeted to beta-cells exhibit hyperinsulinemia and peripheral insulin resistance. *Diabetes* 2000, 49, 1492-1499.
- [309] Sage, A. T., Walter, L. A., Shi, Y., Khan, M. I., *et al.*, Hexosamine biosynthesis pathway flux promotes endoplasmic reticulum stress, lipid accumulation, and inflammatory gene expression in hepatic cells. *Am J Physiol Endocrinol Metab* 2010, 298, E499-511.
- [310] Wakil, S. J., Fatty acid synthase, a proficient multifunctional enzyme. *Biochemistry* 1989, 28, 4523-4530.
- [311] Goodman, J. M., Demonstrated and inferred metabolism associated with cytosolic lipid droplets. *J Lipid Res* 2009, 50, 2148-2156.
- [312] Elahian, F., Sepehrizadeh, Z., Moghimi, B., Mirzaei, S. A., Human cytochrome b5 reductase: structure, function, and potential applications. *Crit Rev Biotechnol* 2014, 34, 134-143.
- [313] Mahler, H. R., Wakil, S. J., Bock, R. M., Studies on fatty acid oxidation. I. Enzymatic activation of fatty acids. *J Biol Chem* 1953, 204, 453-468.
- [314] Kurtz, D. M., Rinaldo, P., Rhead, W. J., Tian, L., *et al.*, Targeted disruption of mouse long-chain acyl-CoA dehydrogenase gene reveals crucial roles for fatty acid oxidation. *Proc Natl Acad Sci U S A* 1998, 95, 15592-15597.
- [315] Kelson, T. L., Secor McVoy, J. R., Rizzo, W. B., Human liver fatty aldehyde dehydrogenase: microsomal localization, purification, and biochemical characterization. *Biochim Biophys Acta* 1997, 1335, 99-110.
- [316] Beller, M., Sztalryd, C., Southall, N., Bell, M., *et al.*, COPI complex is a regulator of lipid homeostasis. *PLoS Biol* 2008, 6, e292.
- [317] Volpicelli-Daley, L. A., Li, Y., Zhang, C. J., Kahn, R. A., Isoform-selective effects of the depletion of ADP-ribosylation factors 1-5 on membrane traffic. *Mol Biol Cell* 2005, 16, 4495-4508.
- [318] Soni, K. G., Mardones, G. A., Sougrat, R., Smirnova, E., *et al.*, Coatamer-dependent protein delivery to lipid droplets. *J Cell Sci* 2009, 122, 1834-1841.
- [319] Olzmann, J. A., Kopito, R. R., Christianson, J. C., The mammalian endoplasmic reticulum-associated degradation system. *Cold Spring Harb Perspect Biol* 2013, 5, 10.1101/cshperspect.a013185 a013185.
- [320] Dudek, J., Pfeffer, S., Lee, P. H., Jung, M., *et al.*, Protein transport into the human endoplasmic reticulum. *J Mol Biol* 2015, 427, 1159-1175.
- [321] Tannous, A., Pisoni, G. B., Hebert, D. N., Molinari, M., N-linked sugar-regulated protein folding and quality control in the ER. *Semin Cell Dev Biol* 2015, 41, 79-89.
- [322] Bulleid, N. J., Disulfide bond formation in the mammalian endoplasmic reticulum. *Cold Spring Harb Perspect Biol* 2012, 4.
- [323] Braakman, I., Hebert, D. N., Protein folding in the endoplasmic reticulum. *Cold Spring Harb Perspect Biol* 2013, 5, a013201.
- [324] la Cour, J. M., Schindler, A. J., Berchtold, M. W., Schekman, R., ALG-2 attenuates COPII budding in vitro and stabilizes the Sec23/Sec31A complex. *PLoS One* 2013, 8, e75309.

- [325] Wallis, A. K., Freedman, R. B., Assisting oxidative protein folding: how do protein disulphide-isomerases couple conformational and chemical processes in protein folding? *Top Curr Chem* 2013, 328, 1-34.
- [326] Galligan, J. J., Petersen, D. R., The human protein disulfide isomerase gene family. *Hum Genomics* 2012, 6, 6.
- [327] Marcus, N., Shaffer, D., Farrar, P., Green, M., Tissue distribution of three members of the murine protein disulfide isomerase (PDI) family. *Biochim Biophys Acta* 1996, 1309, 253-260.
- [328] Alanen, H. I., Williamson, R. A., Howard, M. J., Lappi, A. K., *et al.*, Functional characterization of ERp18, a new endoplasmic reticulum-located thioredoxin superfamily member. *J Biol Chem* 2003, 278, 28912-28920.
- [329] Rowe, M. L., Ruddock, L. W., Kelly, G., Schmidt, J. M., *et al.*, Solution structure and dynamics of ERp18, a small endoplasmic reticulum resident oxidoreductase. *Biochemistry* 2009, 48, 4596-4606.
- [330] Oliver, J. D., van der Wal, F. J., Bulleid, N. J., High, S., Interaction of the thiol-dependent reductase ERp57 with nascent glycoproteins. *Science* 1997, 275, 86-88.
- [331] Zapun, A., Darby, N. J., Tessier, D. C., Michalak, M., *et al.*, Enhanced catalysis of ribonuclease B folding by the interaction of calnexin or calreticulin with ERp57. *J Biol Chem* 1998, 273, 6009-6012.
- [332] Winter, J., Gleiter, S., Klappa, P., Lilie, H., Protein disulfide isomerase isomerizes non-native disulfide bonds in human proinsulin independent of its peptide-binding activity. *Protein Sci* 2011, 20, 588-596.
- [333] Borradaile, N. M., Han, X., Harp, J. D., Gale, S. E., *et al.*, Disruption of endoplasmic reticulum structure and integrity in lipotoxic cell death. *J Lipid Res* 2006, 47, 2726-2737.
- [334] Li, Y., Ge, M., Ciani, L., Kuriakose, G., *et al.*, Enrichment of endoplasmic reticulum with cholesterol inhibits sarcoplasmic-endoplasmic reticulum calcium ATPase-2b activity in parallel with increased order of membrane lipids: implications for depletion of endoplasmic reticulum calcium stores and apoptosis in cholesterol-loaded macrophages. *J Biol Chem* 2004, 279, 37030-37039.
- [335] Chen, J. J., Reid, C. E., Band, V., Androphy, E. J., Interaction of papillomavirus E6 oncoproteins with a putative calcium-binding protein. *Science* 1995, 269, 529-531.
- [336] Yabe, D., Nakamura, T., Kanazawa, N., Tashiro, K., Honjo, T., Calumenin, a Ca²⁺-binding protein retained in the endoplasmic reticulum with a novel carboxyl-terminal sequence, HDEF. *J Biol Chem* 1997, 272, 18232-18239.
- [337] Kanekura, K., Nishimoto, I., Aiso, S., Matsuoka, M., Characterization of amyotrophic lateral sclerosis-linked P56S mutation of vesicle-associated membrane protein-associated protein B (VAPB/ALS8). *J Biol Chem* 2006, 281, 30223-30233.
- [338] Urano, F., Wang, X., Bertolotti, A., Zhang, Y., *et al.*, Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1. *Science* 2000, 287, 664-666.
- [339] De Vos, K. J., Morotz, G. M., Stoica, R., Tudor, E. L., *et al.*, VAPB interacts with the mitochondrial protein PTPIP51 to regulate calcium homeostasis. *Hum Mol Genet* 2012, 21, 1299-1311.

- [340] Sharma, M., Urano, F., Jaeschke, A., Cdc42 and Rac1 are major contributors to the saturated fatty acid-stimulated JNK pathway in hepatocytes. *J Hepatol* 2012, 56, 192-198.
- [341] Walker, J. E., Runswick, M. J., The mitochondrial transport protein superfamily. *J Bioenerg Biomembr* 1993, 25, 435-446.
- [342] Palmieri, L., Pardo, B., Lasorsa, F. M., del Arco, A., *et al.*, Citrin and aralar1 are Ca(2+)-stimulated aspartate/glutamate transporters in mitochondria. *EMBO J* 2001, 20, 5060-5069.
- [343] Komatsu, M., Yazaki, M., Tanaka, N., Sano, K., *et al.*, Citrin deficiency as a cause of chronic liver disorder mimicking non-alcoholic fatty liver disease. *J Hepatol* 2008, 49, 810-820.
- [344] Saheki, T., Kobayashi, K., Iijima, M., Moriyama, M., *et al.*, Metabolic derangements in deficiency of citrin, a liver-type mitochondrial aspartate-glutamate carrier. *Hepatol Res* 2005, 33, 181-184.
- [345] Fromenty, B., Pessayre, D., Inhibition of mitochondrial beta-oxidation as a mechanism of hepatotoxicity. *Pharmacology & Therapeutics* 1995, 67, 101-154.
- [346] Kornblum, C., Nicholls, T. J., Haack, T. B., Scholer, S., *et al.*, Loss-of-function mutations in MGME1 impair mtDNA replication and cause multisystemic mitochondrial disease. *Nat Genet* 2013, 45, 214-219.
- [347] Harmel, J., Ruzzenente, B., Terzioglu, M., Spahr, H., *et al.*, The leucine-rich pentatricopeptide repeat-containing protein (LRPPRC) does not activate transcription in mammalian mitochondria. *J Biol Chem* 2013, 288, 15510-15519.
- [348] Sasarman, F., Brunel-Guitton, C., Antonicka, H., Wai, T., *et al.*, LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria. *Mol Biol Cell* 2010, 21, 1315-1323.
- [349] Liu, L., Sanosaka, M., Lei, S., Bestwick, M. L., *et al.*, LRP130 protein remodels mitochondria and stimulates fatty acid oxidation. *J Biol Chem* 2011, 286, 41253-41264.
- [350] Lei, S., Sun, R. Z., Wang, D., Gong, M. Z., *et al.*, Increased hepatic fatty acids uptake and oxidation by LRPPRC-driven oxidative phosphorylation reduces blood lipid levels. *Front Physiol* 2016, 7, 270.
- [351] Taylor, S. C., Berkelman, T., Yadav, G., Hammond, M., A defined methodology for reliable quantification of Western blot data. *Mol Biotechnol* 2013, 55, 217-226.
- [352] Towbin, H., Staehelin, T., Gordon, J., Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets - procedure and some applications. *Proc Natl Acad Sci U S A* 1979, 76, 4350-4354.
- [353] Bolt, M. W., Mahoney, P. A., High-efficiency blotting of proteins of diverse sizes following sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Anal Biochem* 1997, 247, 185-192.
- [354] Bordeaux, J., Welsh, A., Agarwal, S., Killiam, E., *et al.*, Antibody validation. *Biotechniques* 2010, 48, 197-209.
- [355] Huber, M. D., Vesely, P. W., Datta, K., Gerace, L., Erlins restrict SREBP activation in the ER and regulate cellular cholesterol homeostasis. *J Cell Biol* 2013, 203, 427-436.
- [356] Pearce, M. M., Wormer, D. B., Wilkens, S., Wojcikiewicz, R. J., An endoplasmic reticulum (ER) membrane complex composed of SPFH1 and SPFH2 mediates the ER-

- associated degradation of inositol 1,4,5-trisphosphate receptors. *J Biol Chem* 2009, 284, 10433-10445.
- [357] Feitosa, M. F., Wojczynski, M. K., North, K. E., Zhang, Q., *et al.*, The ERLIN1-CHUK-CWF19L1 gene cluster influences liver fat deposition and hepatic inflammation in the NHLBI Family Heart Study. *Atherosclerosis* 2013, 228, 175-180.
- [358] Xu, Z., Page, R. C., Gomes, M. M., Kohli, E., *et al.*, Structural basis of nucleotide exchange and client binding by the Hsp70 cochaperone Bag2. *Nat Struct Mol Biol* 2008, 15, 1309-1317.
- [359] Sohal, R. S., Arnold, L., Orr, W. C., Effect of age on superoxide dismutase, catalase, glutathione reductase, inorganic peroxides, TBA-reactive material, GSH/GSSG, NADPH/NADP⁺ and NADH/NAD⁺ in *Drosophila melanogaster*. *Mech Ageing Dev* 1990, 56, 223-235.
- [360] Bottinger, L., Oeljeklaus, S., Guiard, B., Rospert, S., *et al.*, Mitochondrial heat shock protein (Hsp) 70 and Hsp10 cooperate in the formation of Hsp60 complexes. *J Biol Chem* 2015, 290, 11611-11622.
- [361] Waterman, I. J., Zammit, V. A., Differential effects of fenofibrate or simvastatin treatment of rats on hepatic microsomal overt and latent diacylglycerol acyltransferase activities. *Diabetes* 2002, 51, 1708-1713.
- [362] Owen, M. R., Corstorphine, C. C., Zammit, V. A., Overt and latent activities of diacylglycerol acyltransferase in rat liver microsomes: possible roles in very-low-density lipoprotein triacylglycerol secretion. *Biochem J* 1997, 323 (Pt 1), 17-21.
- [363] Wurie, H. R., Buckett, L., Zammit, V. A., Evidence that diacylglycerol acyltransferase 1 (DGAT1) has dual membrane topology in the endoplasmic reticulum of HepG2 cells. *J Biol Chem* 2011, 286, 36238-36247.
- [364] Waterman, I. J., Price, N. T., Zammit, V. A., Distinct ontogenic patterns of overt and latent DGAT activities of rat liver microsomes. *J Lipid Res* 2002, 43, 1555-1562.
- [365] Brasaemle, D. L., Thematic review series: adipocyte biology. The perilipin family of structural lipid droplet proteins: stabilization of lipid droplets and control of lipolysis. *J Lipid Res* 2007, 48, 2547-2559.
- [366] Heid, H. W., Moll, R., Schwetlick, I., Rackwitz, H. R., Keenan, T. W., Adipophilin is a specific marker of lipid accumulation in diverse cell types and diseases. *Cell and tissue research* 1998, 294, 309-321.
- [367] Bjellqvist, B., Hughes, G. J., Pasquali, C., Paquet, N., *et al.*, The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* 1993, 14, 1023-1031.
- [368] PhosphoSite Plus: protein page - perilipin 3 (human) <http://www.phosphosite.org/proteinAction.action?id=8568&showAllSites=true>. 2016.
- [369] Guo, X., Zhao, C., Wang, F., Zhu, Y., *et al.*, Investigation of human testis protein heterogeneity using 2-dimensional electrophoresis. *J Androl* 2010, 31, 419-429.
- [370] 'REPRODUCTION-2DPAGE' entry matching: IPI00303882; <http://reprod.njmu.edu.cn/cgi-bin/2d/2d.cgi?IPI00303882>. 2016.
- [371] Sanchez, J. C., Appel, R. D., Golaz, O., Pasquali, C., *et al.*, Inside SWISS-2DPAGE database. *Electrophoresis* 1995, 16, 1131-1151.

- [372] SWISS-2DPAGE: P60709; <http://world-2dpage.expasy.org/swiss-2dpage/protein/ac=P60709#TOP>. 2016.
- [373] Baylin, A., Kabagambe, E. K., Siles, X., Campos, H., Adipose tissue biomarkers of fatty acid intake. *Am J Clin Nutr* 2002, 76, 750-757.
- [374] Chong, M. F., Hodson, L., Bickerton, A. S., Roberts, R., *et al.*, Parallel activation of de novo lipogenesis and stearoyl-CoA desaturase activity after 3 d of high-carbohydrate feeding. *Am J Clin Nutr* 2008, 87, 817-823.
- [375] Kuo, Y. T., Lin, T. H., Chen, W. L., Lee, H. M., Alpha-lipoic acid induces adipose triglyceride lipase expression and decreases intracellular lipid accumulation in HepG2 cells. *Eur J Pharmacol* 2012, 692, 10-18.
- [376] Shi, Y., Cheng, D., Beyond triglyceride synthesis: the dynamic functional roles of MGAT and DGAT enzymes in energy metabolism. *Am J Physiol Endocrinol Metab* 2009, 297, E10-18.
- [377] Cao, J., Zhou, Y., Peng, H., Huang, X., *et al.*, Targeting Acyl-CoA:diacylglycerol acyltransferase 1 (DGAT1) with small molecule inhibitors for the treatment of metabolic diseases. *J Biol Chem* 2011, 286.
- [378] Choi, C. S., Savage, D. B., Kulkarni, A., Yu, X. X., *et al.*, Suppression of diacylglycerol acyltransferase-2 (DGAT2), but not DGAT1, with antisense oligonucleotides reverses diet-induced hepatic steatosis and insulin resistance. *J Biol Chem* 2007, 282, 22678-22688.
- [379] Han, S., Hu, L., Quach, T., Simpson, J. S., *et al.*, Profiling the role of deacylation-reacylation in the lymphatic transport of a triglyceride-mimetic prodrug. *Pharm. Res.* 2015, 32, 1830-1844.
- [380] Lambert, J. E., Ramos-Roman, M. A., Browning, J. D., Parks, E. J., Increased de novo lipogenesis is a distinct characteristic of individuals with nonalcoholic fatty liver disease. *Gastroenterology* 2014, 146, 726-735.
- [381] Seppala-Lindroos, A., Vehkavaara, S., Hakkinen, A. M., Goto, T., *et al.*, Fat accumulation in the liver is associated with defects in insulin suppression of glucose production and serum free fatty acids independent of obesity in normal men. *J Clin Endocrinol Metab* 2002, 87, 3023-3028.
- [382] Moravcova, A., Cervinkova, Z., Kucera, O., Mezera, V., *et al.*, The effect of oleic and palmitic acid on induction of steatosis and cytotoxicity on rat hepatocytes in primary culture. *Physiol Res* 2015, 64 Suppl 5, S627-636.
- [383] Ricchi, M., Odoardi, M. R., Carulli, L., Anzivino, C., *et al.*, Differential effect of oleic and palmitic acid on lipid accumulation and apoptosis in cultured hepatocytes. *J Gastroenterol Hepatol* 2009, 24, 830-840.
- [384] Tamura, S., Shimomura, I., Contribution of adipose tissue and de novo lipogenesis to nonalcoholic fatty liver disease. *J Clin Invest* 2005, 115, 1139-1142.
- [385] Yan, S., Yang, X. F., Liu, H. L., Fu, N., *et al.*, Long-chain acyl-CoA synthetase in fatty acid metabolism involved in liver and other diseases: an update. *World J Gastroenterol* 2015, 21, 3492-3498.
- [386] Stepanova, M., Hossain, N., Afendy, A., Perry, K., *et al.*, Hepatic gene expression of Caucasian and African-American patients with obesity-related non-alcoholic fatty liver disease. *Obes Surg* 2010, 20, 640-650.
- [387] Westerbacka, J., Kolak, M., Kiviluoto, T., Arkkila, P., *et al.*, Genes involved in fatty acid partitioning and binding, lipolysis, monocyte/macrophage recruitment, and inflammation

are overexpressed in the human fatty liver of insulin-resistant subjects. *Diabetes* 2007, 56, 2759-2765.

[388] Knudsen, J., Mandrup, S., Rasmussen, J. T., Andreasen, P. H., *et al.*, The function of acyl-CoA-binding protein (ACBP)/diazepam binding inhibitor (DBI). *Mol Cell Biochem* 1993, 123, 129-138.

[389] Knudsen, J., Neergaard, T. B., Gaigg, B., Jensen, M. V., Hansen, J. K., Role of acyl-CoA binding protein in acyl-CoA metabolism and acyl-CoA-mediated cell signaling. *J. Nutr.* 2000, 130, 294S-298S.

[390] Rasmussen, J. T., Rosendal, J., Knudsen, J., Interaction of acyl-CoA binding protein (ACBP) on processes for which acyl-CoA is a substrate, product or inhibitor. *Biochem. J.* 1993, 292, 907-913.

[391] Jayakumar, A., Tai, M. H., Huang, W. Y., al-Feel, W., *et al.*, Human fatty acid synthase: properties and molecular cloning. *Proc Natl Acad Sci U S A* 1995, 92, 8695-8699.

[392] Harris, C. A., Haas, J. T., Streeper, R. S., Stone, S. J., *et al.*, DGAT enzymes are required for triacylglycerol synthesis and lipid droplets in adipocytes. *J Lipid Res* 2011, 52, 657-667.

[393] Chandel, N. S., Mitochondria as signaling organelles. *BMC Biol* 2014, 12, 34.

[394] Schulz, H., Beta oxidation of fatty acids. *Biochim. Biophys. Acta.* 1991, 1081, 109-120.

[395] Begriche, K., Massart, J., Robin, M. A., Bonnet, F., Fromenty, B., Mitochondrial adaptations and dysfunctions in nonalcoholic fatty liver disease. *Hepatology* 2013, 58, 1497-1507.

[396] Fromenty, B., Pessayre, D., Inhibition of mitochondrial beta-oxidation as a mechanism of hepatotoxicity. *Pharmacol. Ther.* 1995, 67, 101-154.

[397] Icard, P., Poulain, L., Lincet, H., Understanding the central role of citrate in the metabolism of cancer cells. *Biochim Biophys Acta* 2012, 1825, 111-116.

[398] Wiegand, G., Remington, S. J., Citrate synthase: structure, control, and mechanism. *Annu Rev Biophys Biophys Chem* 1986, 15, 97-117.

[399] Lin, R., Tao, R., Gao, X., Li, T., *et al.*, Acetylation stabilizes ATP-citrate lyase to promote lipid biosynthesis and tumor growth. *Mol Cell* 2013, 51, 506-518.

[400] van de Wier, B., Balk, J. M., Haenen, G. R., Giamouridis, D., *et al.*, Elevated citrate levels in non-alcoholic fatty liver disease: the potential of citrate to promote radical production. *FEBS Lett* 2013, 587, 2461-2466.

[401] Sunny, N. E., Parks, E. J., Browning, J. D., Burgess, S. C., Excessive hepatic mitochondrial TCA cycle and gluconeogenesis in humans with nonalcoholic fatty liver disease. *Cell Metab* 2011, 14, 804-810.

[402] Schlichtholz, B., Turyn, J., Goyke, E., Biernacki, M., *et al.*, Enhanced citrate synthase activity in human pancreatic cancer. *Pancreas* 2005, 30, 99-104.

[403] Sanyal, A. J., Campbell-Sargent, C., Mirshahi, F., Rizzo, W. B., *et al.*, Nonalcoholic steatohepatitis: association of insulin resistance and mitochondrial abnormalities. *Gastroenterology* 2001, 120, 1183-1192.

[404] Vance, J. E., Phospholipid synthesis in a membrane fraction associated with mitochondria. *J Biol Chem* 1990, 265, 7248-7256.

- [405] Lewin, T. M., Kim, J. H., Granger, D. A., Vance, J. E., Coleman, R. A., Acyl-CoA synthetase isoforms 1, 4, and 5 are present in different subcellular membranes in rat liver and can be inhibited independently. *J Biol Chem* 2001, 276, 24674-24679.
- [406] Lewin, T. M., Van Horn, C. G., Krisans, S. K., Coleman, R. A., Rat liver acyl-CoA synthetase 4 is a peripheral-membrane protein located in two distinct subcellular organelles, peroxisomes, and mitochondrial-associated membrane. *Arch Biochem Biophys* 2002, 404, 263-270.
- [407] Patergnani, S., Suski, J. M., Agnoletto, C., Bononi, A., *et al.*, Calcium signaling around Mitochondria Associated Membranes (MAMs). *Cell Commun Signal* 2011, 9, 19.
- [408] Gellerich, F. N., Gizatullina, Z., Trumbeckaite, S., Nguyen, H. P., *et al.*, The regulation of OXPHOS by extramitochondrial calcium. *Biochim Biophys Acta* 2010, 1797, 1018-1027.
- [409] Gellerich, F. N., Gizatullina, Z., Arandarcikaite, O., Jerzembek, D., *et al.*, Extramitochondrial Ca²⁺ in the nanomolar range regulates glutamate-dependent oxidative phosphorylation on demand. *PLoS One* 2009, 4, e8181.
- [410] Garcia-Ruiz, C., Baulies, A., Mari, M., Garcia-Roves, P. M., Fernandez-Checa, J. C., Mitochondrial dysfunction in non-alcoholic fatty liver disease and insulin resistance: cause or consequence? *Free Radic Res* 2013, 47, 854-868.
- [411] Pizzo, P., Pozzan, T., Mitochondria-endoplasmic reticulum choreography: structure and signaling dynamics. *Trends Cell Biol* 2007, 17, 511-517.
- [412] Somlyo, A. P., Bond, M., Somlyo, A. V., Calcium content of mitochondria and endoplasmic reticulum in liver frozen rapidly in vivo. *Nature* 1985, 314, 622-625.
- [413] Corbett, E. F., Michalak, M., Calcium, a signaling molecule in the endoplasmic reticulum? *Trends Biochem Sci* 2000, 25, 307-311.
- [414] Allbritton, N. L., Meyer, T., Stryer, L., Range of messenger action of calcium ion and inositol 1,4,5-trisphosphate. *Science* 1992, 258, 1812-1815.
- [415] Parker, I., Yao, Y., Ca²⁺ transients associated with openings of inositol trisphosphate-gated channels in *Xenopus* oocytes. *J Physiol* 1996, 491 (Pt 3), 663-668.
- [416] Rizzuto, R., Pinton, P., Carrington, W., Fay, F. S., *et al.*, Close contacts with the endoplasmic reticulum as determinants of mitochondrial Ca²⁺ responses. *Science* 1998, 280.
- [417] Rizzuto, R., Brini, M., Murgia, M., Pozzan, T., Microdomains with high Ca²⁺ close to IP₃-sensitive channels that are sensed by neighboring mitochondria. *Science* 1993, 262.
- [418] Giacomello, M., Drago, I., Bortolozzi, M., Scorzeto, M., *et al.*, Ca²⁺ hot spots on the mitochondrial surface are generated by Ca²⁺ mobilization from stores, but not by activation of store-operated Ca²⁺ channels. *Mol Cell* 2010, 38, 280-290.
- [419] Csordas, G., Varnai, P., Golenar, T., Roy, S., *et al.*, Imaging interorganelle contacts and local calcium dynamics at the ER-mitochondrial interface. *Mol Cell* 2010, 39, 121-132.
- [420] Giorgi, C., Baldassari, F., Bononi, A., Bonora, M., *et al.*, Mitochondrial Ca(2+) and apoptosis. *Cell Calcium* 2012, 52, 36-43.
- [421] Pinton, P., Ferrari, D., Magalhaes, P., Schulze-Osthoff, K., *et al.*, Reduced loading of intracellular Ca²⁺ stores and downregulation of capacitative Ca²⁺ influx in Bcl-2-overexpressing cells. *J Cell Biol* 2000, 148, 857-862.

- [422] Roy, B., Lee, A. S., The mammalian endoplasmic reticulum stress response element consists of an evolutionarily conserved tripartite structure and interacts with a novel stress-inducible complex. *Nucleic Acids Res* 1999, 27, 1437-1443.
- [423] Nathan, D. F., Vos, M. H., Lindquist, S., In vivo functions of the *Saccharomyces cerevisiae* Hsp90 chaperone. *Proc Natl Acad Sci U S A* 1997, 94, 12949-12956.
- [424] Koch, G., Smith, M., Macer, D., Webster, P., Mortara, R., Endoplasmic reticulum contains a common, abundant calcium-binding glycoprotein, endoplasmin. *J Cell Sci* 1986, 86, 217-232.
- [425] Lee, A. S., The glucose-regulated proteins: stress induction and clinical applications. *Trends Biochem Sci* 2001, 26, 504-510.
- [426] Haas, I. G., Wabl, M., Immunoglobulin heavy chain binding protein. *Nature* 1983, 306, 387-389.
- [427] Thastrup, O., Cullen, P. J., Drobak, B. K., Hanley, M. R., Dawson, A. P., Thapsigargin, a tumor promoter, discharges intracellular Ca^{2+} stores by specific inhibition of the endoplasmic reticulum Ca^{2+} -ATPase. *Proc Natl Acad Sci U S A* 1990, 87, 2466-2470.
- [428] Biswas, C., Ostrovsky, O., Makarewich, C. A., Wanderling, S., *et al.*, The peptide-binding activity of GRP94 is regulated by calcium. *Biochem J* 2007, 405, 233-241.
- [429] Michalak, M., Groenendyk, J., Szabo, E., Gold, L. I., Opas, M., Calreticulin, a multi-process calcium-buffering chaperone of the endoplasmic reticulum. *Biochem J* 2009, 417, 651-666.
- [430] John, L. M., Lechleiter, J. D., Camacho, P., Differential modulation of SERCA2 isoforms by calreticulin. *J Cell Biol* 1998, 142, 963-973.
- [431] Camacho, P., Lechleiter, J. D., Calreticulin inhibits repetitive intracellular Ca^{2+} waves. *Cell* 1995, 82, 765-771.
- [432] Corbett, E. F., Oikawa, K., Francois, P., Tessier, D. C., *et al.*, Ca^{2+} regulation of interactions between endoplasmic reticulum chaperones. *J Biol Chem* 1999, 274, 6203-6211.
- [433] Vassilakos, A., Michalak, M., Lehrman, M. A., Williams, D. B., Oligosaccharide binding characteristics of the molecular chaperones calnexin and calreticulin. *Biochemistry* 1998, 37, 3480-3490.
- [434] Sprenger, R. R., Roepstorff, P., Mass Spectrometry Instrumentation in Proteomics. *eLS* 2012.
- [435] de Godoy, L. M., Olsen, J. V., Cox, J., Nielsen, M. L., *et al.*, Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 2008, 455, 1251-1254.
- [436] Graumann, J., Hubner, N. C., Kim, J. B., Ko, K., *et al.*, Stable isotope labeling by amino acids in cell culture (SILAC) and proteome quantitation of mouse embryonic stem cells to a depth of 5,111 proteins. *Mol Cell Proteomics* 2008, 7, 672-683.
- [437] Lundberg, E., Fagerberg, L., Klevebring, D., Matic, I., *et al.*, Defining the transcriptome and proteome in three functionally different human cell lines. *Mol Syst Biol* 2010, 6, 450.
- [438] Nagaraj, N., Wisniewski, J. R., Geiger, T., Cox, J., *et al.*, Deep proteome and transcriptome mapping of a human cancer cell line. *Mol Syst Biol* 2011, 7, 548.

- [439] Nagaraj, N., Kulak, N. A., Cox, J., Neuhauser, N., *et al.*, System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Mol Cell Proteomics* 2012, 11, M111 013722.
- [440] Cristobal, A., Hennrich, M. L., Giansanti, P., Goerdal, S. S., *et al.*, In-house construction of a UHPLC system enabling the identification of over 4000 protein groups in a single analysis. *Analyst* 2012, 137, 3541-3548.
- [441] Thakur, S. S., Geiger, T., Chatterjee, B., Bandilla, P., *et al.*, Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation. *Mol Cell Proteomics* 2011, 10, M110 003699.
- [442] Mayr, M., Rabilloud, T., Multidimensional separation prior to mass spectrometry: getting closer to the bottom of the iceberg. *Proteomics* 2013, 13, 2942-2943.
- [443] Olsen, J. V., Macek, B., Lange, O., Makarov, A., *et al.*, Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* 2007, 4, 709-712.
- [444] Olsen, J. V., Schwartz, J. C., Griep-Raming, J., Nielsen, M. L., *et al.*, A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. *Mol. Cell Proteomics* 2009, 8, 2759-2769.
- [445] Eliuk, S., Makarov, A., Evolution of Orbitrap Mass Spectrometry Instrumentation. *Annu Rev Anal Chem (Palo Alto Calif)* 2015, 8, 61-80.
- [446] Andrews, G. L., Dean, R. A., Hawkrige, A. M., Muddiman, D. C., Improving proteome coverage on a LTQ-Orbitrap using design of experiments. *J. Am. Soc. Mass. Spectrom.* 2011, 22, 773-783.
- [447] Picotti, P., Clement-Ziza, M., Lam, H., Campbell, D. S., *et al.*, A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* 2013, 494, 266-270.
- [448] Kalli, A., Smith, G. T., Sweredoski, M. J., Hess, S., Evaluation and optimization of mass spectrometric settings during data-dependent acquisition mode: focus on LTQ-Orbitrap mass analyzers. *J. Proteome Res.* 2013, 12, 3071-3086.
- [449] Rabelo, K., Trugilho, M. R., Costa, S. M., Ferreira, A. T., *et al.*, Dataset of proteins mapped on HepG2 cells and those differentially abundant after expression of the dengue non-structural 1 protein. *Data Brief* 2017, 10, 248-263.
- [450] Rabelo, K., Trugilho, M. R., Costa, S. M., Pereira, B. A., *et al.*, The effect of the dengue non-structural 1 protein expression over the HepG2 cell proteins in a proteomic approach. *Journal of proteomics* 2017, 152, 339-354.
- [451] Nesvizhskii, A. I., Aebersold, R., Interpretation of shotgun proteomic data: the protein inference problem. *Mol. Cell Proteomics* 2005, 4, 1419-1440.
- [452] Zhang, B., VerBerkmoes, N. C., Langston, M. A., Uberbacher, E., *et al.*, Detecting differential and correlated protein expression in label-free shotgun proteomics. *J Proteome Res* 2006, 5, 2909-2918.
- [453] Pascovici, D., Handler, D., Wu, J. X., Haynes, P. A., Multiple testing corrections in quantitative proteomics: A useful but blunt tool. *Proteomics* 2016, n/a-n/a.
- [454] Jaksik, R., Iwanaszko, M., Rzeszowska-Wolny, J., Kimmel, M., Microarray experiments and factors which affect their reliability. *Biol Direct* 2015, 10, 46.

- [455] Wu, J. X., Song, X. M., Pascovici, D., Zaw, T., *et al.*, SWATH Mass spectrometry performance using extended peptide MS/MS assay libraries. *Mol Cell Proteomics* 2016, 15, 2501-2514.
- [456] Korla, K., Vadlakonda, L., Mitra, C. K., Kinetic simulation of malate-aspartate and citrate-pyruvate shuttles in association with Krebs cycle. *J Biomol Struct Dyn* 2015, 33, 2390-2403.
- [457] Walther, T. C., Farese, R. V., Jr., Lipid droplets and cellular lipid metabolism. *Annu Rev Biochem* 2012, 81, 687-714.
- [458] Noguchi, Y., Young, J. D., Aleman, J. O., Hansen, M. E., *et al.*, Effect of anaplerotic fluxes and amino acid availability on hepatic lipoapoptosis. *J Biol Chem* 2009, 284, 33425-33436.
- [459] Volmer, R., van der Ploeg, K., Ron, D., Membrane lipid saturation activates endoplasmic reticulum unfolded protein response transducers through their transmembrane domains. *Proc Natl Acad Sci U S A* 2013, 110, 4628-4633.
- [460] Fu, S., Yang, L., Li, P., Hofmann, O., *et al.*, Aberrant lipid metabolism disrupts calcium homeostasis causing liver endoplasmic reticulum stress in obesity. *Nature* 2011, 473, 528-531.
- [461] Suzuki, T., Lu, J., Zahed, M., Kita, K., Suzuki, N., Reduction of GRP78 expression with siRNA activates unfolded protein response leading to apoptosis in HeLa cells. *Arch Biochem Biophys* 2007, 468, 1-14.
- [462] Egnatchik, R. A., Leamy, A. K., Jacobson, D. A., Shiota, M., Young, J. D., ER calcium release promotes mitochondrial dysfunction and hepatic cell lipotoxicity in response to palmitate overload. *Mol Metab* 2014, 3, 544-553.
- [463] Kennedy, D., Samali, A., Jager, R., Methods for studying ER stress and UPR markers in human cells. *Methods Mol Biol* 2015, 1292, 3-18.
- [464] Osowski, C. M., Urano, F., in: Conn, P. M. (Ed.), *Methods in Enzymology*, Academic Press 2011, pp. 71-92.
- [465] Deniaud, A., Sharaf el dein, O., Maillier, E., Poncet, D., *et al.*, Endoplasmic reticulum stress induces calcium-dependent permeability transition, mitochondrial outer membrane permeabilization and apoptosis. *Oncogene* 2008, 27, 285-299.
- [466] Rolo, A. P., Teodoro, J. S., Palmeira, C. M., Role of oxidative stress in the pathogenesis of nonalcoholic steatohepatitis. *Free Radic Biol Med* 2012, 52, 59-69.
- [467] Huang, Q., Yang, L., Luo, J., Guo, L., *et al.*, SWATH enables precise label-free quantification on proteome scale. *Proteomics* 2015, 15, 1215-1223.

APPENDIX

Table 20: File format used by software for proteomics analysis

| Programme | File format | Purpose |
|------------------------------|-------------|---|
| Xcalibur for LTQ-Orbitrap XL | .raw | Generate LC-MS/MS data |
| Proteome Discoverer | .msf | Protein identification |
| Scaffold | .sf3 | Label-free quantification using spectral counting |
| MS Excels | .xlsx | Presenting proteins lists and functional analysis in spreadsheets |

Table 21: Organization and description of supplementary content in the electronic form. Due to the large file size and the number of files, the LC-MS/MS data and the protein identification data has been made available through online repositories. List of identified proteins are best presented on spreadsheets for the want of multiple columns and rows. This table outlines the contents of the folders to help navigate to the required file. Information for each dataset is contained within the name-bearing folder and all folders are organized along the same lines unless indicated otherwise.

| Folder name | Description of contents |
|--|--|
| Scaffold | Label-free quantification performed in Scaffold |
| Consensus list | An .msf file was generated using all LC-MS/MS files for a given biological replicate as input and the list of proteins identified is presented in the .xlsx files. |
| LC1_Pooled.xlsx | Proteins identified and quantified in Scaffold for dataset LC1. |
| LC2_Unpooled.xlsx | Proteins identified and quantified in Scaffold for dataset LC2. |
| LC3_Unpooled.xlsx | Proteins identified and quantified in Scaffold for dataset LC3. |
| LC4_Pooled.xlsx | Proteins identified and quantified in Scaffold for dataset LC4; List of protein enriched in GO and KEGG pathway. |
| 'Analysis of technical variance in dataset LC4 | .sf3 and .xlsx files to compare the technical replicates of each pooled sample |

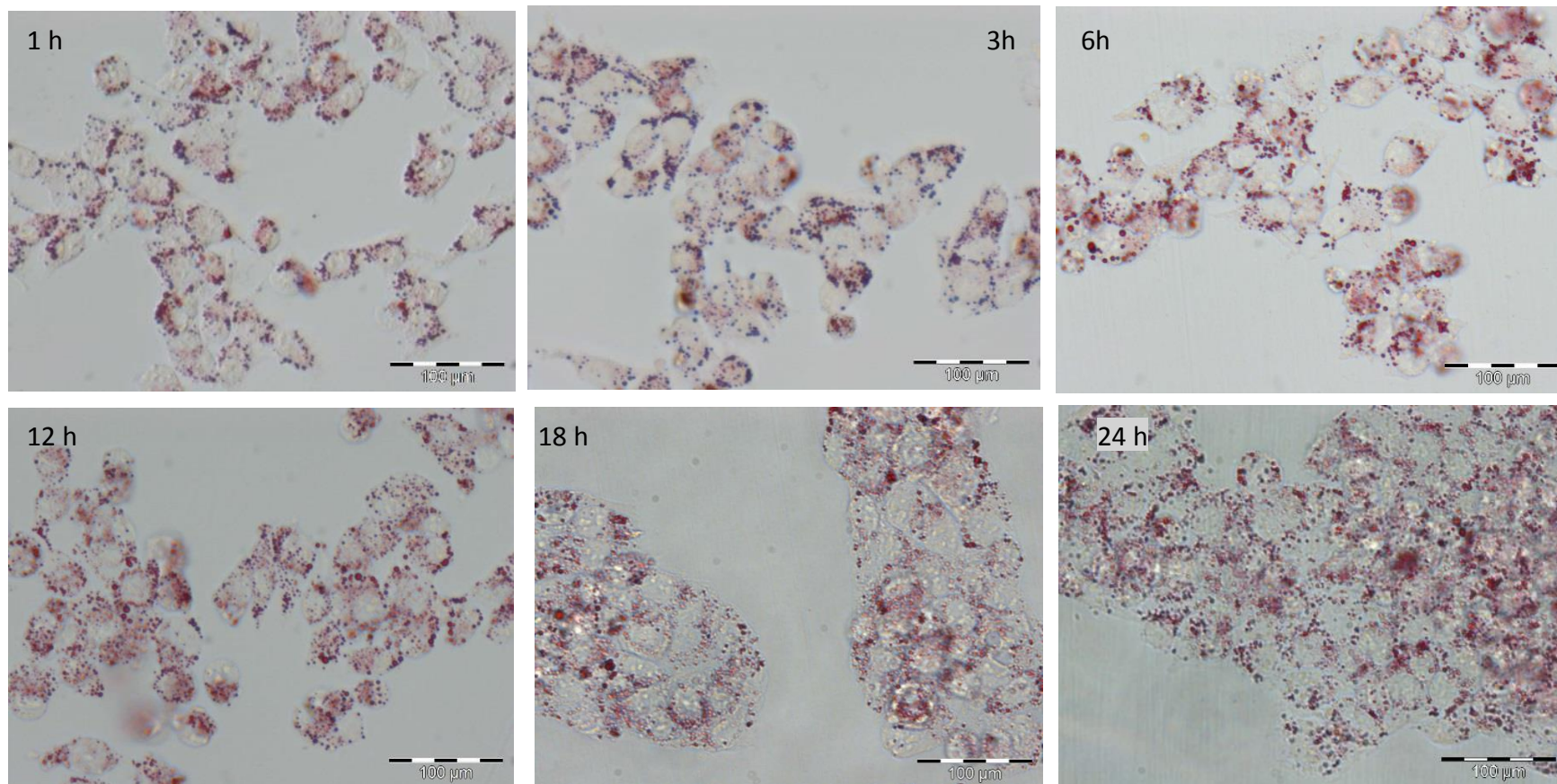


Figure 66: Lipid accumulation over 24 h in control cells.

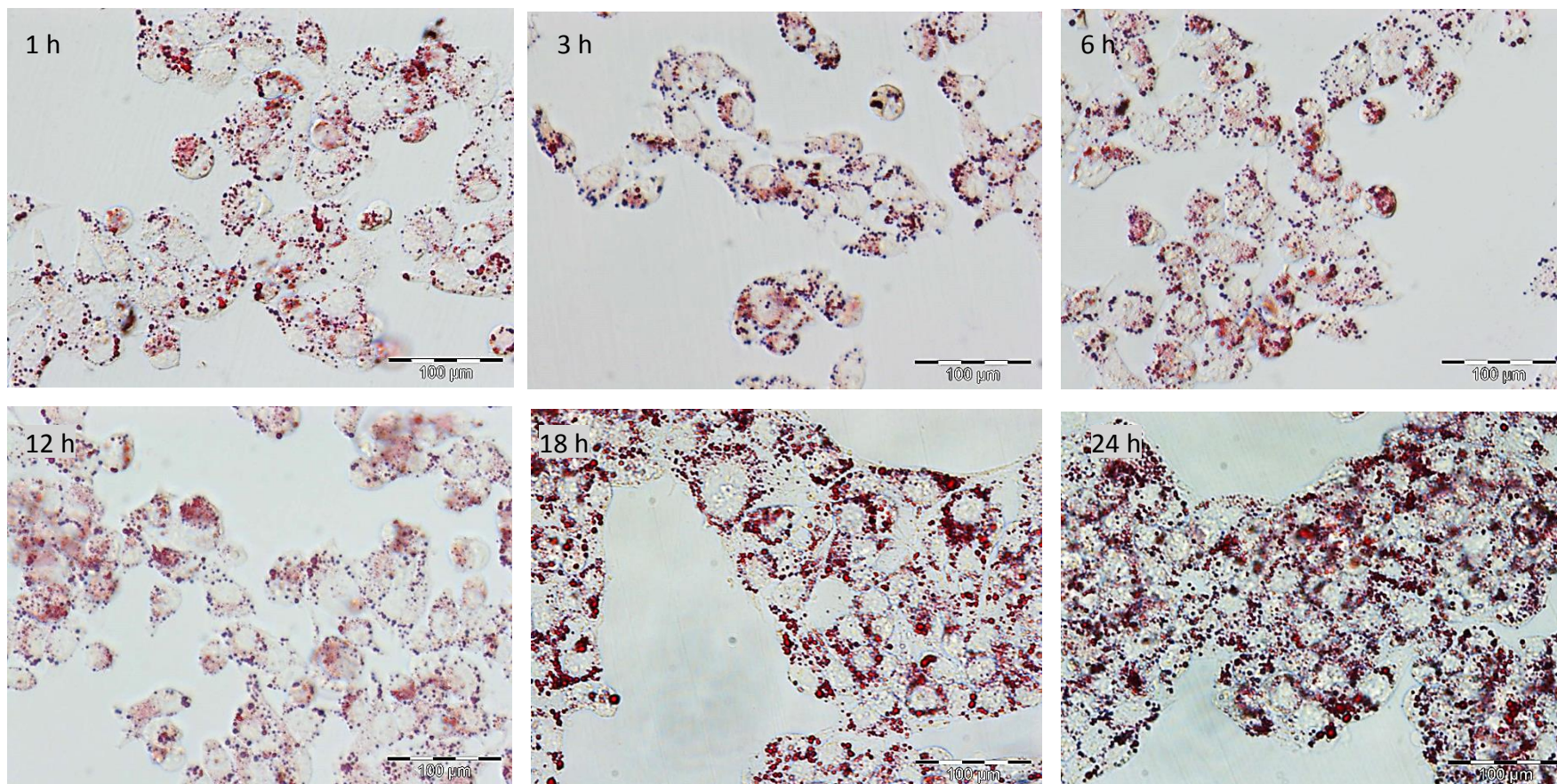


Figure 67: Lipid accumulation over 24 h in HepG2 cells treated with 0.25 mM FA mixture.

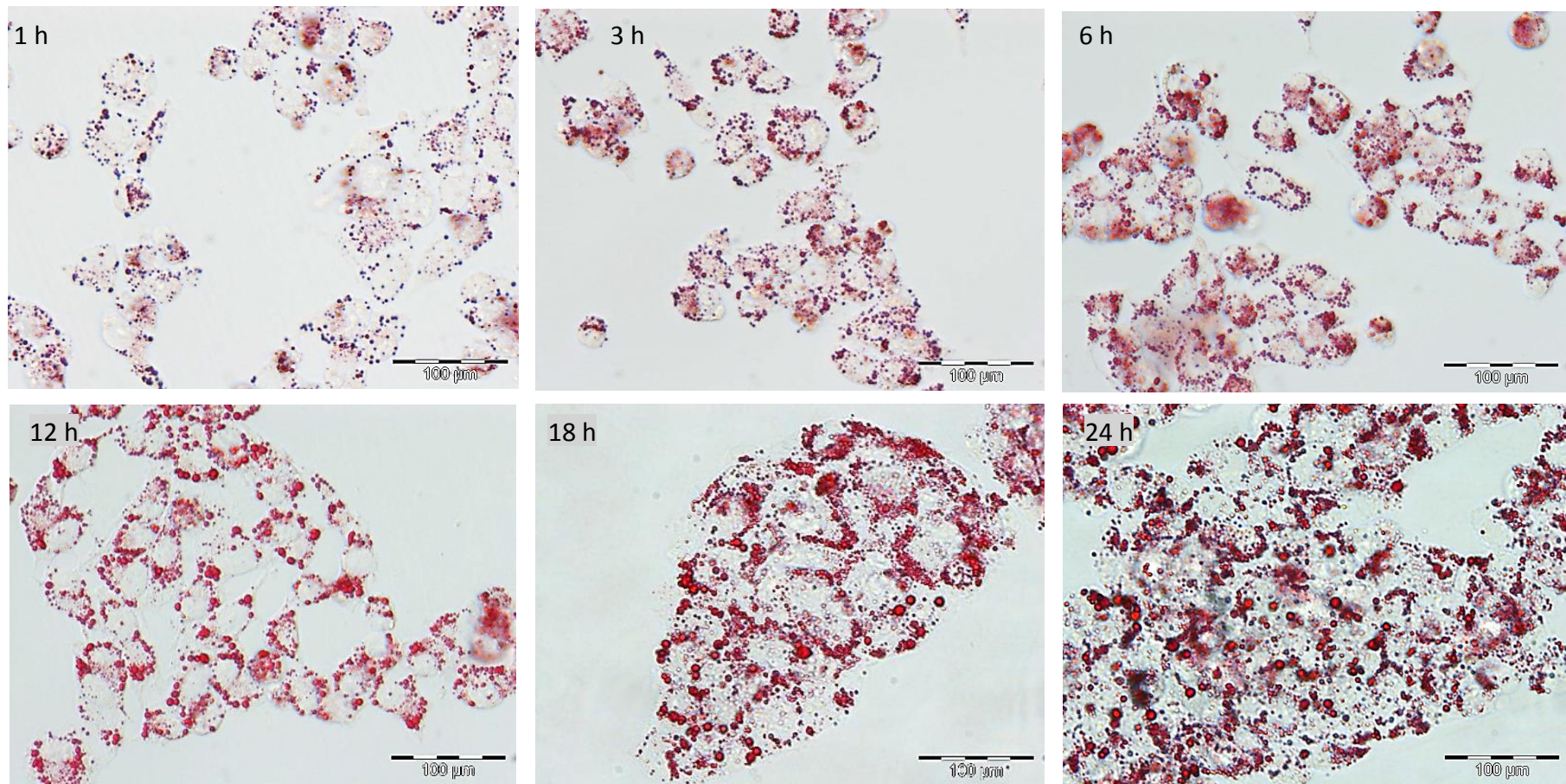


Figure 68: Lipid accumulation over 24 h in HepG2 cells treated with 0.5 mM FA mixture.

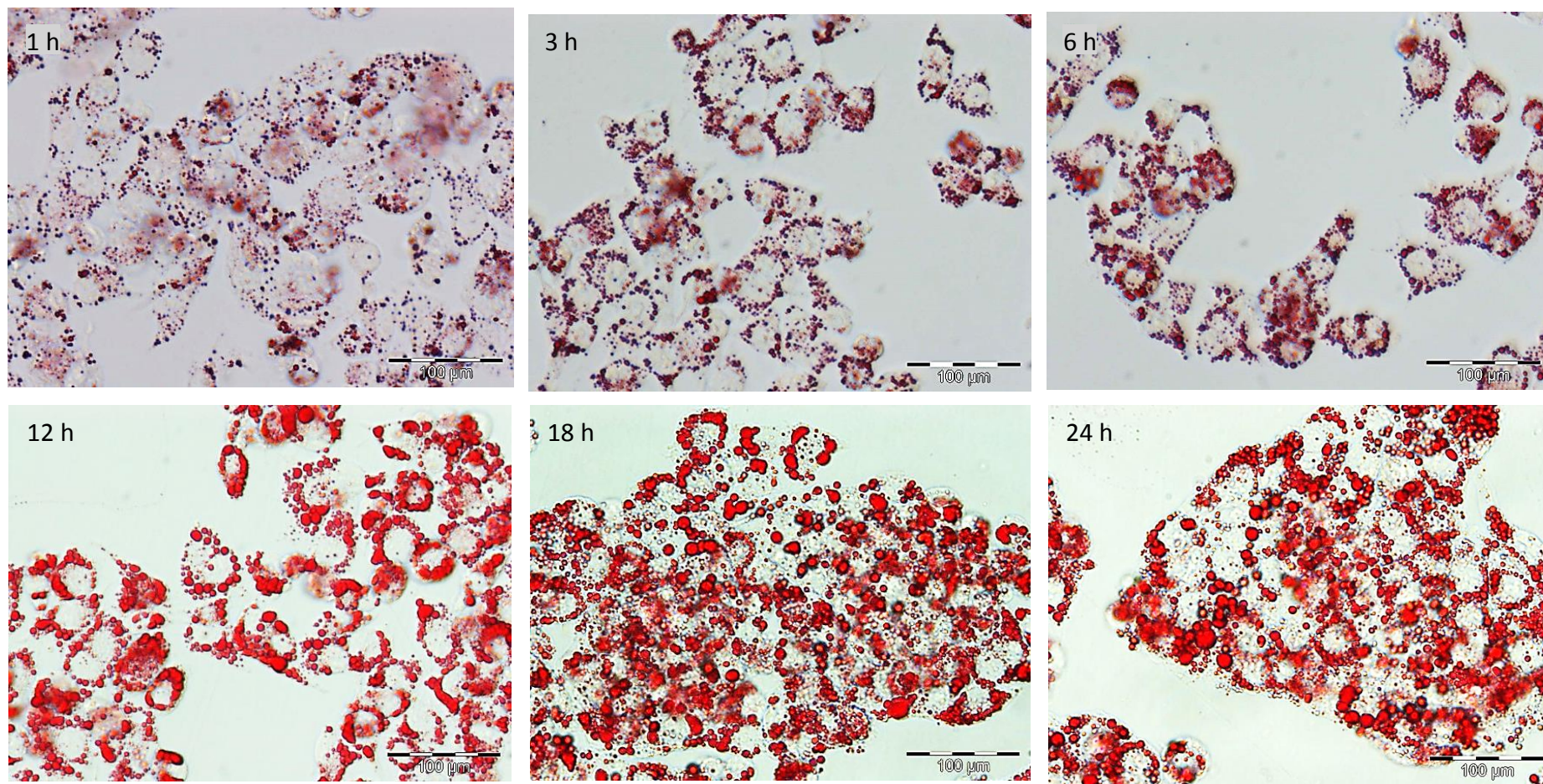


Figure 69: Lipid accumulation in HepG2 cells treated with 2 mM FA mixture.

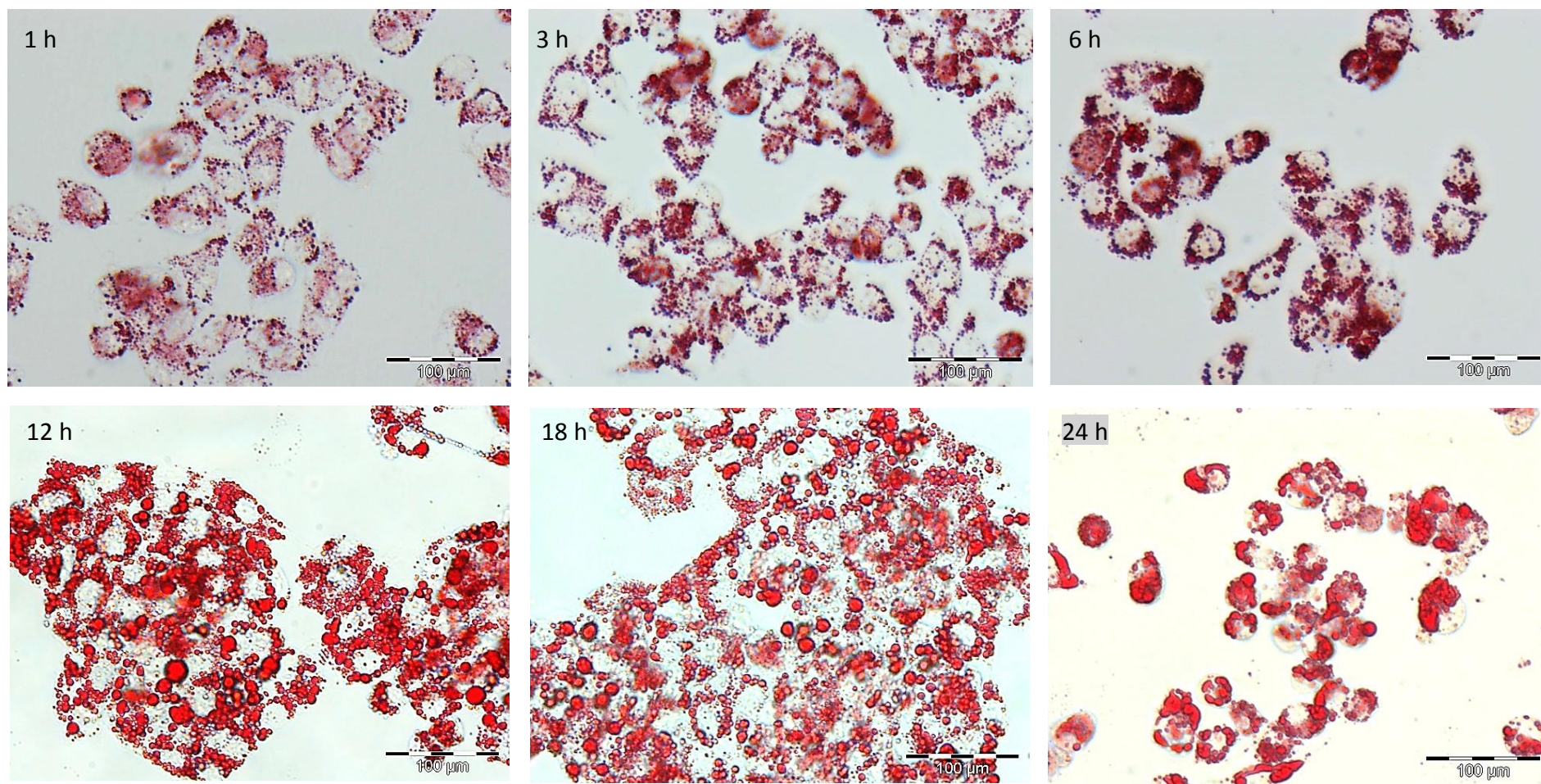


Figure 70: Lipid accumulation in HepG2 cells treated with 3 mM FA mixture. A large number of lysed cells are observed at 24 h.

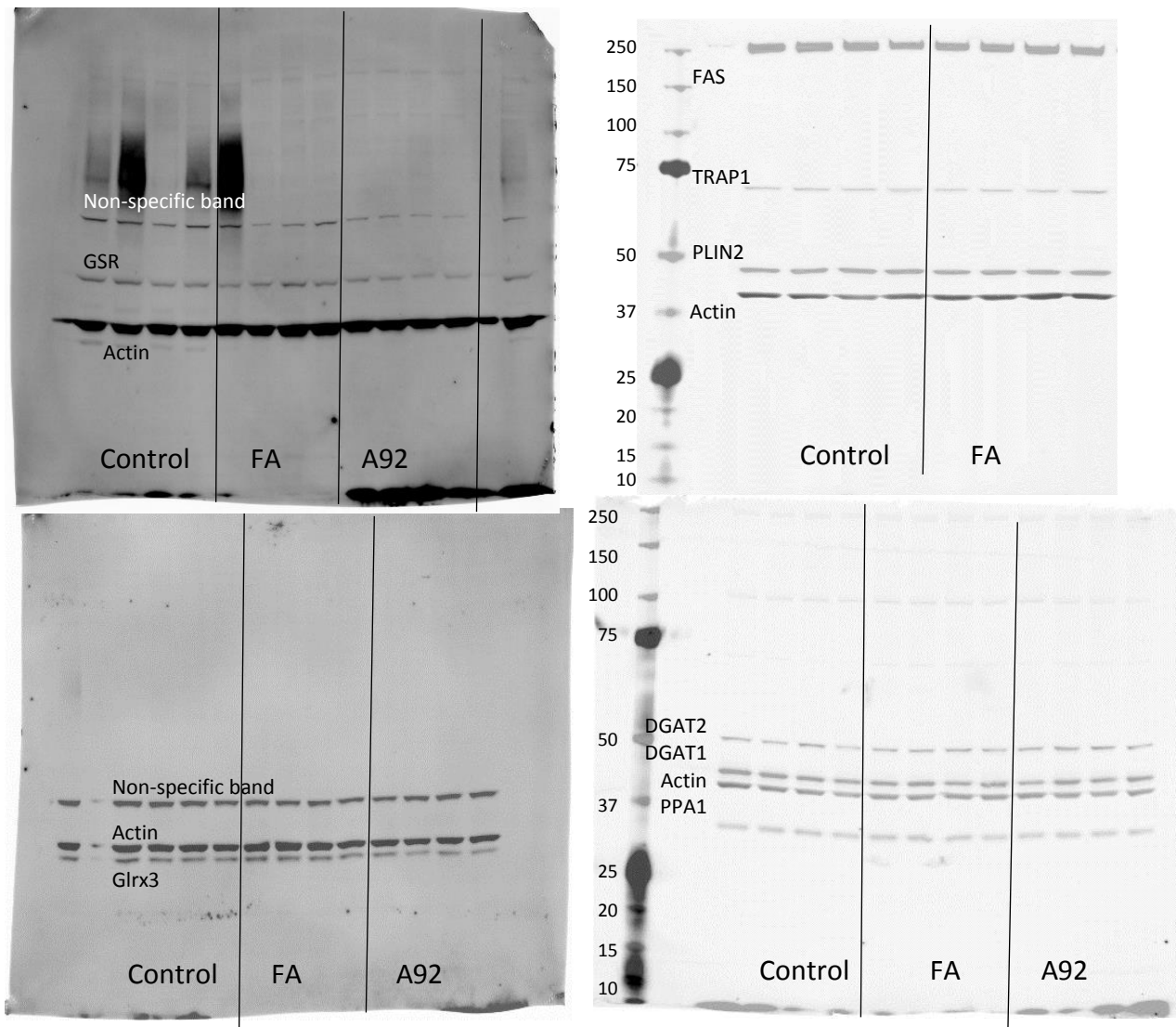


Figure 71: Uncropped images of Western blots for target proteins of dataset LC1.

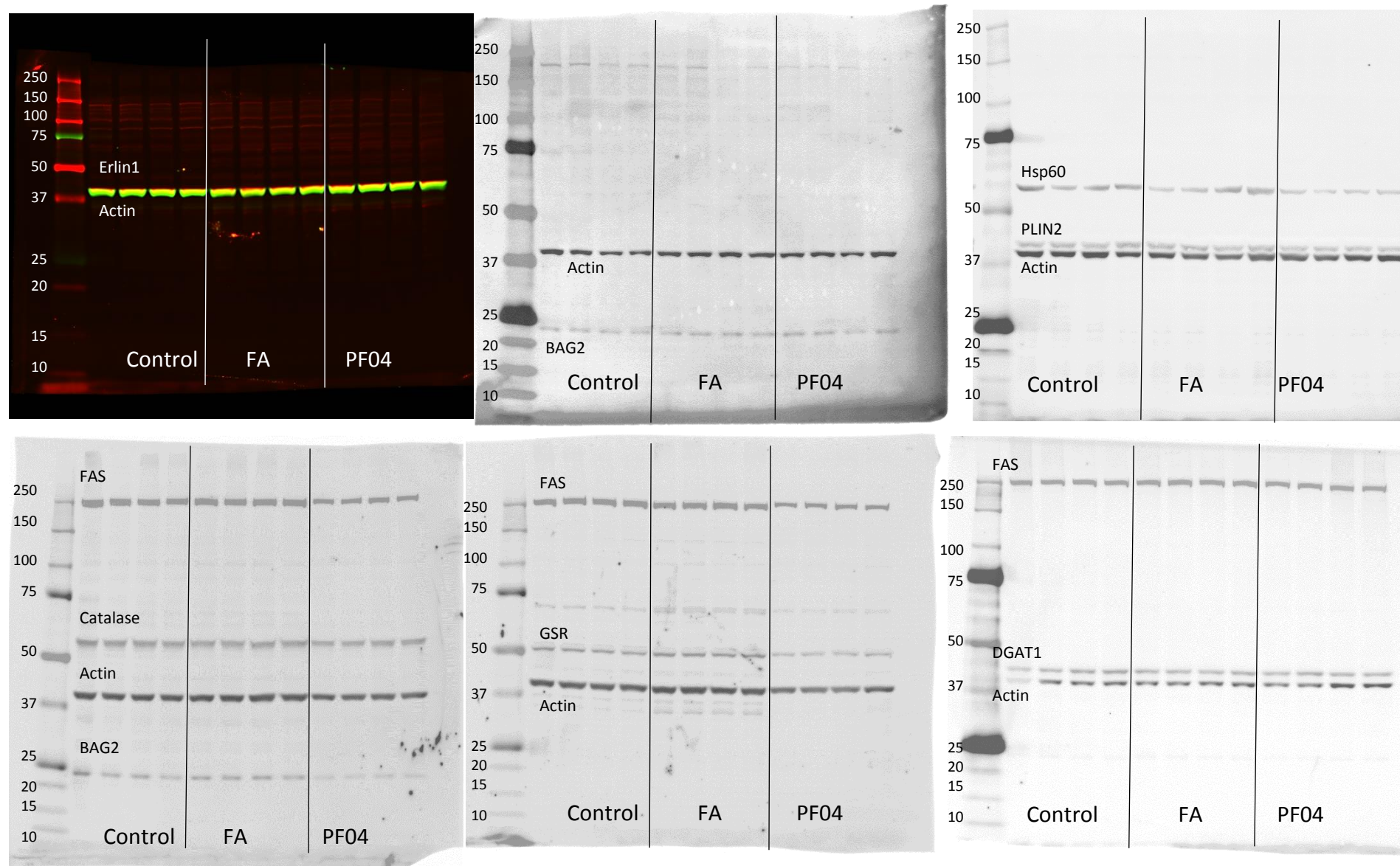


Figure 72: Uncropped images for Western blots of target proteins from dataset LC3 and LC4.