Glycobiology of Adipose-Derived Stem Cell Differentiation

Katherine Wongtrakul-Kish

Bachelor of Medical Science (Hons) (Macquarie University)

This thesis is presented for the degree of

Doctor of Philosophy

Department of Chemistry and Biomolecular Sciences

Macquarie University, Sydney, Australia

March 2016

Table of Contents

Abstract **8**

Declaration **11**

Acknowledgements **12**

Abbreviations **14**

Chapter 1: Introduction **17**

1.1 Mesenchymal Stem Cells 18
1.1.1 What are stem cells? 18
1.1.2 MSCs in regenerative medicine 18
1.1.3 Adipose tissue as a source of MSCs 19
1.1.4 Adipose-derived stem cells in vitro 20
1.2 Adipogenic differentiation of ADSCs 21
1.2.1 Adipogenesis in vivo 21
1.2.2 Adipogenesis in vitro 22
1.2.2.1 Methods to evaluate in vitro adipogenesis 22
1.3 Stem cell differentiation and glycosylation 23
1.3.1 Membrane protein glycosylation 23
1.3.1.1 N- and O-linked glycan biosynthesis 25
1.3.2 Glycoproteins in stem cell differentiation 26
1.3.3 Changes in glycosylation in stem cell differentiation 26
1.4 Analysis of plasma membrane glycoconjugates 30
1.4.1 Strategies to study protein glycosylation 30
1.4.2 Release of protein N- and O-linked glycans 31
1.4.3 Analysis methods for the characterisation of released N- and O-linked glycans 32
1.4.3.1 Separation of released glycans 32
1.4.3.2 Structural analysis of released glycans 34
1.4.3.3 Characterising glycans by MS analysis 35
1.4.3.4 Determination of glycan structure from MS data 37
1.5 The plasma membrane proteome 38
1.5.1 Proteomics of human adipogenic stem cell differentiation 39
1.5.2 Analytical considerations when investigating membrane proteins 39
1.5.2.1 Isolating plasma membranes proteins 40
1.5.2.2 Membrane protein solubilisation and enrichment 40
1.5.3 Mass spectrometry to study membrane proteins 41
1.5.3.1 Quantitative proteomics 42
1.5.3.2 Data analysis in label-free proteomics 43
1.5.3.3 Forming biological conclusions from proteomics experiments 43
1.6 Thesis aims 44

Chapter 2: Evaluating adipose-derived stem cell differentiation *in vitro* **47**

2.1 Introduction 48
2.2 Materials and Methods 50
2.2.1 Materials 50

2.2.2 Lipo-aspirate processing, and adipocyte and ADSC isolation 50
2.2.3 Sourcing media component FBS for the culture of ADSCs and dADSCs 52
2.2.4 Staining and imaging adipocytes and adipose tissue with Oil Red O and Hoechst 55
2.2.5 Adipogenic differentiation of ADSCs in 35 mm^2 wells for time course imaging 55
2.2.6 Adipogenic differentiation in 75 cm^2 flasks 56
2.2.7 Staining dADSCs with Oil Red O and Hoechst 57
2.2.8 Microscopy 57
2.2.9 Collecting ADSCs and dADSCs for total RNA extraction 58
2.2.10 Total RNA extraction from ADSCs, dADSCs and adipocytes 58
2.2.11 qRT-PCR to measure mRNA expression of marker genes 59
2.3. Results 60
2.3.1 Imaging adipose tissue using Differential Interference Contrast and fluorescence microscopy 61
2.3.2 Imaging adipocytes from collagenase-digested adipose tissue 62
2.3.3 In vitro adipogenesis of ADSCs 63
2.3.4 Phase contrast microscopy to visualise in vitro large-scale adipogenesis 66
2.3.5 Comparing in vivo and in vitro products of adipogenic differentiation 70
2.3.6 mRNA expression of cell-marker genes by qRT-PCR 71
2.4 Discussion 73
2.5 Concluding remarks 77
Chapter 3: Characterisation and comparison of *N*- and *O*-linked glycosylation during *in vivo* and *in vitro* adipogenic differentiation 79
3.1 Introduction 80
3.2 Materials and Methods 81
3.2.1 Materials 81
3.2.2 Lipo-aspirate processing, and adipocyte and ADSC isolation 81
3.2.3 Optimisation of cell membrane preparation 82
3.2.4 Comparison of membrane proteins and N-glycosylation with and without the use of Triton X-114 membrane protein enrichment 83
3.2.5 N- and O-linked glycan release from cell membrane glycoproteins 83
3.2.6 Exoglycosidase digestion of reduced N-linked glycans 83
3.2.7 PGC-LC ESI-MS/MS and Data Analysis 84
3.2.8 Lectin dot blotting 84
3.2.9 qRT-PCR to measure MGAT3 and Fut8 mRNA expression levels 85
3.3 Results 86
3.3.1 Evaluating the use of Triton X-114 for membrane protein enrichment 86
3.3.2 N- and O-glycomic profiling of ADSCs over cell culture passages 87
3.3.3 Optimisation of glycan release method 89
3.3.4 PGC-LC ESI-MS/MS separation and identification of isomeric structures 90
3.3.4.1 Identification of isomers with bisecting GlcNAc and lactosamine repeats 90
3.3.4.2 Identification of sialic acid linkage isomers 92
3.3.4.3 Identification of structures containing a terminal Gal-Gal epitope 93
3.3.5 Comparison of ADSCs and adipocytes from the same individual: a model of in vivo differentiation 95

3.3.6 Comparison of ADSCs and dADSCs from the same individual: a model of in vitro differentiation 97
3.3.7 N-glycome of ADSCs, dADSCs and adipocytes: in vitro vs. in vivo adipogenesis 98
3.3.7.1 High Mannose Structures 99
3.3.7.2 Complex Structures 100
3.3.7.3 Hybrid and Pauci Mannose Structures 101
3.3.8 Bisecting GlcNAc structures are characteristic of adipogenesis 101
3.3.9 Cell-specific markers 105
3.3.10 O-glycome of ADSCs, dADSCs and adipocytes 106
3.3.11 Validation of bisecting GlcNAc and core fucosylation expression in differentiation of ADSCs 108
3.3.11.1 Lectin blotting to confirm bisecting GlcNAc expression 108
3.3.11.2 mRNA expression levels of MGAT3 and Fut8 during adipogenesis 109
3.4 Discussion 110
3.5 Concluding remarks 116
Chapter 4: Characterisation and comparison of the plasma membrane proteome during *in vivo* and *in vitro* adipogenic differentiation 117
4.1 Introduction 118
4.2 Materials and Methods 119
4.2.1 Materials 119
4.2.2 Sample collection and membrane protein enrichment 119
4.2.3 SDS-PAGE fractionation of membrane proteins and in-gel digestion 119
4.2.4 Nanoflow RPLC-MS/MS 120
4.2.5 Database search for peptide and protein identification 121
4.2.6 Data processing and quantitation 121
4.2.7 Statistical methods 122
4.2.8 Gene Ontology annotation 122
4.3 Results 123
4.3.1 Data quality 123
4.3.2 Membrane proteomics results 124
4.3.3 Distribution of proteins 125
4.3.3.1 Proteins identified reproducibly in adipocytes alone 125
4.3.3.2 Proteins identified reproducibly in dADSCs alone 128
4.3.3.3 Proteins expressed in cell membrane of both in vivo and in vitro models of adipogenesis 130
4.3.3.4 Proteins identified reproducibly in ADSCs alone 132
4.3.4 Proteins present in all three cell types 134
4.3.5 Pair-wise comparisons of ADSCs with their differentiated adipogenic product cells 137
4.3.5.1 Proteins down-regulated in adipogenesis: in vitro vs. in vivo 138
4.3.5.2 Proteins up-regulated in adipogenesis: in vitro vs. in vivo 142
4.3.6 Expression of ECM related proteins 144
4.3.7 Comparison of mRNA transcript and protein abundance in the three cell types 145
4.4 Discussion 148
4.5 Concluding Remarks 153

Chapter 5: Adipose-derived stem cell differentiation: targeting the cell membrane protein glycosylation 155

5.1 Introduction 156
5.2 Materials and Methods 157
5.2.1 Materials 157
5.2.2 Adipocyte and ADSC isolation from lipo-aspirate 157
5.2.3 Trypsin digestion of SDS-PAGE separated membrane proteins with and without prior in-gel PNGase F de-N-glycosylation 158
5.2.4 PGC-LC ESI-MS/MS analysis of N-linked glycans and RP-nanoLC ESI-MS/MS analysis of de-glycosylated peptides 158
5.2.5 Peptide analysis 159
5.2.6 Protein identification 159
5.2.7 Anti-CD36 Western blotting 160
5.2.8 CD36 mRNA transcript expression by qRT-PCR 160
5.3 Results 161
5.3.1 SDS-PAGE separation of adipocyte, ADSC and dADSC membrane proteins 161
5.3.1.1 Analysis of N-linked glycans released from gel fractions following SDS-PAGE of membrane proteins from native adipocytes, ADSCs and dADSCs 161
5.3.1.2 Identification of abundant glycoproteins in native adipocytes, ADSCs and dADSCs .. 164
5.3.2 Effect of de-glycosylation on determination of CD36 abundance in adipocytes 171
5.3.3 Confirming CD36 differential expression using complementary Western blotting and qRT-PCR 172
5.4 Discussion 174
5.5 Concluding Remarks 176

Chapter 6: General Discussion 177

6.1 Thesis overview 178
6.1.1 Bisecting GlcNAc structures are characteristic of adipocyte N-linked glycosylation 179
6.1.2 The dADSC membrane proteome shares proteins and processes with native adipocytes180
6.1.3 Extracellular matrix-related proteins exhibit differential expression during adipogenesis180
6.1.4 Epigenetic processes potentially underpin in vitro adipogenesis 181
6.1.5 Bisecting GlcNAc N-linked glycans are carried by most adipocyte proteins 181
6.2 Analytical considerations 181
6.2.1 Adipose samples and processing 181
6.2.2 Membrane protein isolation 182
6.2.3 Shotgun proteomics underestimates the abundance of highly glycosylated proteins 183
6.2.4 1D SDS-PAGE separation affects protein resolution and targeted proteomics approaches 183
6.3 Future directions and concluding remarks 183

References 187

Chapter 3 Supplementary Data 207

Table S3.1 208
Table S3.2 220
Table S3.3 221

Chapter 4 Supplementary Data 223

Table S4.1 224
Table S4.2 226
Table S4.3 231
Appendix A 235
1. Ethics Approval 236
2. Contents of Accompanying DVD 238
- Additional Chapter 4 Supplementary Data 238
- Additional Chapter 5 Supplementary Data 238

Abstract

Human adipose tissue is a major source of mesenchymal stem cells, with numbers exceeding those in bone marrow and blood. These stem cells are referred to as adipose-derived stem cells (ADSCs) and have the ability to differentiate into different cell types. *In vitro*, ADSCs can be made to differentiate into the mesenchymal lineages of adipocytes, osteoblasts and chondrocytes, while *in vivo,* ADSCs become pre-adipocytes before further differentiating into mature adipocytes. Protein glycosylation has been shown to change in stem cell differentiation, and while ADSCs have been acknowledged for their therapeutic potential, the investigation of protein glycosylation changes during ADSC adipogenic differentiation has not been reported.

First, adipogenic differentiation of ADSCs *in vitro* was optimised and evaluated by measuring the expression levels of established gene-markers and staining of the accumulated lipid in differentiated cells using Oil Red O. These results confirmed the cells had entered the adipogenic lineage and shared some characteristics with, but did not entirely mirror native, *in vivo* produced adipocytes.

ADSCs, *in vitro* differentiated ADSCs and *in vivo* native mature adipocytes harvested from the same patients were used to investigate the membrane protein *N*- and *O*-linked glycosylation of adipogenesis. Using porous graphitised carbon LC coupled with negative ion ESI-MS/MS, a total of 138 glycan structures carried by the membrane proteins were characterised across the three cell types. Bisecting GlcNAc type *N*-linked structures were detected at a high level (32.1 % of total glycans) in mature adipocytes with some appearing in the differentiating ADSCs (1.9 %), while overall, ADSCs and their *in vitro* differentiated progeny showed similar glycosylation expression profiles. This finding was further correlated by increased mRNA expression of the MGAT3 gene responsible for the enzyme synthesis of this glycan structure type in both the *in vivo* and *in vitro* differentiated ADSCs.

The same cell types (ADSCs, differentiated ADSCs and mature adipocytes) were characterised for their membrane proteome using label-free quantitative shotgun proteomics analysis. Many protein differences were identified between the three cell types, with, importantly, adipocyte-specific proteins found to be up-regulated in both mature adipocytes and differentiating ADSCs. Mitochondrial and lipid-related processes were also up-regulated in both adipogenic cell products, whereas epigenetic-related proteins and processes were found to be up-regulated in *in vitro* differentiated ADSCs alone.

Analysis of the *N*-linked glycans from SDS-PAGE separated membrane proteins of the three cell types revealed that expression of bisecting GlcNAc structures was present on the majority of adipocyte glycoproteins. A more targeted methodology of carrying out proteomic analysis of de-*N*-glycosylated

peptides of the gel-separated proteins unearthed new glycoproteins not detected previously in the shotgun proteomic analysis without de-glycosylation. This approach identified the adipogenic marker, CD36, as the dominant glycoprotein in the adipocyte membrane proteome that was also up-regulated at the mRNA transcript level in both the *in vitro* differentiated ADSCs (7.1-fold increase) and mature adipocytes (102.9-fold increase). This work highlighted the importance of de-*N*-glycosylation of proteins in proteomics experiments for increased identification of glycoproteins.

The systems glycobiology approach by the integration of glycomics, proteomics and transcriptomics analyses in this thesis extended the investigation of membrane protein glycosylation changes in adipose-derived stem cell differentiation. The work provides a framework for future glycoproteomics-based investigations into the differentiation of stem cells into adipocytes, and will allow their related pathologies and potential therapeutic applications to be discovered.

Declaration

I certify that this thesis, titled "Glycobiology of Adipose-Derived Stem Cell Differentiation" is the result of my own work and has not been submitted for higher degree to any other university or institution. This research was approved by the Macquarie University human research ethics committee (Ref # 5201100385).

Katherine Wongtrakul-Kish

Student ID: 40963160

March 2016

Acknowledgements

There are many people who I have shared the past five years with that I would like to acknowledge and thank for their support and encouragement. Unfortunately, there are too many people to list by name here, however the following people deserve specific mention:

First and foremost, my sincerest thanks to my supervisor Professor Nicki Packer, who introduced me to the world of glycobiology. I would like to thank her for her guidance and support, which has been instrumental in my training as a researcher. Also for the opportunity to always have candid discussion about work and life. Her hardworking nature and resilience inspired me to keep at it.

To Dr Daniel Kolarich, who first sparked my interest in mass spectrometric-driven glycomics in my honours year, and for allowing me to spend time with his glyco-group at the Max Planck Institute for Colloids and Interfaces in Berlin during my PhD. Also, thanks to Dr Terry Nguyen-Khuong for his enthusiasm and encouraging self-confidence in my abilities as a mass spectrometrist.

I would particularly like to thank Dr Robyn Peterson, a true inspiration in every sense of the word, for her kindness and generosity, and endless help and encouragement in bringing together this thesis. Robyn, you are one of the best people I know. Also, to Dr Matthew Campbell for his help with this thesis and beyond, and Dr Liisa Kautto for always ensuring a good laboratory working environment. To Professor Paul Haynes, for his generous time and help with my proteomics work, and to Dr Dana Pascovici for her endless patience and help with statistical analyses. Thanks and gratitude also to Debra Birch and Dr Anahit Panesyan for help with microscopy and transcriptomic work.

To the current and past members of the Glyco group at Macquarie University, it has been a pleasure having you all as my peers to share this experience with and I appreciated all our post-group meeting drinks at the uni bar, coffee breaks, as well as our many food excursions. Thanks to all the friends I have made in the group, but most specifically to Dr Albert Lee, Jodie Abrahams, Dr Arun Everest-Dass, Dr Maja Christiansen, Kathrin Stavenhagen, Dr Morten Thaysen-Andersen, Dr Chi-Hung Lin, Dr Jasmine Grinyer, Ling Yen Lee, Dr Clara Cheah, Jenny Chik, Zeynep Sumer-Bayraktar, Wei-Wern Chong, Vignesh Venkatakrishnan and Merrina Anugraham for their friendship and support over the years.

The other main (non-glyco) aspect of this thesis was my time with Associate Professor Ben Herbert's group and their affiliation with Regeneus, without which I would not have had any samples. Ben, thank you for allowing me to be a part of your group and for introducing me to the world of stem cell regenerative medicine. To the members of the Herbert group, particularly Dr Sinead Blaber, whose friendship, patience and support were paramount in a daunting new area of work in which I had no

experience. Thanks also to Hannes Hinneburg for teaching me stem cell differentiation, to Dr Michael Sivell, Dr Michael Medynskyj and Cameron Hill, and to Peter Succar and Liz Karsten particularly in my last year of lab work. Above all, thank you to you all for the many laughs we shared I was always guaranteed a comic escape in all my trips down to your lab!

There are many members of the E8B level 3 labs that I would like to acknowledge, particularly the members of the biomarker lab and my office mates. Thanks to Dr Naveid Ali, whose positive attitude always inspired me, Dr Sarah Randall for always encouraging me, and Dr Karlie Neilson for grounding me and always putting things in perspective. To other BMD lab members: Crystal Semaan, Iniga George, Karthik Kamath, Dr Christoph Krisp, Dr Sibasish Dolai, Juan Martinez-Aguilar, Manveen Sethi and Zhiping Han it wouldn't have been the same without you all.

Finally, to my family both in Australia and Thailand for their never ending love, but most particularly to my parents for giving me courage and conviction in all my pursuits. Everything they have given me has made this possible.

***Biomarker Discovery Lab**, September 2014*

Abbreviations

ABS: α2-3,6 neuraminidase from *Arthrobacter ureafaciens*
ADSC/s: Adipose-derived stem cell/s
APTS: 8-aminopyrene-1,3,6-trisulfonic acid
AUC: area under curve
bp: base pairs
BPC: Base Peak Chromatogram
BSA: Bovine serum albumin
BTG: β1-3,4 galactosidase from bovine testis
CD Marker: Cluster of Differentiation Marker
CE: Capillary Electrophoresis
Cell R: Olympus Cell R Live Imaging System with IX 71 Inverted microscope
CID: collision induced dissociation
dADSC/s: differentiated adipose-derived stem cell/s
DIC: Differential Interference Contrast
DMEM-F12: Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12
DMEM-HG: Dulbecco's Modified Eagle Medium-High Glucose
ECD: electron capture dissociation
EIC: Extracted ion chromatogram
ESI-MS: Electrospray Ionisation Mass Spectrometry
ETD: electron transfer dissociation
FABP4: Fatty acid binding protein 4
FBS: fetal bovine serum
FluoView: Olympus FluoView FV1000 with IX 81 Inverted Confocal Microscope
Fuc: Fucose
Fut8: Fucosyltransferase 8 (α(1,6) fucosyltransferase)
Gal: Galactose
GalNAc: *N*-acetylgalactosamine
GC: Gas chromatography
Glc: Glucose
GlcNAc: *N*-acetylglucosamine
GUSB: β-glucuronidase
HCD: high energy C-trap dissociation
Hex: Hexose
HexNAc: *N*-acetylhexosamine
chromatography
HILIC: Hydrophilic interaction liquid
HPLC: High performance liquid chromatography
IBMX: isobutyl-methylxanthine
IgG: Immunoglobulin G
LC: Liquid Chromatography
MALDI: Matrix-assisted laser desorption ionisation
Man: Mannose
MGAT3: Mannosyl (β-1,4-)-glycoprotein β-1,4-*N*-acetylglucosaminyltransferase
MP: membrane proteins
mRNA: messenger ribonucleic acid
MS/MS or MS2: tandem mass spectrometry
MS: Mass Spectrometry
MSC/s: mesenchymal stem cell/s
N.A: Numerical Aperture
NAN1: α2-3 neuraminidase from *Streptococcus pneumonia*
NeuAc: *N*-acetylneuraminic acid
NeuGc: *N*-glycolylneuraminic acid
NMR: Nuclear magnetic resonance
ns: not significant
NSAF: Normalized Spectral Abundance Factor
NT5E: 5'-nucleotidase, ecto
PBS: Phosphate Buffered Saline
PBST: Tween 20 in Phosphate Buffered Saline
PEI: Ethylene imine polymer Solution
PGC: Porous Graphitized Carbon
PHA-E: Phaseolus vulgaris erthyroagglutinin
PLIN1: Perilipin-1
PLIN4: Perilipin-4
PM: plasma membrane
PNGase A: Peptide-*N*-glycosidase A
PNGase F: Peptide-*N*-Glycosidase F
PPARγ2: Peroxisome proliferator-activated receptor γ2
RP: Reversed phase
SD: Standard Deviation
SDS-PAGE: sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPE: Solid-phase extraction

Sulph: Sulphate
SVF: Stromal Vascular Fraction
TBP: TATA-box binding protein
TOF: time of flight
v/v: volume/volume
w/v: weight/volume
YWHAZ: tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta
2-AA: 2-aminobenzoic acid
2-AB: 2-aminobenzamide

Chapter 1:

Introduction

1.1 Mesenchymal Stem Cells

1.1.1 What are stem cells?

Stem cells are defined by their ability to self-renew and differentiate into different specialised cell types, making them potentially the most versatile cell source for therapeutic use in humans (Zavan, Vindigni *et al.*, 2010). The two classical types of stem cells are embryonic stem cells (ESCs) and adult stem cells (Lin, Xin *et al.*, 2010), with a recent third class emerging called induced pluripotent stem cells (iPSCs), where somatic cells are reprogrammed back to a pluripotent ESC-like state. ESCs are derived from the inner cell mass during the blastocyst stage of development in embryogenesis (Lutolf, Gilbert *et al.*, 2009; Yu and Thomson, 2008; Zuk, 2010), while adult stem cells are found in post-natal tissues (Lindroos, Suuronen *et al.*, 2011; Zuk, Zhu *et al.*, 2001) and their use in scientific research avoids many of the ethical considerations that must be taken into account when working with ESCs.

Adult stem cells are divided into two groups: haematopoietic stem cells and mesenchymal stem cells (MSCs). MSCs are found in homeostatically controlled microenvironments called 'niches', an idea first introduced in 1978 (Schofield, 1978). The niche environment is defined as: all the cells immediately surrounding the MSCs, irrespective of whether they are also stem cells, as well as the extracellular matrix and any soluble molecules (Kolf, Cho *et al.*, 2007). MSCs slowly divide to maintain a constant reservoir in the niche, but a portion are released in response to disease or injury to differentiate and regenerate tissue (Lutolf, Gilbert *et al.*, 2009). The niche is responsible for maintaining MSCs in their undifferentiated state and must regulate stem cell division in conjunction with both tissue and cell repair, and homeostasis requirements (Kolf, Cho *et al.*, 2007; Voog and Jones, 2010). Niches have been reported in adipose tissue, muscle, brain, eye, bone marrow and salivary glands, as well as other organs (Kolf, Cho *et al.*, 2007; Murray, West *et al.*, 2014); however bone marrow, umbilical cord blood and adipose tissue are the three main stem cell sources for human studies and therapeutics (Markarian, Frey *et al.*, 2014).

1.1.2 MSCs in regenerative medicine

Traditionally, the use of MSCs in regenerative medicine focused on the *ex vivo* differentiation of MSCs followed by delivery or re-engraftment to the site of tissue damage (Blaber, Hill *et al.*, 2013). It is now clear that undifferentiated MSCs can induce tissue repair through various mechanisms including immunomodulatory, angiogenic, anti-apoptotic and anti-scarring properties by paracrine signalling (Trayhurn and Wood, 2004). Transplantations of MSC-containing umbilical cord blood and bone marrow from un-related donors have been used in the treatment of graft-versus-host disease (GVHD) in 682 adult acute leukaemia patients (Rocha, Labopin *et al.*, 2004). Acute GVHD grades II, III

and IV were lowered by 13 % in umbilical cord treated patients compared to bone marrow treated patients, and the incidence of chronic GVHD reduced by 16 %. The use of umbilical cord blood in the treatment of Wiskott-Aldrich syndrome, an X-linked primary immunodeficiency causing mortality in childhood, has also shown promise, again with a lowered risk of GVHD, and is a disease typically cured by bone marrow transplantation (Slatter, Bhattacharya *et al.*, 2006; Slatter and Gennery, 2006). Despite MSC-containing umbilical cord blood showing superior therapeutic outcomes, the main challenge with its use is its low MSC content. Although the combination of multiple unrelated umbilical cord blood units can be combined for transplant, the average sample size of a single unit is considered small (Ringden, Okas *et al.*, 2008). Although cord blood is obtained through less invasive methods than the painful process of bone-marrow collection, controversy still surrounds its ability to serve as an alternative source of MSCs to bone marrow, with varying success reported in MSC isolation (Kern, Eichler *et al.*, 2006).

1.1.3 Adipose tissue as a source of MSCs

The discovery of adipose-derived stem cells (ADSCs) heralded a new era for the involvement of adipose tissue in stem cell research and human therapeutics. ADSCs are today, a popular source of stem cells for both research and regenerative medicine (Zuk, 2013), and hold advantages over other traditional sources such as bone marrow. While positive therapeutic outcomes have been observed in the use of bone marrow as described above, the number of MSCs harvested is low in comparison to the numbers collected from adipose tissue (Lindroos, Suuronen *et al.*, 2011; Zavan, Vindigni *et al.*, 2010). Bone marrow-derived stem cells make up approximately 0.0001 0.0004 % of the marrow, whereas up to 2 % of nucleated cells in adipose tissue are stem cells (Strem, Hicok *et al.*, 2005). Subcutaneous adipose depots are an accessible, abundant, and re-plenishable source of MSCs not only for research into tissue engineering, but also for potential autologous regenerative medicine which avoids issues of immunogenicity and rejection (Baer and Geiger, 2012; Mitchell, McIntosh *et al.*, 2006). Potential targets for ADSC treatment include soft tissue injuries which have been trialled in diabetic mice (Amos, Kapur *et al.*, 2010). The use of autologous, adipogenically differentiated ADSCs injected subcutaneously into depressed scars (n=31) showed an average recovery in volume of 74.6 % after 12 weeks (Kim, Kim *et al.*, 2011). There is concern however, with the widespread use of ADSCs in cosmetic cases, for example for their use in post-mastectomy breast cancer reconstruction, as ADSCs may secrete cytokines that increase proliferation of any remaining cancer cells (Zimmerlin, Donnenberg *et al.*, 2011). Other targets for the use of ADSC based therapeutics include: repair of skeletal tissues, ischemic injuries and immune disorders and the progress of preclinical and clinical translational studies in these areas have been comprehensively review by Gimble *et al*. (2012).

1.1.4 Adipose-derived stem cells in vitro

Adipose tissue was once regarded simply as a fat store and energy reservoir (Illouz, Sterodimas *et al.*, 2011; Rosen and MacDougald, 2006), however it is now recognised as a major endocrine organ with roles in modulating metabolism, immunity and satiety (Ferris and Crowther, 2011). There are two types of adipose tissue, white and brown, with brown adipose tissue present in newborns but practically absent in adults leaving white adipose as the predominant type (Lin, Xin *et al.*, 2010). Adipose tissue is a complex, highly vascularised heterogeneous tissue, comprised of several cell types including adipocytes, fibroblasts, macrophages, vascular smooth cells and MSCs (Tapp, Hanley *et al.*, 2009). Each white adipocyte is filled with a single lipid droplet (Lin, Xin *et al.*, 2010). The MSCs found in adipose tissue are sometimes referred to as 'pre-adipocytes' and it was in 2001 that researchers first reported their stem cell-like differentiation capabilities into multiple mesodermal lineages (Zuk, Zhu *et al.*, 2001). This was the first of many studies to confirm the presence of a multi-lineage MSC population in human adipose tissue. These cells, although first termed 'processed lipoaspirate cells' are now more commonly called adipose-derived stem cells (ADSCs).

ADSCs can be harvested from adipose tissue collected in routine liposuction procedures carried out for cosmetic reasons at hospitals or cosmetic clinics. Following harvest of the adipose tissue, the lipo-aspirate is in particulate form and after collagenase digestion and centrifugation, the ADSCs are contained amongst the cells of the stromal vasculature, which functions to support the growth of adipocytes and differentiation of pre-adipocytes into adipocytes *in vivo* (Lin, Xin *et al.*, 2010). The physical separation of ADSCs from other cells in the stromal vascular fraction (SVF) occurs during *in vitro* cell culturing conditions because of the plastic-adherent nature of ADSCs which allows a homogeneous stem cell population of ADSCs to emerge on the tissue culture flask surface after cell culture expansion (Mitchell, McIntosh *et al.*, 2006).

By 2006, biological and clinical interest in ADSCs as an alternative MSC source for potential therapeutic use had increased dramatically, however the increasing number of new investigators to the field resulted in non-standardised harvesting methods and a common definition of an adipose-derived MSC was required. The Mesenchymal and Tissue Stem Cell Committee of the International Society for Cellular Therapy proposed a set of three criteria to define an ADSC and thereby allow confident comparisons of studies from different research groups (Dominici, Le Blanc *et al.*, 2006). The criteria proposed were that a true adipose-derived MSC has:

1. Plastic adherence capability when maintained in standard tissue culture flasks
2. Specific surface antigen expression as measured by flow cytometry, namely CD105, CD73 and CD90, and must lack CD45, CD34, CD14, CD11b, CD79α or CD19 and HLA class II

3. Multipotent mesenchymal lineage differentiation potential, in particular the classical mesenchymal lineages producing osteoblasts (osteogenesis), chondrocytes (chondrogenesis) and adipocytes (adipogenesis).

1.2 Adipogenic differentiation of ADSCs

1.2.1 Adipogenesis in vivo

Adipogenesis is the process of ADSC proliferation followed by differentiation into mature adipocytes (Minteer, Marra *et al.*, 2013). The process occurs in adipose depots across the body, involves marked changes in cell morphology and gene expression, and is dependent on paracrine signals from within the stem cell niche (Ali, Hochfeld *et al.*, 2013) such as insulin-like growth factor-1 (Greenberg and Obin, 2006). At a molecular level, adipogenesis *in vivo* is controlled chiefly by the transcription factor peroxisome proliferation activated receptor-gamma (PPARγ) which stimulates lipid uptake (Galateanu, Dinescu *et al.*, 2012). ADSCs first commit to an adipogenic lineage by becoming pre-adipocytes in a process called 'determination' which is irreversible (Ali, Hochfeld *et al.*, 2013; Rosen and MacDougald, 2006). Determination involves the gradual reduction of development options until the cell is capable of becoming only a single cell type and is the responsibility of tissue-specific genes which are present in all cells (Ali, Hochfeld *et al.*, 2013).

The second stage of adipogenesis is called 'terminal differentiation' and it is at this stage that pre-adipocytes begin to take on the unique morphology of a mature adipocyte (Rosen and MacDougald, 2006). Initially, small lipid-filled vesicles form in the cell, gradually swelling with lipid before eventually fusing to result in a single large lipid vesicle in the mature adipocyte (Figure 1.1) (Moreno-Navarrete and Fernández-Real, 2012). Mature adipocytes are referred to as unilocular and their large central lipid vesicle can be between 80-120 µm in diameter, leaving only a thin layer of cytoplasm (1-2 µm) between the vesicle and the plasma membrane (Shen and Chen, 2008). This results in the nucleus being pushed to the cell periphery and the unique morphology associated with mature adipocytes (Moreno-Navarrete and Fernández-Real, 2012).

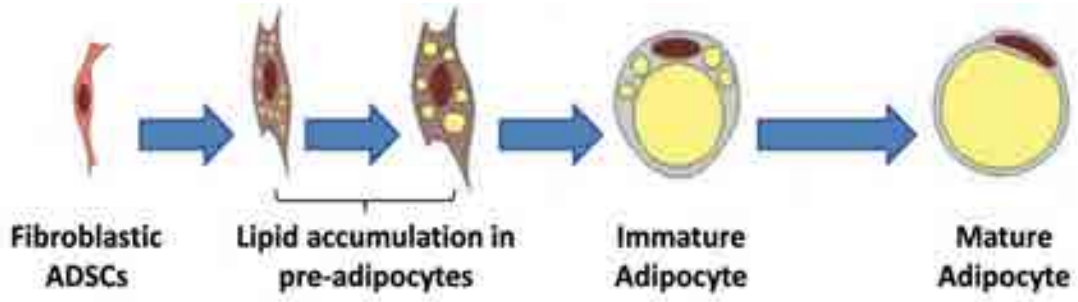


Figure 1.1 The process of adipogenesis: fibroblast-like pre-adipocytes mature into lipid-filled unilocular mature adipocytes. Figure modified from Ali *et al*. (2013)

1.2.2 Adipogenesis in vitro

Adipogenesis *in vitro* encapsulates many of the changes of adipogenesis *in vivo*, with cessation of cell division and the development of morphological changes typical of maturing adipocytes (Bunnell, Estes *et al.*, 2008). This makes *in vitro* models of adipogenesis ideal to study the mechanisms underpinning stem cell differentiation and adipocyte biology (Poulos, Dodson *et al.*, 2010). Adipogenesis *in vitro* is a well characterised process, with the mouse 3T3-L1 pre-adipocyte cell line the most commonly studied, and in which differentiation occurs within approximately one week (Ali, Hochfeld *et al.*, 2013). The three main ingredients used in the adipogenic cocktail to promote differentiation of 3T3-L1 cells are insulin, dexamethasone and 3-isobutyl-1-methylxanthine (IBMX) (Scott, Nguyen *et al.*, 2011).

The adipogenic differentiation of human ADSCs *in vitro* requires the introduction of a cocktail of different inducers that are applied to the cells once they have grown to approximately 90 % confluency. The *in vitro* culture and differentiation of ADSCs remains a non-standardised process, with a variety of protocols reported in the literature within and between ADSCs from the adipose tissue of different species (Katz, Tholpady *et al.*, 2005). Although the main components of the adipogenic differentiation cocktail are similar to the mouse model, and are established and widely accepted, inter-laboratory differences remain. In the adipogenic differentiation of human ADSCs, three to four inducers of adipogenesis are used, such as indomethacin, insulin, IBMX, and dexamethasone (Scott, Nguyen *et al.*, 2011). Indomethacin, and IBMX in conjunction with dexamethasone, are agonists of PPARγ, the key mediator of adipogenesis, while insulin is widely used to promote adipogenic differentiation (Bunnell, Estes *et al.*, 2008; Scott, Nguyen *et al.*, 2011). The adipogenic differentiation protocol used in this thesis was based on a method employed by a pioneer in the field of ADSC biology (Zuk, Zhu *et al.*, 2001) and contains the four inducers listed above.

1.2.2.1 Methods to evaluate in vitro *adipogenesis*

There are several methods for evaluating the adipogenic differentiation of MSCs *in vitro*, such as staining of lipid vesicles, measuring lipoxegenase enzyme activity, and assessing adipocyte-specific gene expression using quantitative real-time polymerase chain reaction (qRT-PCR), and antibody-based protein detection (Bunnell, Estes *et al.*, 2008). Of these methods, the staining of lipid vesicles is the most direct method and available lipid dyes include Nile Red, BODIPY and Oil Red O (Mehlem, Hagberg *et al.*, 2013; Nishimura, Manabe *et al.*, 2008; Zuk, Zhu *et al.*, 2001). Oil Red O is the most commonly used lipid dye and its action is based on its hydrophobic nature which causes it to move from its solvent to associate with accumulating lipids in differentiating stem cells (Mehlem, Hagberg *et al.*, 2013).

The expression of specific marker genes, as determined by qRT-PCR, provides a quantitative indicator to assess adipogenic differentiation *in vitro*. Common adipogenic lineage genes that are up-regulated during *in vitro* adipogenic differentiation and can be used to assess degree of differentiation include fatty acid binding protein (FABP4 or aP2), lipoprotein lipase (LPL), glutathione synthetase (GLUTA), leptin, PPARγ1 and PPARγ2, perilipin, CD36, ADRP and cyclophilin (Bunnell, Estes *et al.*, 2008; Jefcoate, Wang *et al.*, 2008). Conversely, the down-regulation of MSC-specific markers during differentiation such as CD90, CD73 and CD105 (Blaber, Webster *et al.*, 2012; De Gemmis, Lapucci *et al.*, 2006) can also be used to monitor differentiation.

1.3 Stem cell differentiation and glycosylation

1.3.1 Membrane protein glycosylation

The most prominent component of the stem cell surface, which is positioned in such a way to facilitate interaction between the stem cell and its niche, is the glycocalyx. The glycocalyx is present on the surface of all cells and is composed of various glycoconjugate types such as glycoproteins, glycolipids, proteoglycans and glycosaminoglycans (Lanctot, Gage *et al.*, 2007). The formation of glycoconjugates is a co- or post-translational modification with the two main types of glycosylation being the addition of *N*- or *O*-linked glycans to the nascent protein backbone. *N*-linked glycan addition occurs via the covalent attachment of a pentasaccharide core (Mannose$_3$*N*-acetylglucosamine$_2$) to asparagine residues in the consensus sequence Asn-Xxx-Ser/Thr, where asparagine may be followed by any amino acid except proline (Weerapana and Imperiali, 2006). The common *N*-linked glycan pentasaccharide core can be elongated by the addition of sugars to the outer two mannose residues of the core in various linkage configurations resulting in three main classes of *N*-glycans: high-mannose, complex and hybrid, with a fourth, more recently reported class in humans, called pauci mannose type (Thaysen-Andersen, Venkatakrishnan *et al.*, 2015) (Figure 1.2a).

Mucin-type *O*-linked glycans are attached to serine or threonine residues via an *N*-acetylgalactosamine (GalNAc) residue and can have up to as many as eight different core types (Dube, Prescher *et al.*, 2006). Core types 1 and 2 are the two most common *O*-linked glycan cores found in humans and are depicted in Figure 1.2b. Other types of *O*-linked glycans involve attachment to the protein via mannose (*O*-mannosylation), fucose (*O*-fucosylation) or *N*-acetylglucosamine (GlcNAc; *O*-GlcNAcylation) residues. Unlike the translation of gene transcripts into proteins, the assembly of both of *N*- and *O*-linked glycans is a non-template driven process (Bern, Brito *et al.*, 2013; Nairn, Aoki *et al.*, 2012; Nairn, York *et al.*, 2008) aided by the action of approximately 700

proteins including glycosidase and glycosyltransferase enzymes (Moremen, Tiemeyer *et al.*, 2012), co-factors, transporters, and also activated sugar donors (van Kooyk, Kalay *et al.*, 2013).

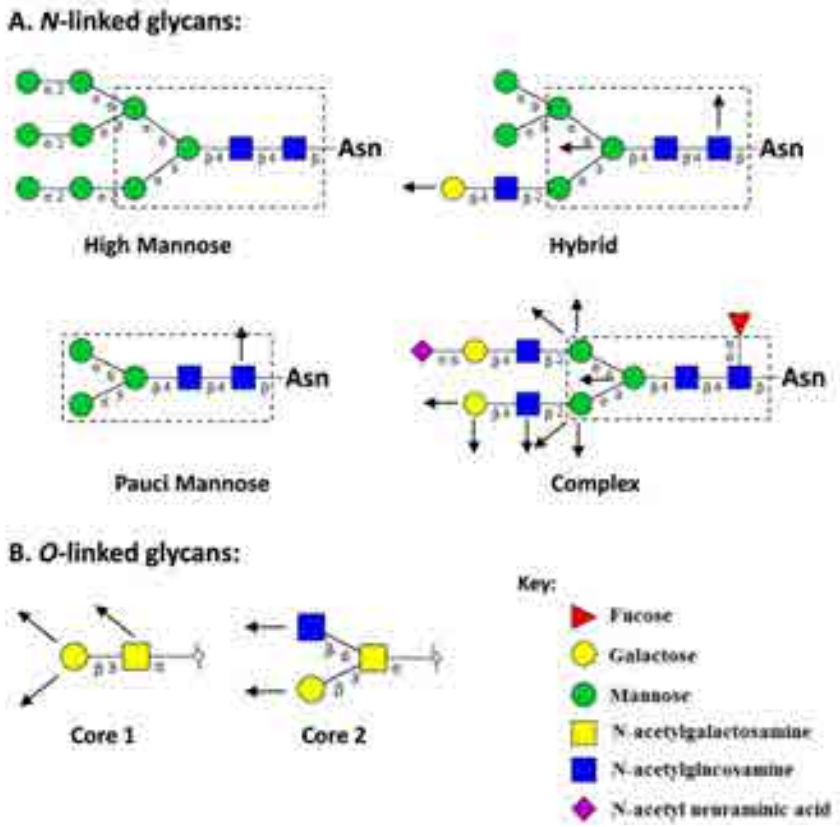


Figure 1.2 *N*-linked glycans are typically classed into four main subtypes: high mannose, hybrid, complex and pauci mannose type which all share in common a pentasaccharide core of $Man_3GlcNAc_2$ (marked in box). Of the eight *O*-linked glycan core types, numbers one and two are the most common in humans. Elongation can occur in different linkage configurations from various positions (marked with arrows).

N- and *O*-glycosylation occurs through the linkage of monosaccharides in linear or branched glycosidic bonds by different pathways to produce a heterogeneous array of glycan structures. The ten monosaccharides used in mammalian glycan production are deoxyhexose or fucose (Fuc), galactose (Gal), mannose (Man), glucose (Glc), *N*-acetylgalactosamine (GalNAc), *N*-acetylglucosamine (GlcNAc), glucuronic acid (GlcA), iduronic acid (IdoA), *N*-acetylneuraminic acid (NeuAc) and xylose (Xyl). These monosaccharides are shown in Figure 1.3 along with their stylised symbolic notation from the Consortium of Functional Glycomics. The glycosylation machinery can also be influenced by epigenetic factors (Anugraham, Jacob *et al.*, 2014), furthering the complexity of this post-translational process.

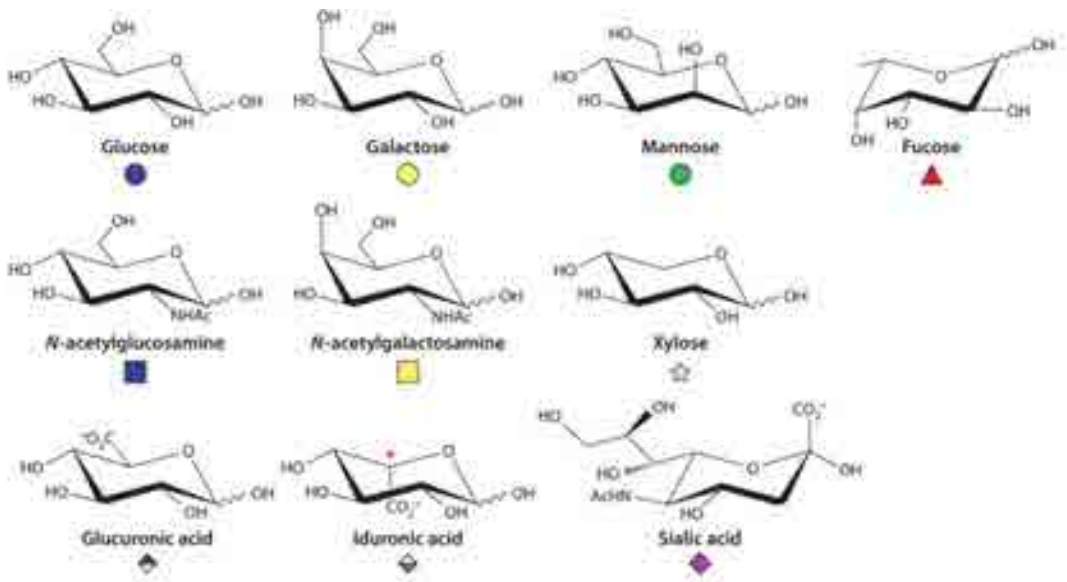


Figure 1.3 Ten monosaccharides used in mammalian glycan production depicted in their chemical structure and symbolic notation from the Consortium of Functional Glycomics, shown as free reducing sugars, *with the exception of iduronic acid which is not found as a free sugar. Figure modified from Rakus and Mahal (2011).

1.3.1.1 N- *and* O-*linked glycan biosynthesis*

The *N*-linked glycan biosynthesis pathway is conserved between species and provides a basis for glycan structural characterisation. The initial steps in the mammalian *N*-linked glycan biosynthetic pathway are shared for all types of *N*-linked glycans, with the most complicated processing occurring in the synthesis of complex-type structures. Briefly, nascent proteins containing the consensus sequence Asn-Xxx-Ser/Thr are translocated to the endoplasmic reticulum (ER) where a lipid-linked, 14 residue glycan precursor is transferred *en bloc* to the asparagine residue by the oligosaccharidetransferase (OST). The non-reducing end monosaccharides of the precursor are then trimmed by various exoglycosidases before a glucose-capped structure containing eight mannose residues is translocated to the Golgi apparatus. Here, glycans are further trimmed and processed. Those that continue to carry five to nine Man residues are classed as high mannose, whereas the addition of the first GlcNAc residue to the trimmed mannose core in the *cis*-Golgi marks a branching point from which hybrid-type or complex-type structures are synthesised. The addition of GlcNAc to produce branching or linear structures is typically followed by a single galactose residue, which may be subsequently extended by NeuAc (sialylation) and/or Fuc (fucosylation), or GlcNAc (including lactosamine repeats: Gal-GlcNAc) and/or sulphation or phosphorylation. Modifications also occur to the *N*-linked glycan core such as the addition of an α1-6 linked fucose to the reducing-end GlcNAc residue or the β1-4 linked addition of GlcNAc to the inner most mannose residue of the core, termed a 'bisecting GlcNAc'. The biosynthetic *N*-linked glycan pathway results in high complexity and diversity and dramatically different *N*-glycosylation patterns can be observed in different cells or

tissues, and also on the same protein in different tissues such as the differential glycosylation of low density lipoprotein receptor-related protein 1 in liver and brain tissues (May, Bock *et al.*, 2003). Examples of tissue-specific unique glycosylation also exist, for example the addition of poly-sialic acid residues to neural cell-adhesion molecule (NCAM) in the brain (Lanctot, Gage *et al.*, 2007).

In contrast, *O*-linked glycan biosynthesis follows a very different pathway. The majority of *O*-linked glycans are typically found in abundance in secreted fluids such as mucus and saliva, and are formed by a process termed 'mucin-type' *O*-glycosylation. *O*-linked glycans can occur as linear or branched structures and their synthesis begins with the addition of a GalNAc to the serine or threonine residues of a peptide. The GalNAc can be extended by GlcNAc, Gal, Fuc and NeuAc, along with potential sulphation of Gal and GlcNAc residues. Of the eight *O*-linked glycan core types, core numbers 5-8 are rarely detected in humans, and of cores 1-4, core 1 and 2 are the most commonly identified.

1.3.2 Glycoproteins in stem cell differentiation

Of the glycoconjugates present on the membrane glycocalyx of a cell, glycoproteins are the major functional components, and have been implicated in many different cellular processes including ligand binding, bacterial attachment and cell differentiation (Hart and Copeland, 2010). The Notch membrane glycoprotein provides an example of where glycosylation affects protein function, in this case cell surface ligand-binding and activation of pathways involving cell lineage determination. Notch is glycosylated with both *N*- and *O*-linked glycans, and its developmental role is determined by the addition of an *O*-linked fucose (*O*-fucosylation) residue to serine/threonine repeats in certain epidermal growth factor repeats. The fucose provides a substrate for the addition of a GlcNAc residue and it is the addition of these two residues that allows further elongation and modulates Notch ligand-binding (Moloney, Panin *et al.*, 2000; Moloney, Shair *et al.*, 2000).

1.3.3 Changes in glycosylation in stem cell differentiation

As membrane glycoconjugates are at the interface between a cell and its external environment, much of the work investigating stem cell glycosylation has been done with the aim of identifying glycosylated cell surface signatures. For example, the glycolipid SSEA-3 and SSEA-4 epitopes, which consist of five to six monosaccharides attached to a lipid via a ceramide tail, are commonly used markers to identify embryonic stem cells as they decrease in abundance upon differentiation, making them ideal markers for ESC purification (Lanctot, Gage *et al.*, 2007).

A number of studies have characterised changes in global protein glycosylation (glycomics) during stem cell differentiation in both human and non-human stem cells derived from various sources (Table 1.1). Investigations into glycoprotein glycosylation changes during iPSC differentiation have

also been undertaken and are primarily carried out using a high throughput approach with lectin arrays (Saito, Onuma *et al.*, 2011; Tateno, Toyota *et al.*, 2011; Toyoda, Yamazaki-Inoue *et al.*, 2011; Wang, Nakagawa *et al.*, 2011). Mass spectrometry (MS)-based investigations of the global glycoprofile of human stem cell differentiation have used bone marrow-derived, embryonic and hematopoietic stem cells (Hamouda, Ullah *et al.*, 2013; Hasehira, Tateno *et al.*, 2012; Heiskanen, Hirvonen *et al.*, 2009; Hemmoranta, Satomaa *et al.*, 2007; Satomaa, Heiskanen *et al.*, 2009). These studies, their key findings and the analytical methods used are summarised in Table 1.1, including investigations that have involved profiling and comparison of the glycome of stem cells with other cells types, but not necessarily with their differentiated progeny (An, Gip *et al.*, 2012; Fujitani, Furukawa *et al.*, 2013). One limitation of these studies is the lack of a 'control' sample, i.e. an *in vivo* produced differentiated cell product for the lineage under investigation, with which to compare and evaluate changes occurring in the *in vitro* differentiation. To our knowledge, the work of Hamouda *et al*. (2013) is the first to investigate the protein glycosylation changes occurring in *in vitro* adipogenic differentiation. The Hamouda *et al*. (2013) study used MALDI-TOF/TOF MS analysis on permethylated glycans released from bone marrow-derived MSCs compared with their adipogenically differentiated counterparts. The results indicated key differences upon adipogenic differentiation, particularly the over expression of two structures in the MSCs with composition (HexNAc)2 (Fuc)1 + (Man)3(GlcNAc)2 and (Hex)2 (HexNAc)2 (Fuc)3 + (Man)3(GlcNAc)2, and two structures over expressed in the differentiated cells with composition (Hex)3(HexNAc)3(Fuc)1 + (Man)3(GlcNAc)2 and (Hex)4 (HexNAc)4 (Fuc)1 + (Man)3(GlcNAc)2. These structures offer promising markers of MSC adipogenic differentiation stage. Their use of tandem MS provided structural glycan information which aided the characterisation of the glycans, however the absence of chromatography prior to MS would result in an absence of isomeric and low abundance glycan structural information.

Table 1.1 Summary of studies characterising and comparing protein glycosylation profiles of human undifferentiated stem cells with their differentiated progeny and other cell types

MSC source	Differentiation lineage/Progeny	Analysis method	Observations	Reference
STUDIES CHARACTERISING CHANGES IN GLYCOSYLATION AFTER STEM CELL DIFFERENTIATION:				
Bone marrow	Osteogen c	- MALDI-TOF/TOF MS - ^{1}H NMR - Exog ycos dase treatment	**- Upon differentiation, *N*-linked glycans:** ↓ H gh mannose-type structures (78 to 58 %) ↑ Comp ex-type structures (12 to 25 %) ↑ Hybr d-type structures (9 to 16 %) **- Features of MSCs:** α2,3-s a y ated structures **- Features of differentiated MSCs:** su phated or phosphory ated *O*- nked g ycans	(He skanen, H rvonen *et al.*, 2009)
Bone marrow	Ad pogen c	- MALDI-TOF/TOF MS - Exog ycos dase treatment	**- Upon differentiation, *N*-linked glycans:** ↑ B antennary fucosy ated structures ↑ 16.5% s a y at on ↑ 5.7% h gh mannose-type structures ↓ Fucosy ated, afucosy ated tr - and tetra-antennary structures **- MSCs showed over expression of structures with composition:** (Hexose)6+(HexNAc)5+(Fuc)1 and (Hexose)7+(HexNAc)6+(Fuc)1 **- Adipogenic progeny showed over expression of structures with composition:** (Hexose)3+(HexNAc)4+(Fuc)1 and (Hexose)5+(HexNAc)4+(Fuc)3	(Hamouda, U ah *et al.*, 2013)
Derma f brob asts (represent ng d fferent ated ce state)	20187 human PSCs from derma f brob asts (represent ng und fferent ated ce state)	- RP-HPLC -MALDI-TOF MS - Conf rmat on w th g ycosy transferase gene express on	**Upon induction of pluripotency:** ↑ H gh mannose-type structures - Sh ft from α2,3 nked s a c ac ds to α2,6 nked - New express on of structures conta n ng α1,2 nked fucose res dues and type 1 LacNAc (Ga β1-3G cNAc) structures	(Haseh ra, Tateno *et al.*, 2012)
ESCs	Embryo d bod es	- MALDI-TOF MS - NMR - Exog ycos dase d gest on - Lect n b nd ng	- Neutra g ycans compr sed two-th rds of the ESC *N*- nked g ycome - H gh mannose and b antennary comp ex-type *N*- nked g ycans were unchanged dur ng d fferent at on ↓ 12% structures conta n ng ≥ 2 fucose res dues	(Satomaa, He skanen *et al.*, 2009)

MSC source	Differentiation lineage/Progeny	Analysis method	Observations	Reference
Hematopo et c stem ce s from cord b ood express ng CD133	Hematopo et c stem ce s from cord b ood not express ng CD133 (CD133 s ost dur ng stem ce d fferent at on)	- MALDI-TOF/TOF MS -Exog ycos dase d gest on - NMR -F ow cytometry us ng ect ns - Measurement of g ycan b osynthes s-re ated genes w th m croarray ana ys s	- B antennary comp ex-type *N*- nked g ycans were enr ched n CD133⁺ce s ↑ H gh mannose-type structures ↑ α2,3 nked s a c ac ds n CD133⁺	(Hemmoranta , Satomaa *et al.*, 2007)
ESCs	12-day d fferent ated embryo d bod es as a mode to study ear y stage d fferent at on	Lect n and ant - carbohydrate ant body mmunof uorescence sta n ng	**Upon differentiation, *N*-linked glycans:** - ↑ 70% n α2,3- nked NeuAc-Ga ep topes - ↓ 32% b sect ng G cNAc structures - A g ycans n both und fferent ated and 12-day d fferent ated ce s conta ned α1,6- nked fucose - 15% of g ycans n ESCs conta ned α1,2- nked fucose wh e th s feature was absent n embryo d bod es	(Weame, W nter *et al.*, 2006)
12-day d fferent ated embryo d bod es	28-day d fferent ated embryo d bod es	Lect n and ant - carbohydrate ant body mmunof uorescence sta n ng	**From 12 to 28 days, increases were observed in:** - S a y at on of α-Ga NAc res dues - Ga β1,3-Ga NAc (T ant gen) ep tope - D -fucosy ated Ga β1-4G cNAc ep tope	(Weame, W nter *et al.*, 2008)
Cells analysed:		**Analysis method**	**Observations**	
STUDIES COMPARING GLYCOSYLATION PROFILES BETWEEN STEM CELLS AND OTHER SOMATIC CELL TYPES:				
a) ESCs b) PSCs der ved from umb ca t ssue, amn ot c and dec dua t ssue c) Norma human ce ne d) Ep the a carc noma ce ne		MALDI-TOF MS	- A ce types conta ned a h gh percentage of h gh mannose-type g ycans - ESC and PSCs expressed s gn f cant y ower eve s of pauc mannose-type structures - Extended core 1 or core 2 *O*- nked g ycans term nat ng w th β1–3 nked ga actose structures were character st c of ESC and PSCs	(Fuj tan , Furukawa *et al.*, 2013)
a) ESCs b) IMR-90 f brob asts c) Norma breast ce ne		- Fou r er transform on cyc otron resonance MS - F ow cytometry - Immunof uorescence	- H gher express on of h gh mannose type structures and both α2,3 and α2,6 nked s a c ac ds n ESCs - *O*- nked g ycans are pr mar y neutra o gosacchar des n ESCs	(An, G p *et al.*, 2012)

HexNAc = *N*-acety hexosam ne; MALDI-TOF/TOF = Matr x ass sted aser desorpt on t me-of-f ght/t me-of-f ght; NMR = nuc ear magnet c resonance; RP HPLC = reversed phase h gh performance qu d chromatography

1.4 Analysis of plasma membrane glycoconjugates

1.4.1 Strategies to study protein glycosylation

Currently, there is no single glyco-analysis technique that provides a complete protein glycosylation profile of a biological system, resulting in the need to use orthogonal techniques that target various aspects of protein glycosylation. These techniques provide different structural details and can be broadly divided into different levels: monosaccharide analysis, analysis of released glycans (glycomics), glycoproteomics and proteomics (Fernandes and Spencer, 2008; Moh, Thaysen-Andersen *et al.*, 2015).

Proteomics-based approaches do not address glycan structure but can be used to identify proteins and glycoproteins present in a sample. The amount of information required for the structural characterisation of the glycoproteins produced by a cell is much greater than of unmodified proteins due to the heterogeneity and structural diversity found at the glycan level (Thaysen-Andersen and Packer, 2014). Bottom up 'shotgun' methods are most common in proteomics investigations, whereby proteins are identified based on the MS/MS analysis of their peptide fragments and by matching to gene-based predicted peptide sequences. In addition, the analysis of enzymatically de-glycosylated samples can provide information about *N*-glycosylation sites by a 0.9840 Da mass shift arising from deamidation of previously glycosylated asparagine residues (Pan, Chen *et al.*, 2011). However, this technique must be used with caution due to naturally occurring deamidation that can result in false-positive glycosylation site identification (Leymarie, Griffin *et al.*, 2013; Palmisano, Melo-Braga *et al.*, 2012). Proteomics workflows and methods will be further discussed in Section 1.5 and Chapter 4.

Monosaccharide analysis provides compositional information about the monosaccharide residues present in the sample and can be used as a guide in downstream structural elucidation in glycomics analyses (Mariño, Bones *et al.*, 2010). In the analysis of non-mammalian samples, monosaccharide analysis can be very useful in determining the biosynthetic glycosylation machinery present in the originating sample.

Glycomics-based approaches focus on the structural characterisation and quantitation of glycans globally released from a biological sample but do not involve identification of the peptide or protein carrying the glycans. The analysis of the released glycans faces the challenge of a high degree of compositional similarity due to the limited number of monosaccharide building blocks. Structures must be distinguished from one another also in terms of linkage, branching and anomericity, with the added complexity of the presence of isobaric isomers. This complexity adds to the challenges of the other 'omics analyses where the main challenge is the large number of analytes (Moh, Thaysen-

Andersen *et al.*, 2015). Also, in comparison to other 'omics methodologies such as proteomics and genomics, the technological advancement in structural characterisation of glycans has been slow. The approaches used in the analysis of released glycans will be discussed in greater detail in Section 1.4.3 and Chapter 3.

The complete characterisation of glycoproteins requires the integration of proteomics and glycomics information and remains highly challenging at a system-wide level (Pan, Chen *et al.*, 2011). Recent glycoproteomics-based investigations of complex protein mixtures resulted in glycoprotein identification, glycosylation site identification/s, determination of degree of site occupancy and the monosaccharide compositions of the attached glycans at each site (Parker, Thaysen-Andersen *et al.*, 2013). Current whole cell/tissue glycoproteomic analyses however are limited in their capacity for detailed glycan structural characterisation (Moh, Thaysen-Andersen *et al.*, 2015). The fragmentation information gained from glycopeptide MS/MS is not sufficient for determining glycan structures, so ideally a global glycomics analysis is carried out in parallel to provide more detailed structural information. Full glycoproteomics analysis presently can be more readily achieved when dealing with an isolated, purified glycoprotein (Sumer-Bayraktar, Kolarich *et al.*, 2011; Sumer-Bayraktar, Nguyen-Khuong *et al.*, 2012).

1.4.2 Release of protein N- *and* O-*linked glycans*

Various methods currently exist for the chemical or enzymatic release of *N*- and *O*-linked glycans from glycoproteins for global glycomics analyses. Mammalian *N*-linked glycans are typically released using endoglycosidase peptide-*N*-glycosidase F (PNGase F). PNGase F cleaves *N*-linked glycans by hydrolysis of the bond between the asparagine residue of the protein and the linked reducing end GlcNAc of the *N*-linked glycan and results in reducing glycans that can be reduced using sodium borohydride (Jensen, Karlsson *et al.*, 2012). Reduction of the glycan is recommended to overcome increased complexity caused by the separation of the α- and β-anomers of reducing sugars in subsequent chromatography. Peptide-*N*-glycosidase A (PNGase A) is an alternative endoglycosidase with broader specificity that can be used to release glycans from glycopeptides and is more suitable for the analysis of glycans from plant and insect samples that have structures resistant to PNGase F cleavage. *N*-linked glycans that contain fucose linked to the reducing end GlcNAc in an α1,3 linkage are an example of such structures and can be cleaved by PNGase A. This linkage type does not occur in mammalian glycosylation, as the fucose at this position is added in a α1,6 linkage (Fernandes and Spencer, 2008).

An enzymatic method for removal of *O*-linked glycans currently exists and involves the use of an O-glycanase; however, this enzyme is only specific for Core 1 type *O*-linked glycans (GalNAc-Gal) that do

not carry any extensions. Chemical methods for *O*-linked glycan release are therefore typically favoured. Reductive β-elimination is the most common chemical method and involves using highly basic conditions to cleave the glycan-peptide bond, without glycan degradation or peeling of the reduced terminus (Mechref, 2011; Moh, Thaysen-Andersen *et al.*, 2015). An alternative chemical release method involves the use of anhydrous hydrazine, which produces *O*-linked glycans that may be further derivatised or labelled. Hydrazinolysis can also be used to release *N*-linked glycans by using higher temperatures and a longer incubation period compared to *O*-linked glycan release (Fernandes and Spencer, 2008). The major concern of release of glycans by hydrazinolysis is peeling of the reducing terminus, although reportedly this can be avoided through the addition of EDTA to the reaction mixture (Kozak, Royle *et al.*, 2012). Other non-reductive *O*-linked glycan release methods include the use of dimethylamine and ammonia which are reported to produce low levels of peeling (Maniatis, Zhou *et al.*, 2010).

1.4.3 Analysis methods for the characterisation of released N- *and* O-*linked glycans*

Following *N*- and *O*-linked glycan release, various technologies are available for their analysis including capillary electrophoresis (CE), nuclear magnetic resonance (NMR), lectin arrays, and gas and liquid chromatography (GC and LC respectively) coupled with tandem mass spectrometry (MS/MS) (Robinson, Artpradit *et al.*, 2012).

Released glycans can be analysed in their native state and can also be further derivatised by permethylation or labelled with various high performance liquid chromatography (HPLC)-MS compatible labels including 2-aminobenzamide (2-AB), 2-aminobenzoic acid (2-AA) or procainamide (Fernandes and Spencer, 2008; Klapoetke, Zhang *et al.*, 2010). Derivatisation is performed to allow UV detection of glycans and sometimes to increase sensitivity of MS detection. The additional clean-up methods that are required post-derivatisation along with the risk of incomplete derivatisation can however result in sample loss or skewing (Moh, Thaysen-Andersen *et al.*, 2015).

1.4.3.1 Separation of released glycans

Due to the structural complexity found in a heterogeneous pool of released glycans, analysis of the diversity and of isomeric structures is only possible by the use of separation methods coupled to various detectors. Capillary electrophoresis (CE) offers high separation efficiency of glycan samples, including isomeric structures, and has the added advantage of automation of high throughput samples. This has been performed by using a multi-capillary format as designed for high throughput DNA sequencing (Callewaert, Geysens *et al.*, 2001). Separation of glycans is mediated by an electric field, but because the majority of glycans are not charged, labelling with a negatively charged chromophore or fluorophore is required (Szabo, Guttman *et al.*, 2010). The most common

fluorophore used is 8-aminopyrene-1,3,6-trisulfonic acid (APTS), and allows sample detection with laser induced fluorescence. Comparison between the CE trace and that of a glycan standard subsequently allows a possible identification of the structure but does not allow for detailed structural analysis (Rakus and Mahal, 2011). The coupling of CE to MS/MS provides a platform with which to comprehensively characterise glycans with high sensitivity. However, this coupling requires consideration of some areas of incompatibility, for example a discrepancy in required flow rates (Mechref, 2011) and the higher analyte concentration required for MS.

An alternative to CE, and a commonly used separation method, is HPLC and has the added advantage of better compatibility with MS and other analytic devices (Rakus and Mahal, 2011). Although HPLC has a lower throughput, it offers greater resolving power than CE with reversed (RP) and normal phases being the two main types of chromatography used.

For analysis using RP chromatography, glycans must first be derivatised for retention on the RP column by permethylation or addition of a chromophoric label (Ruhaak, Zauner *et al.*, 2010). This increase in overall hydrophobicity allows retention and reliable separation. RP separation of permethylated glycans has been successfully employed in the analysis of fetuin, α-1 acid glycoprotein and ribonuclease B where a miniaturised solid-phase permethylation and online solid-phase purification system was utilised (Desantos-Garcia, Khalil *et al.*, 2011). An additional strategy in RP separations is the incorporation of a lipophilic ion pairing reagent to the mobile phase which can enhance the retention of charged molecules on the stationary phase (Kailemia, Ruhaak *et al.*, 2013). This technique relies on an important choice of ion pairing agent and also the pH of the system. Successful coupling of RP ion pairing to ESI-MS/MS has been achieved in the analysis of negatively charged glycosaminoglycans (Thanawiroon, Rice *et al.*, 2004) which are characteristically long, linear, sulphated alternating disaccharide units (Estrella, Whitelock *et al.*, 2009), but to our knowledge has not been used in the analysis of *N*- and *O*-linked glycans.

Hydrophilic interaction liquid chromatography (HILIC) on amide columns is a type of normal phase chromatography where the mobile phase solvent is an organic solvent, less polar than the mobile aqueous phase (Alpert, 1990). In HILIC separations, a binary gradient is employed with the percentage of the aqueous fraction increased up to 50 % for glycan sample elution (Zauner, Deelder *et al.*, 2011). Glycan separation using HILIC is reliant on many different sample factors such as size, charge, composition, linkage and glycan branching (Mariño, Bones *et al.*, 2010). Ultimately the discerning feature that determines retention is based on the number of polar groups, resulting in predictable elution patterns that aid subsequent data interpretation (Zauner, Deelder *et al.*, 2011). Derivatisation or labelling of glycans aids in the separation of structural isomers and the use of a

chromophore label can also allow the incorporation of UV-based detection and quantitation to be incorporated into a HILIC HPLC-LC-MS/MS analysis. The natural properties of HILIC amide materials and their affinity for highly polar hydrophilic analytes make it a more suitable separation surface for glycans in comparison to RP (Zauner, Deelder *et al.*, 2011).

While non-derivatised glycans can also be analysed using HILIC, chromatography using porous graphitised carbon (PGC) chromatography is an attractive alternative. Thomsson *et al*. (1999) compared the separation power of an amine-bonded HILIC column with PGC in the analysis of sulphated mucin oligosaccharide alditols. Isomeric structures were separated on both columns while co-elution also occurred, demonstrating their use as complementary methods. Although its exact mechanism is not well understood, particular glycan structural features have been reported to influence elution behaviour (Pabst, Bondili *et al.*, 2007), for example isomers carrying a bisecting GlcNAc elute before their non-bisected counterparts (Stavenhagen, Kolarich *et al.*, 2015). Elution time is also influenced by sialylation, with isomers carrying α2-6 linked sialic acid residues eluting before isomers carrying α2-3 linked sialic acids (Anugraham, Jacob *et al.*, 2014; Nakano, Saldanha *et al.*, 2011). PGC-LC ESI-MS/MS analyses have been successfully utilised in a number of different sample types: saliva (Everest-Dass, Jin *et al.*, 2012), tammar wallaby milk (Wongtrakul-Kish, Kolarich *et al.*, 2013), ovarian cancer cell lines (Anugraham, Jacob *et al.*, 2014), human tears (Nguyen-Khuong, Everest-Dass *et al.*, 2015), sputum from cystic fibrosis patients (Venkatakrishnan, Thaysen-Andersen *et al.*, 2015), human serum (Chu, Niñonuevo *et al.*, 2009), dried blood spots (Ruhaak, Miyamoto *et al.*, 2012) and chicken intestinal mucins (Struwe, Gough *et al.*, 2015) amongst others. Due to the strong capability for separation of closely related glycan structures by PGC, this method was chosen for separation of the glycans released from membrane glycoproteins in this thesis.

1.4.3.2 Structural analysis of released glycans

NMR is the most effective analytical technique in that the spectra provides highly detailed information regarding glycan composition, linkage, anomeric conformation and the presence of any chemical modifications (Robinson, Artpradit *et al.*, 2012). NMR is a useful tool when three-dimensional information about a glycan sample is required. Although NMR allows recovery of sample material for further use after analysis, large quantities (typically > 1 mg) of a highly purified starting sample (>95 %) are required (Mariño, Bones *et al.*, 2010), thus limiting its application to complex biological samples. Minor components may also be lost during sample clean up and preparation.

Lectin arrays are also often used to give structural glycan detail and involve a panel of lectins immobilised at high density in an array format similar to gene microarrays (Robinson, Artpradit *et al.*, 2012). Lectin arrays are used as a high throughput method in which glycan-lectin binding is assessed

through application of the glycan sample to the lectin array surface. Each lectin has a specificity for particular glycan structural characteristics and binding is determined through fluorescent detection of labelled ligands. It must be noted that not every glycan has a corresponding lectin for recognition, and some lectins bind to several different structural types (Katrlik, Svitel *et al.*, 2010) making interpretation of precise structural determinants difficult.

MS is currently the most used method for the characterisation of released *N*- and *O*-linked glycan structures. MS and MS/MS can provide information about glycan composition, modifications, sequences and linkages (Robinson, Artpradit *et al.*, 2012). Coupled with HPLC, separation of complex mixtures can be achieved, thus allowing a greater depth of information to be obtained including separation of isomers.

1.4.3.3 Characterising glycans by MS analysis

Advancements in MS techniques have enabled the highly detailed characterisation of glycans in recent years and unearthed glycosylation changes in a number of different diseases (Stavenhagen, Kolarich *et al.*, 2015). Ionisation of the glycan sample is typically conducted using either matrix assisted laser desorption ionisation (MALDI) or electrospray ionisation (ESI) (Kailemia, Ruhaak *et al.*, 2013; Robinson, Artpradit *et al.*, 2012).

MALDI: For MALDI, the sample is mixed with an organic solvent and matrix solution before spotting onto a target plate. The dried sample spot is then irradiated with an ultraviolet laser, the energy is absorbed by the matrix and transferred to the analyte, thereby ionising it. Advantages of MALDI include its ability to ionise at high mass ranges and tolerance to contaminants (Kailemia, Ruhaak *et al.*, 2013). Unfortunately, the high energy conditions in MALDI can lead to loss of acidic species such as sialic acid from the structures, although this can be minimised by the use of different strategies to stabilise the acidic groups. Methylation of the carboxyl group of sialic acid or permethylation of whole glycans are approaches to stabilise sialylated structures in MALDI-MS (Morelle and Michalski, 2007), but can also result in the loss of other negatively charged, unstable modifications such as sulphates and phosphates (Kailemia, Ruhaak *et al.*, 2013).

ESI: ESI provides an alternative ionisation method which is considered 'softer' in comparison to MALDI, and is useful in the analysis of structures containing labile modifications without derivatisation. Analyte ions are first passed through a fine tipped capillary with a potential of between 1-4 kV applied at the tip, with the resulting potential difference between the tip and capillary generating charged sample droplets (Kailemia, Ruhaak *et al.*, 2013). The positioning of the tip at the mass spectrometer inlet in the presence of high temperatures and drying gas vaporises the droplets as they enter the mass spectrometer, producing mass/charge (*m/z*) values for each glycan

from which the mass and thus the possible monosaccharide composition can be determined. ESI can result in the formation of multiply charged ions which adds another layer of complexity to data analysis, but which also facilitates detection of structures with high mass that would not otherwise be detected in mass spectrometers with limited mass range. In contrast to MALDI however, ESI is less tolerant to contaminants. Consequently, an important consideration in the use of ESI is the avoidance or removal of detergents, such as sodium dodecyl sulfate (SDS), and of salt contaminants often used in sample preparation which may have high ionisation efficiency and mask glycan ion signals (Robinson, Artpradit *et al.*, 2012).

Ionisation polarity: A second consideration concerning ionisation in MS workflows is the choice of ionisation polarity mode. Acidic glycans analysed in positive ion mode typically have low ionisation efficiency compared to neutral structures. This low ionisation of sialylated structures can be combated by the use of negative ion mode; however, under negative ionisation, sialylated structures will ionise in preference to neutral structures and produce 'ion suppression' which can impact quantitation.

MSn fragmentation: While glycan monosaccharide composition can be predicted from *m/z* values at the MS level, MS/MS or MSn fragmentation provides information that enables structural characterisation of the glycan sample. The most commonly used fragmentation method for carbohydrate analysis is collision-induced dissociation (CID), which involves the collision of an inert gas, typically argon or helium, with the selected precursor ion. Two types of fragments are generated from this: glycosidic fragments between monosaccharide residues, and cross-ring fragments of hexose residues which provide glycosidic bond linkage information. The nomenclature for glycan fragmentation was first proposed by Domon and Costello (1988) and is shown in Figure 1.4. Glycosidic cleavages (generating B and Y ions) dominate the MS/MS spectra of positively charged ions while the more information-rich cross-ring fragments (A and X ions) dominate MS/MS spectra of negatively charged ions. Fragmentation in negative ion mode is more specific than in positive ion mode due to the generation of a $[M-H]^-$ ion which fragments in a predictable manner (Everest-Dass, Abrahams *et al.*, 2013; Harvey, Crispin *et al.*, 2008). Diagnostic ions for structural features such as antenna composition, fucose substitution and presence of a bisecting GlcNAc residue are therefore prominent in negative ion mode and readily aid glycan structural assignment.

Other fragmentation methods such as high energy C-trap dissociation (HCD) have been demonstrated as a viable alternative to CID and avoids problems such as neutral loss and fragment rearrangement events during MS/MS (Kenny, A. Hayes *et al.*, 2011; Kenny, Issa *et al.*, 2011), while

electron transfer dissociation (ETD) is commonly used in combination with CID in the analysis of glycopeptides (Mechref, 2012).

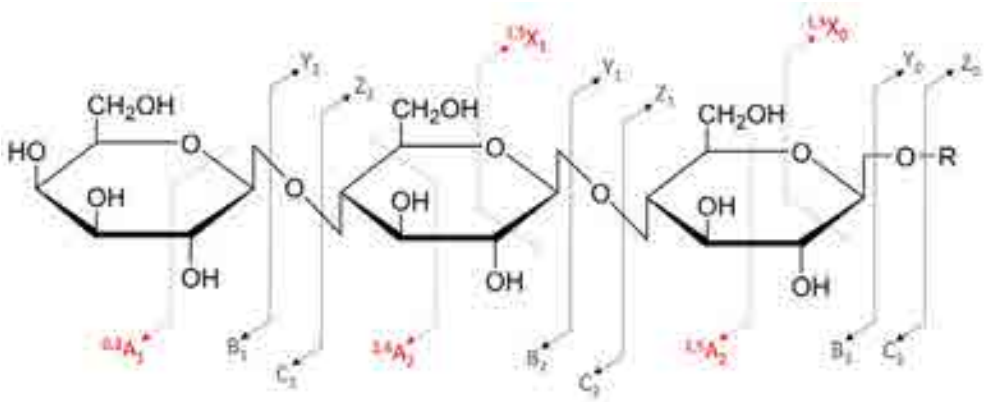


Figure 1.4 Nomenclature for CID of glycans as proposed by Domon and Costello (1988). Cross-ring fragments are marked in red broken lines.

1.4.3.4 Determination of glycan structure from MS data

Structural characterisation of glycans from MS data is still a largely a manual process, although developments have been made in the automation of structural assignment from MS data such as through the use of the GlycoQuest feature within the commercially available ProteinScape software (Chamrad, Blueggel *et al.*, 2007). For manual interpretation of MS data, several free online tools exist such as GlycoMod (Cooper, Gasteiger *et al.*, 2001), GlycoWorkBench (Ceroni, Maass *et al.*, 2008) and UniCarb-DB (Campbell, Nguyen-Khuong *et al.*, 2014). GlycoMod provides basic but useful information about the possible glycan monosaccharide compositions from a given *m/z* value, thus aiding subsequent structural assignment. Once a possible glycan composition is known, the structural details may be assigned through the use of MSn fragmentation data which contains diagnostic ions as mentioned above (and further discussed in Chapter 3). This may be done manually and is a time-consuming and laborious process, or through matching theoretical lists of fragment masses against a list of peaks derived in the MSn spectrum using tools such as GlycoWorkBench. To support the continued development of glycomics-based bioinformatics tools for data analysis, databases such as UniCarb-DB, which contain glycomics-derived LC-MS/MS data on *N*- and *O*-linked glycans released from glycoproteins, provides a public repository necessary to support and assist glycomics research.

Glycan quantitation: For LC-MS/MS analysis of glycans, there are two main methods of sample quantitation. Absolution quantitation compares the measured amount of a glycan of interest to an internal (in-sample spiking) or external (pre-measured using a calibration curve) standard of known amount. On the other hand, relative quantitation refers to the amount of a glycan of interest

expressed as a percentage of total glycans or as a percentage of a reference glycan such as the most abundant glycan in the sample (Moh, Thaysen-Andersen *et al.*, 2015).

The main caveat in MS-based absolute quantitation of glycans is the different ionisation efficiencies of glycans with different structural characteristics as mentioned above. Chromatography conditions in PGC-LC-ESI-MS/MS such as the type of organic solvent used has been shown to influence ionisation efficiency of acidic glycans in negative ion mode (Pabst and Altmann, 2008), potentially further compounding this problem. The use of internal glycan standards with differing structural features for relative quantitation could combat this, whereby glycans are compared only to standards sharing similar structural features. Unfortunately such a set of glycan standards covering all glycan species of interest does not currently exist (Moh, Thaysen-Andersen *et al.*, 2015). Relative quantitation using MS detection can be performed through isolation of specific glycan *m/z* values, extraction of the chromatogram (EIC) for each *m/z,* and performing area-under-curve based quantitation. Comparison of relative intensities of each glycan, or only of glycans with shared characteristics between samples, rather than the absolute quantities in each sample, then combats the issue of differing ionisation properties.

Complementary exoglycosidase treatment: Although LC-MS/MS can provide structural and quantitation information, orthogonal techniques are sometimes required to confirm structural assignments. A parallel approach that utilises these analytical technologies before and after treatment, is the use of exoglycosidase enzymes which cleave specific terminal residues at the non-reducing end of a glycan (Mariño, Bones *et al.*, 2010). These enzymes can have broad or specific linkage specificity, thereby providing or confirming linkage information that is not always readily discernible in MS/MS fragmentation data. The use of exoglycosidases results in HPLC peak shifts which are then compared to the parent profile. For example, a sialylated glycan sample treated with a narrow α2,3 linkage specificity neuraminidase enzyme would result in the loss of peaks containing α2,3 linked sialylated structures and an increase in peaks containing their neutral (desialylated) or partially sialylated α2,6 linked counterparts.

1.5 The plasma membrane proteome

The plasma membrane (PM) is a phospholipid bi-layer comprised of lipids, proteins and carbohydrates. Transport of molecules across the membrane is facilitated by membrane proteins which are classed into three main types: integral proteins that span the bi-layer or are embedded in the inner or outer face of the bi-layer; peripheral proteins that do not span the bi-layer but are associated with either the inner or outer face of the bi-layer; and lipid-anchored proteins (Singer and Nicolson, 1972). In addition to changes in global protein glycosylation, the composition and

abundance of PM proteins can be significantly altered during cell differentiation (Cordwell and Thingholm, 2009) and therefore their identification can aid in further understanding of this process and as discussed above, adds an additional 'omics level to the analysis of protein glycosylation.

1.5.1 Proteomics of human adipogenic stem cell differentiation

Much of the proteomic-based work investigating adipogenesis is in the context of obesity-related disorders such as insulin resistance and type II diabetes (Prior, Larance *et al.*, 2011) and has been carried out using the mouse 3T3-L1 cell line (DeLany, Floyd *et al.*, 2005; Kim, Kim *et al.*, 2015). The use of human ADSCs as a model of *in vitro* adipogenesis has been limited, as have targeted membrane proteomic studies of differentiation in humans. A study of the whole cell protein lysate of adipogenically differentiated ADSCs revealed differential expression of cytoskeletal, metabolic, redox, protein degradation, and heat shock proteins/chaperones (DeLany, Floyd *et al.*, 2005). The researchers concluded that further investigation of heat shock proteins was warranted with respect to the aetiology of obesity and type-2 diabetes. Although human ADSCs are not widely used as a model for adipogenesis, other human MSCs such as bone marrow-derived MSCs have been utilised for this purpose and may help to avoid concerns of differences between human and murine adipogenesis that arise from use of 3T3-L1 cell lines. Bone-marrow derived stem cells have been used in investigations of the total proteome in adipogenesis (Lee, Lee *et al.*, 2006) and also in targeted PM proteomics (Jeong, Ko *et al.*, 2007). The latter study found an increase in proteins involved in lipid metabolism and trafficking after differentiation but focused particularly on ectodomain-containing membrane proteins which include solute carriers and ATPases in an attempt to define new cell surface markers. These protein types were found to decrease after differentiation while several proteins involved in lipid metabolism and trafficking were up-regulated.

1.5.2 Analytical considerations when investigating membrane proteins

PM proteins are typically present in relatively low abundance in the cell proteome and therefore the more abundant cytoplasmic proteins can pose challenges for membrane enrichment from a cell lysate (Sprenger and Jensen, 2010). Membrane proteins are also poorly soluble due to the high hydrophobicity of the trans-membrane domain, a problem that is not encountered with the more soluble cytoplasmic proteins (Gundry, Boheler *et al.*, 2008). This difference in solubility can result in skewing of protein quantitation and/or identification in MS-based proteomics studies (Barrera and Robinson, 2011). Therefore, PM proteins are usually isolated prior to analysis, as described in more detail below.

1.5.2.1 Isolating plasma membranes proteins

PM proteins can be isolated by chemical or physical techniques (Alpert, 1990; Cordwell and Thingholm, 2009; Gundry, Boheler *et al.*, 2008; Sprenger and Jensen, 2010). One example of a chemical method for isolating membrane proteins is aqueous two-phase purification. Here, the membrane and membrane proteins are separated from soluble proteins based on their affinity for two immiscible aqueous phases (Kumar, Galaev *et al.*, 2007; Schindler and Nothwang, 2006). The two phases can be either two different hydrophilic polymers, for example dextran or polyethylene glycol, or a polymer and a low molecular weight solute such as a salt (Kumar, Galaev *et al.*, 2007; Schindler and Nothwang, 2006).

Physical methods to isolate PM proteins typically involve sub-cellular fractionation through ultra-centrifugation or differential centrifugation where organelles are separated based on their different densities (Josic and Clifton, 2007; Sprenger and Jensen, 2010). Sodium carbonate is sometimes used to isolate the plasma membrane and improve solubility (Josic and Clifton, 2007). A recent report of the use of a discontinuous sucrose gradient to fractionate cellular organelles shows promise to selectively avoid inclusion of ER and Golgi membrane proteins in membrane proteomics studies (Lin, Chik *et al.*, 2015).

None of the above methods results in the complete removal of non-PM proteins from a sample (Cordwell and Thingholm, 2009) but downstream solubilisation steps can help to further isolate PM proteins as discussed below. Despite continued efforts to improve methods in PM protein analysis, integral PM proteins particularly are under-represented in proteomics analyses and contamination with intracellular organelle membranes is a problem which is difficult to fully resolve (Josic and Clifton, 2007).

1.5.2.2 Membrane protein solubilisation and enrichment

A suitable diluent must be chosen for the solubilisation of PM proteins after membrane isolation that is able to solubilise PM proteins, disrupt the lipid bi-layer and be compatible with downstream ESI-MS analysis (Lee, Kolarich *et al.*, 2009). Strong non-ionic detergents such as Triton X-114 can be used to separate hydrophobic membrane proteins from the more hydrophilic cytoplasmic protein carryover from the membrane isolation procedure (Bordier, 1981). The separation mechanism of Triton X-114 is based on a temperature sensitive aqueous-detergent phase partitioning in which the hydrophobic proteins are contained in the detergent phase (Cordwell and Thingholm, 2009). Sample cleanup prior to MS analysis is most commonly achieved using acetone or TCA precipitation (Josic and Clifton, 2007). Membrane glycoproteins are typically hydrophobic and their isolation using this method has been successfully applied previously (Lee, Kolarich *et al.*, 2009; Nakano, Saldanha *et al.*,

2011). Subsequent quantitation of the amount of isolated membrane protein can be difficult however, due to problems with compatibility of the detergents with reagents required for Bradford (Bradford, 1976) and BCA protein assays (Walker, 1996). These assays are sensitive to high levels of Triton X-114 and SDS, along with commonly used chaotropes such as urea and thiourea, and concentrations of these reagents must therefore be checked for compatibility with protein assays. Membrane proteins solubilised in chaotropes such as urea are compatible for the PVDF-based release of *N*- and *O*-linked glycans from membrane glycoproteins described above.

1.5.3 Mass spectrometry to study membrane proteins

The most commonly used approach in proteomics investigations is the 'shotgun' proteomic method whereby proteins proteolytically digested into peptides are separated prior to detection with MS-MS (Yates, Ruse *et al.*, 2009). This 'bottom up' approach can involve the use of various protease enzymes such as trypsin, chymotrypsin, Lys-C and Arg-C. 'Top down' approaches, where masses of intact proteins and their fragments are used for protein identification, are also employed, typically with the use of alternative fragmentation methods such as ETD and electron capture dissociation (ECD) which provide more complete backbone sequencing. The hallmark of this approach is the ability to produce 100 % coverage of a protein sequence at both the intact and fragment ion levels (Kelleher, 2004). The use of two-dimensional (2D) sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) to fractionate complex samples is also used in top down proteomics and can provide high resolution of protein isoforms, allow greater proteome coverage and is also compatible with ensuring solubility of hydrophobic membrane proteins (Coorssen, 2013).

In bottom up proteomics, proteins can also be fractionated to decrease sample complexity using one-dimensional (1D) SDS-PAGE and is performed prior to protease digestion. In 1D SDS-PAGE 'slice and dice' experiments, gel bands are usually 'sliced' into equal portions, or particular bands of interest are selected, before each piece is further 'diced' into small pieces before proteolytic digestion. This is an efficient way to ensure solubilisation of hydrophobic membrane proteins and detergent cleanup before MS-MS (Neilson, Keighley *et al.*, 2013). Prior to MS analysis, LC is the most commonly used method for peptide sample separation, with HPLC instruments usually coupled directly to mass spectrometers with an ESI source (Yates, Ruse *et al.*, 2009). RP chromatography is the most commonly used type of chromatography for separation of digested peptides prior to MS (Gundry, White *et al.*, 2009; Josic and Clifton, 2007; Robinson, Artpradit *et al.*, 2012) with the added advantage of the ability of chromatography to concentrate and desalt the sample using MS-compatible solvents.

1.5.3.1 Quantitative proteomics

Although MS is not an inherently quantitative method, the field of MS-based proteomics has moved beyond simple protein identification with various means of quantifying changes in protein abundance. There are two main approaches used in quantitative proteomics: the use of stable isotope labelling or label-free techniques. Although labelling strategies require expensive isotope labels and specific software, they are often considered to be more accurate for protein quantitation (Neilson, Ali *et al.*, 2011). Common labelling techniques involve modifying peptides with isobaric tags for relative or absolute quantitation (iTRAQ) (Ross, Huang *et al.*, 2004), labelling proteins with isotope-coded affinity tags (Gygi, Rist *et al.*, 1999), or labelling proteins by incorporation of stable isotope labels with amino acids in cell culture (SILAC) (Ong, Blagoev *et al.*, 2002). The addition of a metabolic label early in a cell's growth cycle can offer a highly accurate identification of the proteome and avoid anomalies from subsequent sample handling (Bantscheff, Lemeer *et al.*, 2012). Disadvantages of these collective approaches however are the limitations in the number of samples that can be processed in an experiment and problems of compatibility with certain sample types.

Label-free quantitation approaches are not affected by these limitations and have relatively straight forward data-analysis requirements without needing specific software (Neilson, George *et al.*, 2014). Quantitation techniques for label-free proteomics studies fall into two main groups: 1) area under the curve (AUC) or signal intensity measurement based on the precursor ion spectra; and 2) spectral counting, which is based on counting the number of peptides assigned to a protein in an MS/MS experiment (Neilson, Ali *et al.*, 2011). AUC quantitation involves the integrated measurement of chromatographic peaks for a given peptide and is based on the principle that the area under a chromatographic elution curve is proportional to the amount of the eluting compound. This method however relies on accurate and reproducible chromatography for proper alignment across different chromatographic profiles, and loses accuracy as peptide mixture complexity increases. Spectral counting relies on the principle that as more of a digested protein is analysed in a MS, more peptides belonging to the protein will be identified (Neilson, Keighley *et al.*, 2013). Therefore, the number of resultant spectra assigned to a protein can be used as a measure of the relative protein amount within the sample. This approach has been modified to accommodate for the length of a protein affecting the number of spectral counts (SpC), i.e. longer proteins will produce more peptides. The use of a normalised spectral abundance factor (NSAF) provides an improved measure of protein abundance. This involves dividing the SpC of a protein by its length (L) and then normalised by dividing by the sum of all SpC/L for all proteins in a sample (Zybailov, Mosley *et al.*, 2006). It is important to note that NSAF-based quantitation can be influenced by experimental parameters of

the MS, where settings such as dynamic exclusion of precursor ions and instrument scan speed and affect a peptide's detection (Bantscheff, Lemeer *et al.*, 2012).

The proteomic analysis in this thesis was carried out using a label-free quantitative shotgun approach with SDS-PAGE RP-LC ESI-MS/MS and quantitation carried out using NSAFs.

1.5.3.2 Data analysis in label-free proteomics

Large volumes of data are produced in a label-free experiment and require processing for protein identification and quantitation, statistical assessment and determination of biological significance. Protein identification is performed by matching spectra to predicted, *in silico*-generated peptide fragments in a database through the use of scoring algorithms such as SEQUEST (Eng, McCormack *et al.*, 1994), MASCOT (Pevzner, Mulyukov *et al.*, 2001) and X!Tandem (Craig and Beavis, 2004). The score given to an identified peptide reflects the probability of how well the experimental data matches the theoretical data contained in the database and therefore the likelihood of a true match. False positive identifications can be minimised by searching spectra against a reversed or randomised database where any matches would be presumed false (Johnson, Davis *et al.*, 2005). This can also allow the calculation of a false discovery rate at both the peptide and protein level. Furthermore, protein identifications can be filtered by setting of a scoring threshold to minimise false positives and maximise true identifications (Neilson, Ali *et al.*, 2011). The search engine used for protein identification in this thesis was the Global Proteome Machine (Craig, Cortens *et al.*, 2004) software which uses a freely available version of the X!Tandem algorithm, GPM-XE Tornado, and relies on probability-based scoring like that used in MASCOT.

An important issue in bottom up proteomics regardless of whether a labelling or label-free method has been employed is that the same peptide sequences can occur in multiple proteins; this is termed the protein inference problem. Such peptides can be matched in a database to one or more proteins during protein identification and currently there are three methods used to deal with this: allow shared peptides to be counted multiple times, remove shared peptides and focus only on unique peptides (containing an exclusive peptide sequence not found in multiple proteins) or distribute shared peptides across the homologous proteins (Neilson, Ali *et al.*, 2011). The X!Tandem algorithm assigns equal counts to shared peptides (Gammulla, Pascovici *et al.*, 2011). It is important to note that the protein inference problem is also dependent on the database used for peptide matching and the date of its use.

1.5.3.3 Forming biological conclusions from proteomics experiments

A challenge in the generation of large proteomics datasets is the extraction of relevant biological information. Gene ontology (GO), interaction networks and mapping of biological pathways are

commonly incorporated into proteomics workflows (Malik, Dulla *et al.*, 2010). GO is a controlled ontology designed to ensure the standardised use of terms describing protein function with each term falling under one of three categories: biological process, molecular function or cellular component (Ashburner, Ball *et al.*, 2000). Several GO analysis tools exist such as AmiGO (Carbon, Ireland *et al.*, 2009), GOMiner (Zeeberg, Feng *et al.*, 2003), WEGO (Ye, Fang *et al.*, 2006) and DAVID (Dennis, Sherman *et al.*, 2003). DAVID offers the function to cluster similar GO terms making it the most useful tool for GO annotation and helps detect over- or under-represented GO terms in a dataset in addition to a statistical assessment as to whether this representation is significant with respect to a background dataset (the default is the human proteome) (Malik, Dulla *et al.*, 2010).

For pathway analysis, currently there are around 300 pathway resources available, the most comprehensive of which are KEGG (Kanehisa, Araki *et al.*, 2008) and Reactome (Joshi-Tope, Gillespie *et al.*, 2005) which are the biggest pathway databases. One concern with pathway analysis is that it is not always ontology-driven, causing ambiguity and inconsistency between databases (Malik, Dulla *et al.*, 2010). For information about protein-protein interaction, interaction network tools such as the Agile Protein Interaction DataAnalyzer (Prieto and De Las Rivas, 2006), STRING 8 (Jensen, Kuhn *et al.*, 2009) and Cytoscape (Shannon, Markiel *et al.*, 2003) can provide useful information about protein activities in a complex as most proteins do not operate alone. These tools however rely on the input of relatively large protein lists in order to obtain significant or meaningful biological insights.

1.6 Thesis aims

The therapeutic potential of human ADSCs in tissue engineering and regeneration of adipose tissue has been recognised, and as ADSCs increasingly move into the clinical setting, a comprehensive knowledge of ADSC biology is required.

Much of the current knowledge of adipogenesis has been based on the murine cell line 3T3-L1 as a model (Wang, Scherer *et al.*, 2014), however 3T3-L1 is an immortal cell line and murine in origin which raises concerns about differences in adipogenesis between humans and mice (DeLany, Floyd *et al.*, 2005). Here, we use human ADSCs, a relatively easily obtainable source of human MSCs, as a suitable alternative with which to study the adipogenic differentiation process *in vitro,* and in comparison with the mature adipocytes formed *in vivo* from ADSCs in the same person (Figure 1.5). The highly dissimilar *in vivo* and *in vitro* adipose differentiation environments will be compared in terms of the phenotypic differences in membrane protein glycosylation of the adipogenic cells produced.

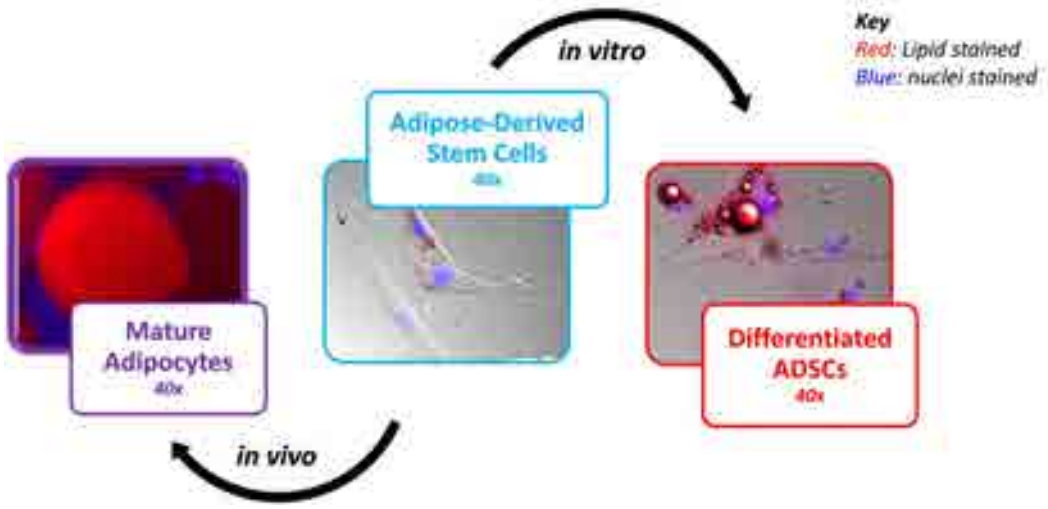


Figure 1.5 Model for the study of adipose-derived stem cell adipogenesis. Mature adipocytes and ADSCs from the same lipo-aspirate provide a model of *in vivo* adipogenesis while ADSCs and *in vitro* differentiated ADSCs provide a model of *in vitro* adipogenesis

Specific Aims:

The overall aim of the work presented in this thesis was to investigate changes in *in vitro* compared to *in vivo* adipogenesis in the context of protein glycosylation, using the model described above. More specific aims were:

1. To establish a large-scale method for adipogenic differentiation of ADSCs *in vitro* for membrane glycomic and proteomic analysis (Chapter 2).
2. To improve knowledge and understanding of the role of membrane glycoproteins in ADSC differentiation by the characterisation of glycans (Chapter 3) and membrane proteins (Chapter 4) from un-differentiated (ADSC) and differentiating (dADSC) stem cells compared to the fully differentiated native mature adipocytes.
3. To develop a protein-specific integrated glycomics and proteomics approach to further investigate specific glycoprotein changes that occur during ADSC differentiation (Chapter 5).

Chapter 2:

Evaluating adipose-derived stem cell differentiation *in vitro*

2.1 Introduction

Cell differentiation involves large morphological changes through cytoskeletal remodelling in the formation of the product or terminally differentiated cell. These changes can be visualised through the use of various microscopy techniques which form complementary biomolecular approaches to investigating biological questions and processes. In this opening study, the *in vitro* adipogenic differentiation of adipose-derived stem cells (ADSCs) was optimised and evaluated using fluorescence microscopy and measurement of adipocyte-specific gene marker expression. These *in vitro* differentiated cells were compared with *in vivo*-derived adipocytes.

The adipogenic differentiation of ADSCs *in vitro* involves the introduction of a cocktail of different inducers that are applied to the cells once they have grown to approximately 90 % confluency. The adipogenic differentiation protocol used in this thesis was based on a method employed by a pioneer in the field of ADSC biology (Zuk, Zhu *et al.*, 2001) and contains four of widely used adipogenic inducers: indomethacin, insulin, IBMX and dexamethasone. Insulin is widely used to promote adipogenic differentiation *in vitro*, while indomethacin, and IBMX in conjunction with dexamethasone are agonists of peroxisome proliferator-activated receptor gamma (PPAR γ), the key mediator of adipogenesis (Scott, Nguyen *et al.*, 2011).

The adipogenic differentiation of ADSCs *in vivo* occurs in two stages: commitment to the adipogenic lineage in a process called determination, followed by terminal differentiation where the cells take on the unique unilocular morphology of a mature adipocyte (Rosen and MacDougald, 2006). Initially, small lipid-filled vesicles form in the cell, gradually swelling with lipid before eventually fusing to result in a single large lipid vesicle in the mature adipocyte (Moreno-Navarrete and Fernández-Real, 2012). This causes the adipocyte nucleus to be pushed to the cell's periphery, resulting in the characteristic morphology of native mature adipocytes. ADSCs and pre-adipocytes themselves resemble fibroblasts in shape (Moreno-Navarrete and Fernández-Real, 2012) and a morphological comparison of ADSCs and adipocytes can be conducted using microscopy.

Imaging whole adipose tissue faces many challenges as it is a very dense material and sections must be cut very finely to avoid light scattering interfering with image detail, which is typical in thick samples (Shen and Chen, 2008). The lipid-filled nature of the cells also has implications for how the light is refracted in the sample, with a high refractive index often resulting in a kaleidoscope-like effect and obscuring structural detail. This can be avoided by selecting thin areas of tissue to image, and also by optical sectioning using a confocal microscope (St Croix, Shand *et al.*, 2005). This technique provides valuable information about the different layers of a cell or whole tissue, and can also remove out-of-focus areas by imaging in a single plane of focus.

The imaging of adipocytes is also a challenging task, particularly in comparison with *in vitro* cultured cells adhered to culture vessels. Adipocytes are the major component of adipose tissue, a loose connective tissue with a heterogeneous mix of cell types. Mature adipocytes are fragile cells which make it difficult to immobilise them on a surface for imaging, particularly in the use of inverted microscopes where a cover slip must be applied and the slide subsequently mounted face-down onto the microscope stage. Since they are composed of high lipid content, adipocytes also typically float in cell culture medium making them difficult to handle.

The use of different stains helps to visualise cellular characteristics. Commonly used stains for the lipids contained in adipocytes include Nile Red, BODIPY and Oil Red O (Mehlem, Hagberg *et al.*, 2013; Nishimura, Manabe *et al.*, 2008; Zuk, Zhu *et al.*, 2001). When stained in conjunction with a nucleic acid dye, the characteristic unilocular adipocyte with its nucleus pushed to the cell periphery can be identified. In adipogenic differentiation of ADSCs *in vitro*, while there are several ways to assess degree of differentiation, staining is the most direct (Bunnell, Estes *et al.*, 2008). For areas of sample that are unstained, the additional use of optical methods to introduce contrast can help visualisation. Two such methods are phase contrast or differential interference contrast (DIC).

In addition to staining and microscopy, to further verify and compare morphological changes occurring in stem cell differentiation, expression levels of lineage-specific marker genes can be measured using quantitative real-time polymerase chain reaction (qRT-PCR) (Zuk, Zhu *et al.*, 2001). Common adipogenic lineage genes that are up-regulated during *in vitro* adipogenic differentiation include fatty acid binding protein (FABP4 or aP2), lipoprotein lipase (LPL), glutathione synthetase (GLUTA), leptin, PPAR γ1 and PPAR γ2, perilipin, ADRP and cyclophilin (Bunnell, Estes *et al.*, 2008; Jefcoate, Wang *et al.*, 2008). Conversely, the down-regulation of mesenchymal stem cell-specific markers during differentiation such as CD90, CD73 and CD105 (Blaber, Webster *et al.*, 2012; De Gemmis, Lapucci *et al.*, 2006) can also be used to monitor differentiation. When performing qRT-PCR, it is important to choose a stable housekeeping gene to normalise expression levels of genes of interest. However, many cytoskeletal changes occur during differentiation, with the commonly used housekeeping gene for the product β-actin, exhibiting varying expression levels during cell differentiation (Fink, Lund *et al.*, 2008) and thus not able to be used. Recent studies have reported genes coding for TATA-binding protein (TBP), tyrosine 3/tryptophan 5-monooxygenase activation protein (YWHAZ) and β-glucuronidase (GUSB) to be the most stable housekeeping genes during MSC differentiation (Fink, Lund *et al.*, 2008; Ragni, Vigano *et al.*, 2012).

Previous work has demonstrated changes in *in vitro* adipogenesis that mimic morphological characteristics of mature adipocytes and display elevated levels of adipogenic marker genes

(Hamouda, Ullah *et al.*, 2013), therefore this chapter investigates whether these morphological and genetic differences are reflected in the samples obtained for this thesis. The overall objective of the work presented in this chapter was to develop an *in vitro* method for the generation of sufficient quantities of differentiated ADSCs of an adipogenic lineage for subsequent use in glycomic and proteomic analyses (Chapters 3, 4 and 5). To assess the differentiation process and compare to *in vivo* adipogenesis, various microscopy techniques were used to image ADSCs along with their adipogenic progeny, and phenotypic characteristics were compared to those of native adipocytes from the same individual. Changes in gene expression, particularly of known adipocyte marker genes, were monitored by qRT-PCR.

2.2 Materials and Methods

2.2.1 Materials

Collagenase (from *Clostridium histolyticum*), phosphate buffered saline (PBS) tablets, isobutyl-methylxanthine (IBMX), dexamethasone, paraformaldehyde, chloroform, Oil Red O and γ-irradiated FBS was purchased from Sigma (St Louis, MO, USA). Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 with GlutaMAX™ (DMEM-F12), Dulbecco's Modified Eagle Medium-High Glucose (DMEM-HG), Penicillin-Streptomycin solution, Trypsin EDTA 0.25 % and insulin (human recombinant, zinc) were obtained from Invitrogen (Life Technologies, Australia). Hoechst 33342 was from Life Technologies (Australia). Fetal bovine serum (FBS) was purchased from Bovogen (Australia) and saline solution from Baxter Bioscience (Australia). cOmplete Protease Inhibitor Cocktail tablets were purchased from Roche Diagnostics (Basel, Switzerland). Uncoated, γ-irradiated 35 mm^2 glass bottom Microwell dishes with a 14 mm diameter microwell with No. 1.0 thickness glass, were purchased from MatTek Corporation (MA, USA). HPLC-grade water was obtained using a Milli-Q® Synthesis water purification system from Millipore. 175cm^2 and 75cm^2 flasks were purchased from Corning (NSW, Australia). RNase-Free DNase Set, RNeasy Mini Kit and QuantiTect Reverse Transcription Kit were purchased from Qiagen (Victoria, Australia). GoTaq qPCR Master Mix was purchased from Promega (Sydney, Australia), and qRT-PCR primers were purchased from IDT Technologies (Australia). Trizol and nuclease-free water was purchased from Ambion (Life Technologies, Australia). Ethylene imine polymer solution (PEI) was from Fluka Analytical (Sigma, USA), and liquid blocker super PAP pen and Fluoro-Gel were from ProSciTech (Australia).

2.2.2 Lipo-aspirate processing, and adipocyte and ADSC isolation

Human lipo-aspirates from various abdominal and torso regions were obtained from patients undergoing elective liposuction procedures for cosmetic purposes at various cosmetic surgery clinics

in Sydney, NSW. Written consent was obtained from the patients who participated in this study. The use of lipo-aspirate samples for this work was approved by Macquarie University's Human Research Ethics Committee (Ref: 5201100385) and the ethics approval can be found in Appendix A. Processing of lipo-aspirate material was conducted using a modified procedure previously reported (Blaber, Webster *et al.*, 2012). Approximately 600 ml of lipo-aspirate from each sample was washed in pre-warmed saline to remove blood cells and tumescent fluid remaining from the liposuction procedure. Samples were then treated with 0.5 mg/ml collagenase (from *Clostridium histolyticum*) in a 37 °C water bath for one hour with periodic mixing every 15 minutes, followed by centrifugation at 1000 *g* for ten minutes. The resulting floating adipocytes and pelleted stromal vascular fraction (SVF) cells were washed separately with saline and centrifuged again. Washed adipocytes were flash frozen using liquid nitrogen before storing at -80 °C for later analysis.

Table 2.1 Cell culture media components and amount used in ADSC maintenance culture and adipogenic differentiation in different culture vessels

Media Type	Used For	Components	Usage (ml/cm^2)
Standard Media	Expansion and culture of ADSCs in T175 cm^2 flasks at passage zero, before transfer to 75 cm^2 flasks for passages one to three	DMEM/F-12 10 % FBS 1 % Penicillin-Streptomycin	0.2
Adipogenic Differentiation Media (modified from Zuk, Zhu *et al.*, 2001)	30-day differentiation of ADSCs in a) 35mm^2 wells and b) 75cm^2 flasks	DMEM-HG 10 % FBS 1 % Penicillin-Streptomycin 0.5 mM IBMX 1 µM Dexamethasone 10 µM Insulin 200 µM Indomethacin	a) 0.8 b) 0.26

The stem cell-containing SVF pellet was resuspended in Standard Media (culture medium components outlined in Table 2.1) The number of nucleated cells harvested were counted using a FacsScan flow cytometer (Becton Dickinson, USA) using the protocol outlined by Blaber, Webster *et al.* (2012). Based on the counts of viable cells, the volume of SVF mixture corresponding to two million viable nucleated cells was seeded into separate 175 cm^2 culture flasks as passage zero. The volume of sample in each flask was then topped up to 30 ml with Standard Media. ADSCs were left to adhere to the flask and proliferate for two to three days before the first media change, which removed non-adherent cells including red blood cells. Confirming the remaining adherent cells as mesenchymal stem cells using the described SVF harvest method has been previously carried out by CD Marker characterisation using flow cytometry (Blaber, Webster et al., 2012). Testing of multiple mesenchymal lineage differentiation ability was also undertaken to confirm plasticity (Blaber, Webster et al., 2012). ADSCs were grown to approximately 80 % confluency before passaging using

Trypsin EDTA 0.25 % and grown in 75 cm^2 flasks for subsequent passages. Media changes were performed every three days and cells passaged at approximately 80 % confluency until passage three where ADSCs were allowed to grow to 90 % confluency before either:

a) Collection for differentiation in 35 mm^2 wells (Section 2.2.5)
b) Induction of adipogenic differentiation in the same 75 cm^2 flasks (Section 2.2.6)
c) Collection for later RNA isolation and qRT-PCR (Section 2.2.9).

The work undertaken in this chapter is illustrated in Figure 2.1.

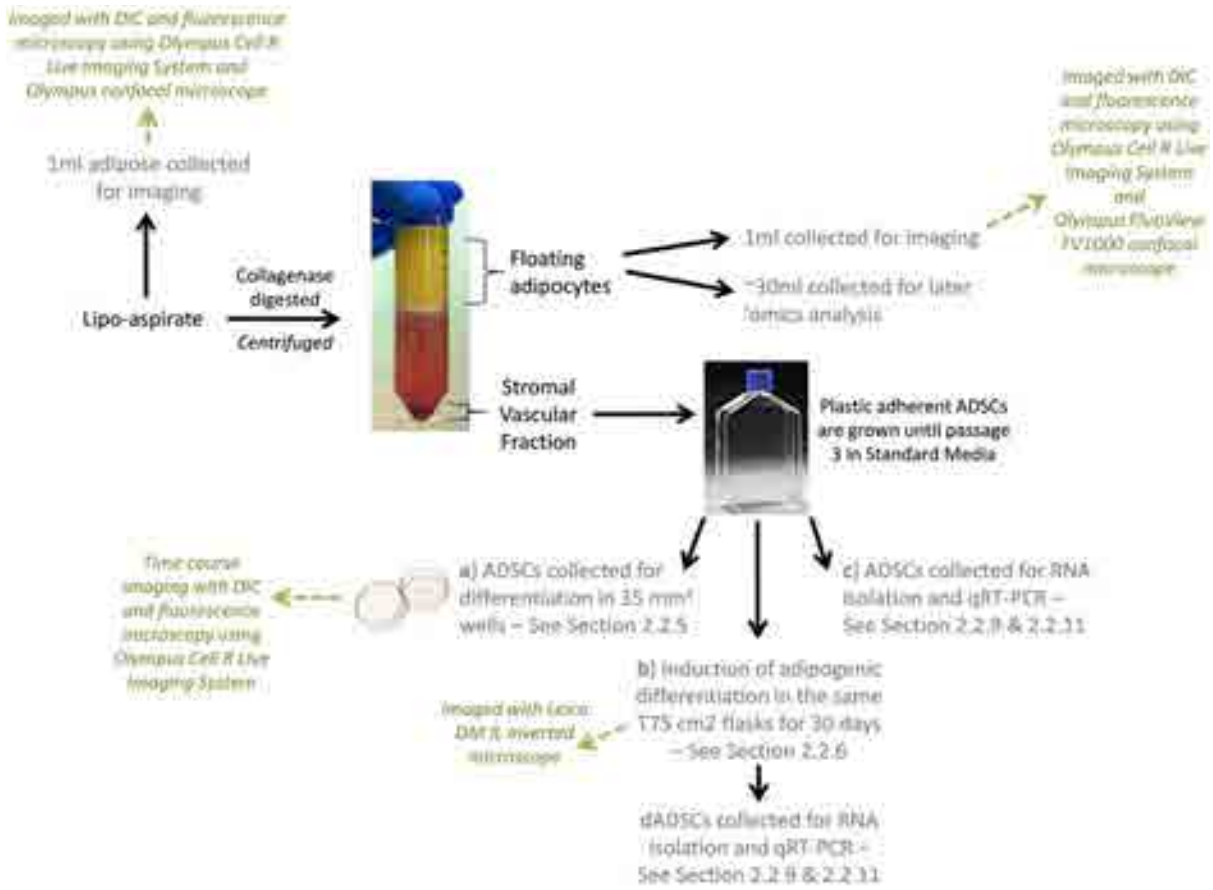


Figure 2.1 Flowchart of methods undertaken in this chapter with corresponding methods section/s indicated. (35 mm^2 microwell image from http://glass-bottom-dishes.com/pages/)

2.2.3 Sourcing media component FBS for the culture of ADSCs and dADSCs

Over the course of this thesis, ADSCs were cultured using DMEM-F12 media supplemented with γ-irradiated FBS from Sigma. Using phase contrast microscopy to routinely check the growth of ADSCs, it was discovered that cells cultured with γ-irradiated FBS lost their typical fibroblastic phenotype over approximately the first two cell passages, with stress fibres becoming evident (marked with arrows; Figure 2.2b) and beginning to form clusters despite having been a homogenous single-cell layer (Figure 2.2a). When the cells reached passage two, to discount the possibility that γ-radiation

had decreased the concentration of active serum proteins through denaturation (as demonstrated by Martel, Wu *et al.*, 2010), the percentage of serum was increased to 20 % to try to account for lower levels of active proteins. This however resulted in the cells becoming more senescent in appearance (Figure 2.2c). Concurrently, the γ-irradiated FBS was removed from other flasks and replaced with non γ-irradiated FBS. After overnight incubation the cells appeared to be more fibroblastic (Figure 2.2d), however after two subsequent passages (Figure 2.2e-f) appeared to revert to the senescent appearance of passage two (Figure 2.2b), indicating that they could not be recovered for further growth. This medium was replaced by non-irradiated FBS obtained from Bovogen (Australia) for the experiments described in this thesis.

In summary, it should be noted that γ-irradiated FBS supplemented culture medium had disastrous effects on the ability to grow healthy ADSCs. γ-irradiated FBS is used for ensuring possible contaminating bacteria are inactivated, but its successful use in cell culture seems to be dependent on the cell line, with some cell lines exhibiting higher growth rate in irradiated FBS than in the non-irradiated serum control (Nims, Gauvin *et al.*, 2011). Zuk *et al*. (2001) tested ten different serum types from three manufacturers in the cell culture of ADSCs and reported that a specific sera did not seem to be required, as there were no adverse effects on morphology, growth rate or differentiation capacity. However, the work presented here (Figure 2.2) indicates that the growth and morphology of human ADSCs is severely affected by the used of γ-irradiated FBS and special consideration should be undertaken when choosing serum for ADSC cell culture

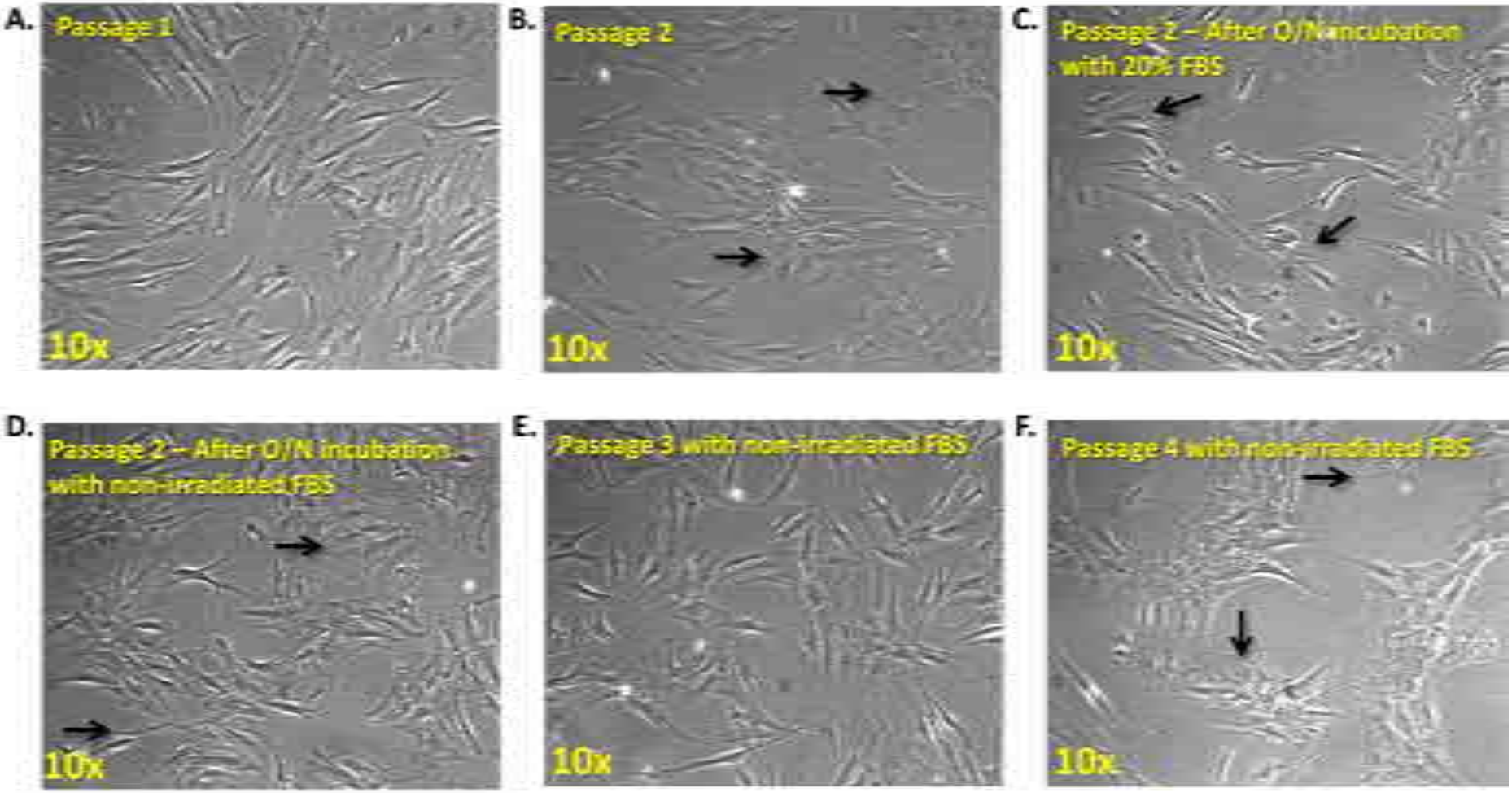


Figure 2.2 ADSCs cultured in DMEM/F12 media supplemented with γ-irradiated FBS (**A & B**) begin to lose their fibroblastic shape and cluster together. Increasing the FBS concentration to 20 % (**C**) resulted in an increase of senescent phenotypes, while replacement with non-γ-irradiated FBS initially indicated ADSCs could be recovered (**D & E**). After two subsequent passages, these cells no longer resembled a homogenous layer of fibroblastic cells (**F**). Formation of stress fibres is marked by arrows.

2.2.4 Staining and imaging adipocytes and adipose tissue with Oil Red O and Hoechst

Fixation: To stain adipocytes from collagenase-digested adipose tissue, 1 ml adipocytes were first transferred to an eppendorf tube and fixed in 8 % (v/v) paraformaldehyde. For undigested adipose tissue, approximately 1 ml of adipose tissue from the original lipo-aspirate was transferred into an eppendorf tube using a 20 ml serological pipette and fixed in 8 % (v/v) paraformaldehyde. As the adipose tissue is broken up during the lipo-suction procedure, it is not a solid tissue mass. Rather it is comprised of small clumps of tissue which can be easily collected but require a large pipette tip opening for collection, such as the 20 ml serological pipette.

Preparation of slides: Microscope slides were treated with 0.1 % (v/v) PEI solution to help cell adherence. Wells to contain cells and tissue were drawn using a liquid blocker super PAP pen. A 20 µl drop of adipocytes in solution was applied to a treated microscope slide. For adipose tissue, a small amount of material was scraped out of the eppendorf tube using a wooden toothpick and also applied to a microscope slide.

Staining: To decrease sample perturbation, filter paper was used to gently draw off liquid in the following staining steps for both adipocytes and adipose tissue. Oil Red O is a hydrophobic stain therefore after samples were rinsed with water, isopropanol was added for five minutes to dehydrate the cells. To prepare a working solution, Oil Red O (0.3 % w/v in isopropanol) was diluted with water 3:2 and passed through a 0.45 µm filter. It was then applied to adipocytes and adipose tissue for no longer than 30 min to avoid dye precipitation (Mehlem, Hagberg *et al.*, 2013). Stain was removed and samples washed with PBS until no longer pink. For fluorescence microscopy, nuclei were stained with 2 µg/ml Hoechst 33342 in PBS for 15 min. Samples were then washed once with PBS and then mounted in Fluoro-Gel. A cover slip was placed on cells and tissue, and the edges sealed using clear nail polish. This pushed thinner sample areas to the sample periphery which were used for imaging.

2.2.5 Adipogenic differentiation of ADSCs in 35 mm^2 wells for time course imaging

The adipogenic differentiation protocol based on that reported by Zuk *et al.* (Zuk, Zhu *et al.*, 2001) was performed on ADSCs in 35 mm^2 glass bottomed Microwell dishes. Glass is a transparent and ideal surface for microscopy imaging of cells using an inverted microscope and DIC imaging. ADSCs cultured in 75 cm^2 flasks were grown until approximately 90 % confluency in the third passage before treatment with Trypsin EDTA 0.25 % (w/v) to collect the cells. Cells were counted using a haemocytometer before 100, 000 ADSCs were seeded onto separate 12 glass bottomed dishes and Standard Media applied. Eight dishes (Dishes 1-8; Figure 2.3) were allocated for time course imaging of adipogenic differentiation. A schematic of the experimental layout is shown in Figure 2.3. The cells

in these dishes were allowed to reach at least 90 % confluency before the adipogenic differentiation medium was added. A high confluency rate is needed before inducing differentiation as ADSCs do not continue to proliferate after induction, and some ADSCs die upon the addition of the differentiation cocktail (S. Blaber; personal communication). A high confluency has also been reported to increase overall efficiency of adipogenic differentiation (Bunnell, Flaat *et al.*, 2008). The cocktail components are outlined in Table 2.1 and differ from the method reported by Zuk *et al*. (Zuk, Zhu *et al.*, 2001) by the use of DMEM-HG and penicillin/streptomycin instead of plain DMEM and antibiotic/anti-mycotic as previous work from our lab demonstrated a higher degree of differentiation is achieved when DMEM-HG is used. The differentiation medium was added at a volume of 0.8 ml/cm^2 and carefully changed twice a week. Before a media change, a dish was taken for staining and imaging, further described in Section 2.2.7. For example, Dish 2 was stained and imaged eight days into the 30-day differentiation period and had undergone a media change on the fifth day when Dish 1 was imaged. The ADSCs were imaged before the differentiation cocktail was added to Dishes 1-8 to have a 'before' comparison of ADSCs at 90 % confluency. ADSCs in the remaining four dishes (Dishes 9-12; Figure 2.3) were kept and maintained in Standard Media.

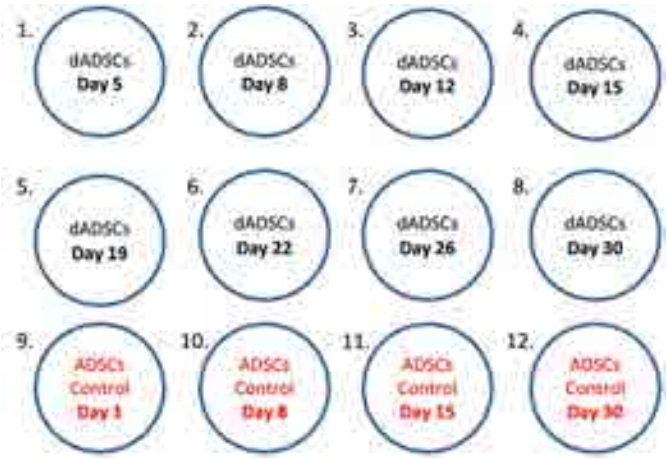


Figure 2.3 Allocation of dishes for time course imaging of *in vitro* adipogenic differentiation of ADSCs. Dishes one to eight were used to observe the 30-day differentiation of ADSCs and one dish was taken for imaging every three to four days. Dishes 9 to 12 contained ADSCs which were used as undifferentiated controls

2.2.6 Adipogenic differentiation in 75 cm^2 flasks

In order to have enough material for glycomics and proteomics analysis, ADSCs were differentiated in 75 cm^2 flasks. Once cells reached approximately 90 % confluency at passage three, Standard Media was replaced with Adipogenic Differentiation media (Table 2.1) in half the flasks for 30 days. ADSCs in the other flasks were collected as described in Section 2.2.8.

The differentiation media was added at a volume of 0.26 ml/cm^2 to mirror as much as possible the generous volume used in the 35 mm^2 dishes. In the first two weeks, media was changed twice a week with old media removed and fresh media added. After this time the differentiated cells began to lift off the base of the flask due to lipid accumulation. To minimise the loss of these more differentiated cells through media changes, old media was removed and discarded for the last time at the beginning of the third week. Fresh media was added and the flasks topped up with more fresh media at the beginning of the fourth week. Therefore, no media was discarded in the last two weeks and at the end of the 30 days all media was collected. Generous volumes of media were used throughout the differentiated period, particularly to avoid cells being exposed to high concentrations of accumulated cell metabolic waste when old media was not being removed.

2.2.7 Staining dADSCs with Oil Red O and Hoechst

To confirm adipogenic differentiation, dADSCs were stained with Oil Red O for lipid. Media was carefully removed from wells and the culture vessel rinsed with PBS. Extra care was taken in these steps to avoid dislodging any loosely bound dADSCs. Cells were fixed by addition of 4 % paraformaldehyde to the wells for 30 min at room temperature and washed with PBS prior to staining. Oil Red O staining was performed in the wells as described in Section 2.2.4, however the stain was applied only for 5 min as dADSCs are less dense than adipocytes and adipose tissue. Nuclei were stained with 2 µg/ml Hoechst 33342 in PBS for 15 min then washed once with PBS before imaging.

2.2.8 Microscopy

For adipocytes, adipose tissue and dADSCs from 35mm^2 dishes, DIC and wide-field epi-fluorescence of Oil Red O and Hoechst staining was imaged on a Cell R Live cell imaging platform equipped with an IX 71 inverted microscope (Olympus, Japan). 20X Plan Super Apochromat N.A 0.75 and 40X U Plan Super Apochromat N.A 0.95 objectives were used, along with excitation filters BP377/50, BP527/27 (for illumination system, Xenon lamp) and a DAPI/FITC/TxREd triple band dichroic mirror and emission filter. Imaging and image processing was carried out using xCellence rt software.

Also for adipocytes and adipose tissue, confocal imaging was conducted using a FluoView FV1000 confocal microscope equipped with an IX 81 inverted microscope (Olympus, Japan). Transmitted light and confocal imaging was performed using 20X UPLSAPO N.A 0.75 and 40X UPLASAPO N.A 0.9 objectives. Semiconductor lasers LD405 and LD559 were used for excitation of Hoechst 3342 and Oil Red O respectively and a spectral fluorescent detector for collection of fluorescent emission signals. The FluoView imaging and image processing software was used.

Phase contrast microscopy of dADSCs from 75 cm^2 flasks was carried out using a Leica DM IL inverted microscope.

2.2.9 Collecting ADSCs and dADSCs for total RNA extraction

Once ADSCs reached approximately 90 % confluency at the third passage, the cells from half the flasks were harvested for later analysis using a cell scraper. Media was removed from the flasks and the cells gently washed three times with cold PBS to remove residual media and FBS proteins. Cells were then collected in a 2 ml solution of a protease inhibitor in PBS (one cOmplete Protease Inhibitor Cocktail tablet solubilised in 50 ml PBS). ADSCs were pelleted and flash frozen with liquid nitrogen before storing at -80 °C for total RNA extraction.

At the end of the 30-day differentiation, dADSCs in 75 cm^2 flasks that were not used for staining were collected. First, the media was collected into Falcon tubes to collect any floating dADSCs cells. The differentiated cells formed a suspension throughout the media solution, making them difficult to collect, therefore approximately 20 ml cOmplete Protease Inhibitor in PBS was added to each tube to decrease the density of the solution. Tubes were then spun at 14, 000 g for 10 minutes, causing some of the cells to clump together and enabling some of the surrounding liquid to be removed with a pipette. The addition of cOmplete Protease Inhibitor Cocktail in PBS and centrifugation was repeated another 1-2 times for each tube and eventually allowed for the differentiated cells to form a pellet or be collected using a pipette. These cells were added to those collected from the flasks which were scraped and collected as described for the ADSCs above. dADSC pellets were also flash frozen with liquid nitrogen before storing at -80 °C.

2.2.10 Total RNA extraction from ADSCs, dADSCs and adipocytes

Total RNA extraction and analysis was conducted in biological triplicate. For ADSC and dADSC samples, cells were collected and frozen from 1 x 75 cm^2 flask, equating to approximately the same amount of starting material. Without thawing first, cell pellets were homogenised in 1 ml Trizol reagent by pipetting up and down several times. 200 µl chloroform was added and the samples mixed by inverting the Falcon tube several times followed by centrifugation at 12, 000 *g* for 15 min at 4 °C. Three x 12 ml of mature adipocytes from collagenase-digested adipose tissue were collected; one from the same individual that the ADSCs and dADSCs were obtained from, and two other individual adipose samples. Due to the high lipid content of adipocytes and the large starting volume, a modified version of the above Trizol treatment was applied based on Guan and Yang (2008). Without thawing, adipocytes were homogenised with 30 ml Trizol reagent and the Falcon tube inverted several times and then spun at 12, 000 *g* for 10 min at 4 °C. The top layer of lipid was completely removed, before transferring the pink layer of Trizol to a fresh tube, leaving behind the

pellet which contains cellular membranes, polysaccharides and high molecular weight DNA (Guan and Yang, 2008). 13 ml chloroform was then added to the Trizol homogenate and the tubes shaken gently before centrifugation. Following chloroform addition and centrifugation for ADSCs, dADSCs and adipocyte samples, the upper aqueous phase was collected and mixed with an equal volume of 70 % ethanol and the bottom layer discarded. Further cleanup of crude RNA contained in the upper phase was performed using the RNeasy Mini Kit following the manufacturer's instructions. 500 µl of the ethanol mixture was applied to the RNeasy spin columns. Because of the larger volumes used in the processing of adipocytes, the resulting ethanol mixture was approximately 12 ml. This volume was passed through a single RNeasy spin column for each sample in 500 µl amounts followed by centrifugation until all the solution had been passed through the column.

Following cleanup using the RNeasy columns, RNA from all samples were treated with DNase I enzyme for removal of contaminating genomic DNA using the RNase-Free DNase Set. DNase I was added and incubated at room temperature for 10 min, before removal by cleanup using the RNeasy spin columns once again. RNA was eluted from the column in 2 x 15 µl nuclease-free water before determining concentration and purity using a NanoDrop 1000 Spectrophotometer. Sample purity was determined using the ratio of absorbance readings taken at 260 and 280 nm (A260/280 ratio).

For cDNA synthesis, 1 µg of RNA stock from each sample was reverse-transcribed using the QuantiTect Reverse Transcription Kit according to the manufacturer's instructions. The resulting cDNA was then diluted to a working concentration of 2 ng/µl and aliquots stored at -70 °C until further mRNA expression analysis. Remaining RNA was also diluted to a concentration of 2 ng/µl and aliquots stored at -70 °C.

2.2.11 qRT-PCR to measure mRNA expression of marker genes

qRT-PCR was performed in 10 µl PCR reactions containing 2.5 µl of primers (2 µM of forward and reverse primers), 2.5 µl cDNA (2 ng/µl) or RNA (2 ng/µl) and 5 µl GoTaq qPCR Master Mix. The primer sequences used for adipocyte marker genes and the stem cell marker genes are shown in Table 2.2. qRT-PCR was performed using a Mastercycler Realplex (Eppendorf International, Germany) and reactions performed in technical duplicate for each sample. Cycling conditions were as follows: initial denaturation at 95 °C for 2 min, followed by 40 amplification cycles of 95 °C for 15 s, 55 °C for 15 s and 68 °C for 15 s. Expression values were normalised to the corresponding TATA-box binding protein (TBP) expression level within the same sample and reported as $2^{-\Delta CT}$ (Schmittgen and Livak, 2008). For comparison of mRNA expression of ADSCs, dADSCs and adipocytes, Student's t-tests (unpaired with unequal variance) was performed using GraphPad Prism 6. *P*-values < 0.05 were considered significant.

Elimination of contaminating genomic DNA was confirmed in all samples by setting up parallel reaction mixes containing RNA as the reaction template molecule, run alongside cDNA as the reaction template. For all primer pairs, primer efficiency analysis was performed as described by Brzoska and Hassan (2014), using cDNA templates over several orders of magnitude to ensure gene amplification with minimal primer-dimers. All primers were designed using the Primer-BLAST software (Ye, Coulouris et al., 2012) (available at http://www.ncbi.nlm.nih.gov/tools/primer-blast/), specifying products to be between 50-150 bp and otherwise using default settings. For each gene, when primers targeting different transcript variants were presented in the Primer-BLAST software, primers targeting the longest gene transcript or the most transcript variants were selected.

Table 2.2 qRT-PCR primer pair sequences

Gene Name	Symbo	Forward (5'-3')	Reverse (5'-3')	Pr mer Target/s	Product S ze (bp)
Adipogenic marker genes:					
Peroxisome proliferator-activated receptor γ2	PPARγ2	ACCCAGAAAGCGATTCCTTC	ACGGAGCTGATCCCAAAGTT	Transcr pt var ant 2	112
Perilipin-1	PLIN1	CAAGGAAGAGTCAGCCCCTG	GCACGGTGTATCGAGAGAGG	Transcr pt var ant 1	101
Perilipin-4	PLIN4	CCAGTTCCAGGACCACAGAC	CTACACTGAGCACATCCGGG	D fferent transcr pt var ants	133
Fatty acid binding protein 4	FABP4	ATGGGGGTGTCCTGGTACAT	AACTCTCGTGGAAGTGACGC	Spec f c to FABP4	119
Mesenchymal stem cell marker gene:					
5'-nucleotidase, ecto (CD73)	NTSE	TGGTGGAGATGGGTTCCAGA	CCGACCTTCAACTGCTGGAT	Transcr pt var ant 2	107
Housekeeping genes:					
tyrosine 3-monooxygenase/ tryptophan 5-monooxygenase activation protein, zeta	YWHAZ	AACAGCAGATGGCTCGAGAA	TGCTTGTGAAGCATTGGGGA	D fferent transcr pt var ants	98
TATA-box binding protein	TBP	GCACCACTCCACTGTATCCC	GCTGCGGTACAATCCCAGAA	Transcr pt var ant 1	116
β-glucuronidase	GUSB	AACAAGCATGAGGATGCGGA	TCTGCATAGGGGTAGTGGCT	Transcr pt var ants 1 & X1	129

2.3. Results

Various microscopic techniques were used to investigate and compare the morphology of ADSCs and their *in vitro* and *in vivo* progeny. Adipocytes have a unique morphology, making microscopy an ideal

tool to determine whether the phenotypic changes that occur during *in vitro* adipogenesis mirror that of a native adipocyte. Large-scale differentiation of ADSCs *in vitro* was also evaluated by measuring gene expression of the adipocyte marker genes PPARγ2, PLIN1, PLIN4 and FABP4, and a stem cell marker CD73 by qRT-PCR.

2.3.1 Imaging adipose tissue using Differential Interference Contrast and fluorescence microscopy

Adipose tissue was stained with fluorescent Oil Red O for lipid, and nuclei stained with fluorescent Hoechst before imaging on the Cell R Live Cell imaging platform. The merged DIC and fluorescence image is shown in Figure 2.4a, while fluorescence is shown in Figure 2.4b. Fluorescence imaging enabled visualisation of the loose connective nature of adipose tissue and the SVF cells contained are seen as chains of blue nuclei (Figure 2.4b). The use of DIC to image the adipose tissue was not required as evident in Figure 2.4a, where the merged image is convoluted and lacks specific cellular details. Using fluorescence alone, the membranes of the individual adipocytes are still visible and the blue nuclei more easily discernible.

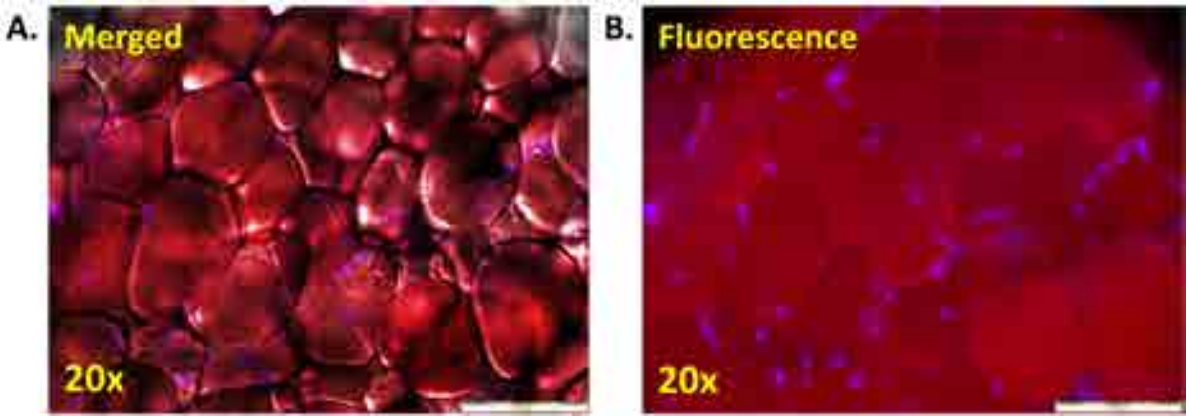


Figure 2.4 Adipose tissue was stained with Oil Red O and nuclei stained with Hoechst 33342. Tissue was imaged using the Olympus Cell R – Live Cell imaging platform. The same slide area was imaged with fluorescence and DIC imaging. The merged DIC and fluorescence images are shown in (**A**) and the fluorescence alone in (**B**). Images were taken at 20x magnification and the scale bar denotes 100 µm.

Adipose tissue was also imaged on an Olympus FluoView 1000 inverted confocal microscope. The superior focusing ability of the confocal microscope was able to capture sharper images (Figure 2.5) in comparison to the Cell R system. Again, the fluorescence image alone (Figure 2.5b) provided better overall structural information aided by the contrast created by the lipid and nucleic stains with the black background. In both the fluorescence and merged images there are blue nuclei seen near the periphery of some adipocytes, however it is not possible to distinguish these as adipocyte nuclei or the nuclei of adjacent stromal vascular cells. Therefore, fluorescence imaging was the superior method for imaging adipose tissue in both imaging systems although ultimately the confocal microscope provided sharper images and greater structural information.

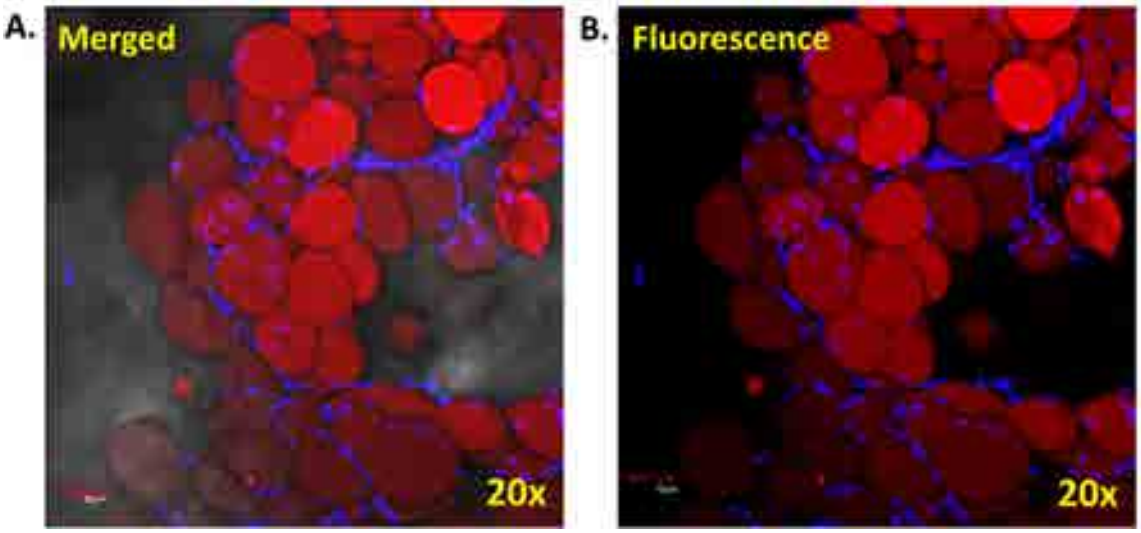


Figure 2.5 Adipose tissue stained with Oil Red O and counter stained with Hoechst were imaged using an Olympus FluoView 1000 confocal microscope. Tissue was imaged at 20x magnification with (**A**) and without DIC overlay (**B**). Scale bars in the lower left hand corner indicate 50 µm.

2.3.2 Imaging adipocytes from collagenase-digested adipose tissue

In order to determine the characteristic adipocyte phenotype with its unilocular lipid vesicle and nucleus pushed to the cell's periphery, adipocytes from collagenase-digested adipose tissue were imaged. Digestion of adipose tissue and subsequent centrifugation causes the SVF cells to pellet, helping to minimise interfering SVF in staining and imaging. Despite this, imaging individual adipocytes proved a challenging task. When placed on a microscope slide, adipocytes were not always a single monolayer of cells, but clumped together increasing the sample density. By imaging around the sample periphery which contained areas flattened by the microscope slide cover slip, an individual adipocyte was imaged on both the Cell R (Figure 2.6a) and FluoView confocal microscope (Figure 2.6b). DIC images were also taken but are not shown in this example as they convolute the image detail, as was seen in Figure 2.4 and 2.5. Using the FluoView confocal microscope enables the overlay of several optical sections and provides a sharper image than the Cell-R as confocal imaging allows for the more precise visualisation of fluorescent signals (Shen and Chen, 2008). The image taken with the Cell R contains high background from Hoechst staining which interferes with the contrasting detail that fluorescence can provide as also seen in Figure 2.6b. Despite this, the adipocyte nucleus is still visible on the left edge of the cell in both images. When imaging adipocytes, some contaminating nuclei from chains of SVF cells can still remain therefore great care was taken to focus on areas where such chains were not present to ensure that what is seen are nuclei from adipocytes. Since there are no obvious SVF nuclei chains in Figure 2.6 this indicates the nucleus on the left edge does belong to adipocyte in both images.

Because adipocytes are fragile cells, many burst during the collagenase digestion and staining process, releasing their lipid contents which form lipid droplets visible under microscopy. Unstained or stained with Oil Red O alone, these lipid droplets are almost morphologically identical in appearance to true adipocytes. They can usually be distinguished however, as they form perfectly spherical shapes compared with an adipocyte, and when stained with Hoechst the absence of a nucleus in the lipid droplets also sets them apart. Some lipid droplets can be seen below the adipocyte in Figure 2.6b, and some to the top left of the image with motion blur from the overlay of image sections taken over time. These lipid droplets are marked with arrows.

The different imaging systems also resulted in differences in depth of field. In Figure 2.6a, the adipocyte is at the forefront of the image, with some adipocytes in the background to the right and below. When using wide-field fluorescence, both in-focus and out-of-focus light from other focal planes are captured in the image (St Croix, Shand *et al.*, 2005). Imaging of the adipocytes was also complicated by the high light refractive index of lipid which decreases the ability to focus on the sample and thus renders an out-of-focus appearance (Figure 2.6a). This problem was avoided by using the confocal microscope (Figure 2.6b) which, like imaging adipose tissue, proved to be the preferred method for visualising adipocytes.

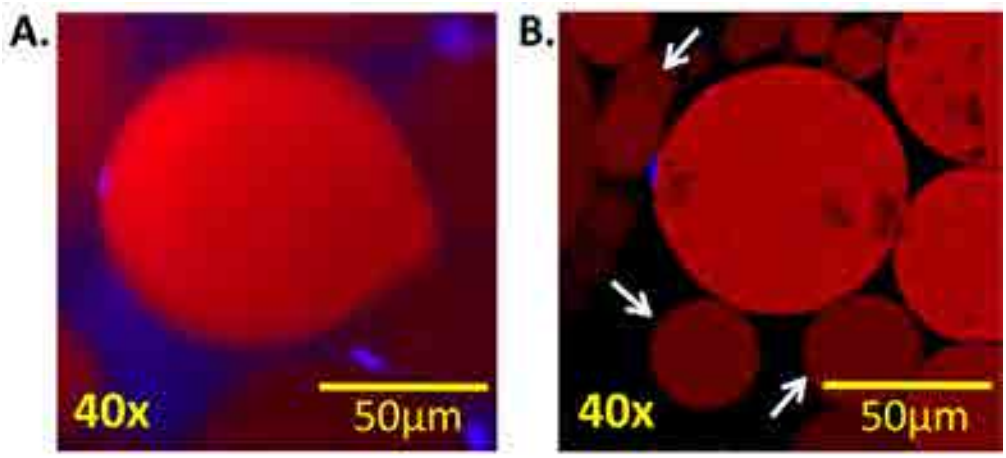


Figure 2.6 Adipocytes from collagenase digested adipose tissue were stained with Oil Red O for lipid and nuclei stained with Hoechst. Cells were imaged at 40x magnification using fluorescence microscopy on the Olympus Cell R (**A**) and Olympus FluoView confocal microscope (**B**). The scale bar denotes 50 µm. Lipid droplets are marked with arrows

2.3.3 In vitro *adipogenesis of ADSCs*

To image the *in vitro* differentiation of ADSCs over time, ADSCs were grown in 35mm^2 glass-bottomed dishes. Because the microscopes used in this work are inverted microscopes, using a culture vessel with a thin glass bottom allows for clear imaging of cells as the lens objectives are below the stage facing up. Cells were imaged at 20x magnification to provide an overall indication of

percentage of cells differentiating, and also at 40x magnification to provide a more detailed view of the specific morphological changes and degree of differentiation taking place (Figure 2.7). The degree of Oil Red O red staining seen in red, allows the accumulation of lipid in the cells to be visualised. While lipid vesicles can be viewed without staining, the use of Oil Red O is not only an established test for adipogenesis *in vitro* (Bunnell, Estes *et al.*, 2008; Zuk, Zhu *et al.*, 2001), but the presence and accumulation of red stain helps visualise the cellular changes occurring. After the addition of Differentiation Media, the number of red-stained lipid vesicles began to increase in number (Day 5 to Day 8), before increasing in size due to lipid accumulation (Day 12 onwards).

The investigation by Zuk *et al*. (2001), from which the present adipogenic differentiation method is based, differentiated ADSCs for 14 days, however 30 days was chosen in order to differentiate as many of the cells as possible. By Day 15, and especially Day 19 onwards, cell-free areas appeared in the dish. As the ADSCs progress further into the adipogenic differentiation the cells began to lift off the flask, most likely because they were accumulating with lipid. The inset image in Day 19 shows a partially adhered cell now in a different plane of focus. Some of these loosely-adhered cells were also lost in the staining process, further contributing to cell-free areas. Overall, given that the ADSCs were at least 90 % confluent at Day 1, the cell-free areas indicate that the ADSCs have differentiated as the control cells were still adherent. By Day 26 and Day 30, very few remained adhered to the dish, so it can be estimated that at least 90 % of ADSCs had differentiated and lifted off. Also by these time points, some differentiated cells had a morphology more like that of native adipocytes with a central large lipid vesicle taking up most of the cell and the nucleus pushed to the cell edge (inset Day 26 and 30, Figure 2.7).

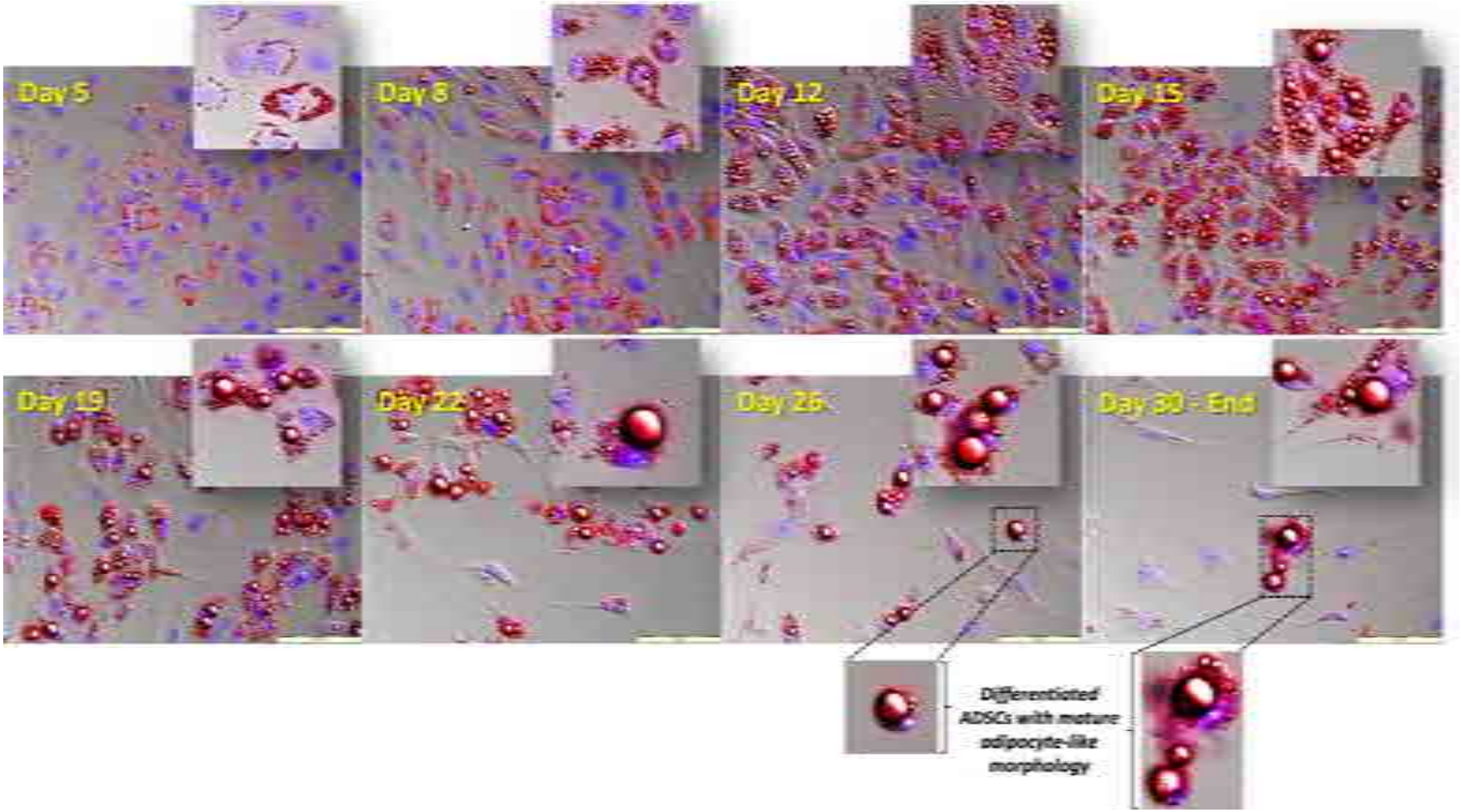


Figure 2.7 Representative time course imaging of *in vitro* adipogenic differentiation of ADSCs from one biological replicate over 30 days. At each differentiation time point, a dish was taken for staining with Oil Red O for lipid, and nuclei counter stained with Hoechst. Cells were imaged using DIC and fluorescence microscopy on the Cell-R. Cells were imaged at 20x magnification (scale bar denotes 100 µm). Inset images were taken at 40x magnification.

A time course of control ADSCs maintained in Standard Media was also undertaken as shown in Figure 2.8 which includes an image just prior to induction of adipogenic differentiation, at Day 8, Day 15, and the end of the differentiation period at Day 30. Interestingly, some small red vesicles were already present in ADSCs prior to induction of differentiation. An increase in ADSC confluency was seen over the three subsequent time points however the size of the lipid vesicles remained small, unlike the treated ADSCs as seen in Figure 2.7. This indicated that ADSCs maintained in prolonged culture (30 days) without passaging do not progress into an adipogenic lineage.

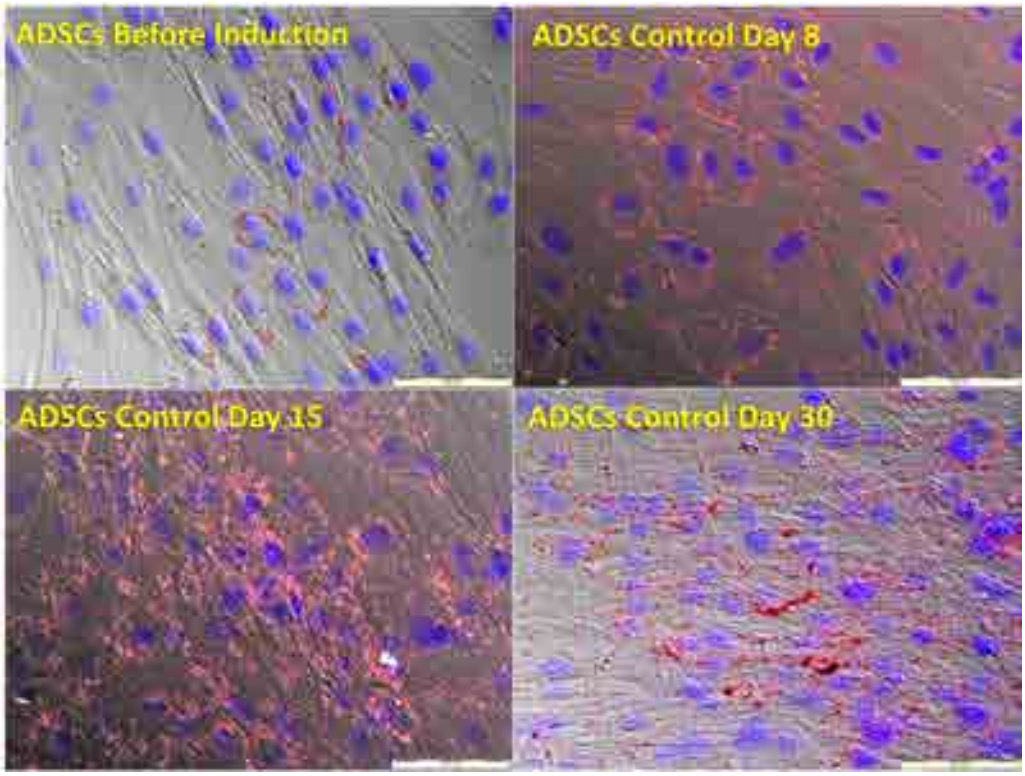


Figure 2.8 Time course imaging of ADSC controls just prior to induction of adipogenic differentiation, and during differentiation at Day 8, 15 and 30. Cells were stained with Oil Red O for lipid and nuclei stained with Hoechst. Cells were imaged at 20x magnification and the scale bar denotes 100 µm.

2.3.4 Phase contrast microscopy to visualise in vitro *large-scale adipogenesis*

For the purpose of investigating the changes in membrane glycosylation during adipogenic differentiation of ADSCs using glycomics and proteomics analyses, a larger amount of material was required. In membrane glycomic analyses of colon cancer cell lines by a colleague at Macquarie University, a sufficient amount of material could be successfully analysed from cells grown to near confluency in a 145 cm^2 culture dish (J. Chik, personal communication). Differentiation was therefore conducted in larger culture vessels than the 35 mm^2 dishes described in the previous section which have a culturing surface area of 9.6 cm^2. ADSCs were grown to approximately 90 % confluency in 75 cm^2 culture flasks before induction of adipogenic differentiation for seven days and were then

stained with Oil Red O to confirm and check the differentiation rate (Mehlem, Hagberg *et al.*, 2013). Cells were viewed with phase contrast microscopy using a Leica DM IL inverted microscope alongside an undifferentiated stained control (Figure 2.9). The amount of red staining present after seven days was similar to that of the Day 5 time point when differentiating ADSCs in 35 mm^2 dishes (Figure 2.7). This indicated that the differentiation process was still in its infancy, given the adipocyte-like morphology attained by cells at the later time points in the smaller dishes (Figure 2.7).

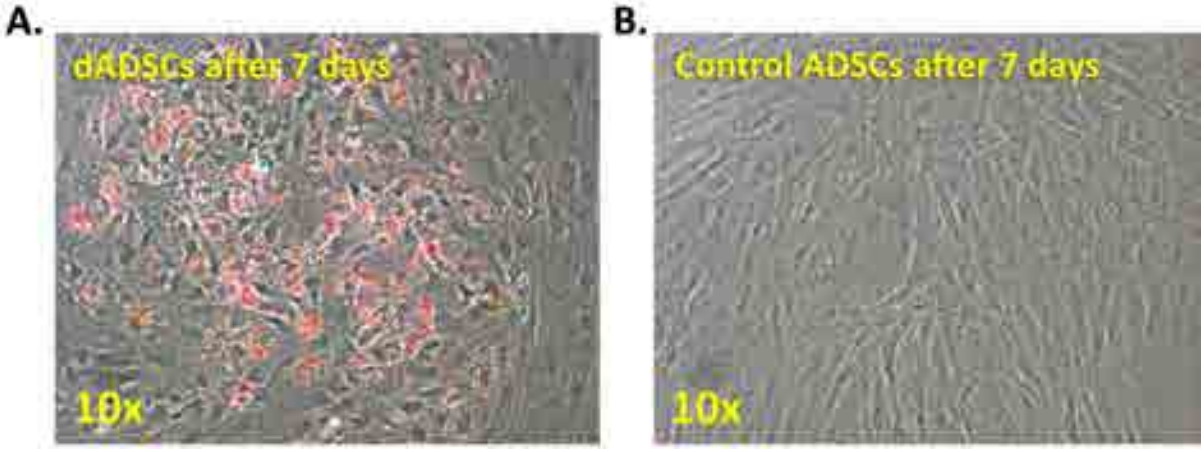


Figure 2.9 Phase contrast microscopy image of dADSCs after seven days of adipogenic differentiation confirmed adipogenic differentiation had commenced (**A**), while the controls remained undifferentiated (**B**). Both cell types were stained with Oil Red O stain for lipid.

Adipogenic differentiation in 75 cm^2 flasks was therefore conducted for the same time (30 days) to obtain the degree of differentiation seen in 35 mm^2 dishes. To avoid losing loosely-adhered dADSCs in the Oil Red O staining process, cells were not stained prior to imaging to more accurately estimate the percentage of cells that had differentiated (Figure 2.10). As mentioned previously, the lipid vesicles in dADSCs can be imaged without Oil Red O staining due to the different light refractive properties of lipid and the rest of the cell. dADSCs were thus imaged at 4x and 10x magnification by phase contrast microscopy (Figure 2.10a & b) to gain an overview of the percentage of differentiated cells, and at 40x magnification (Figure 2.10c) to compare the degree of differentiation. Large cell-free areas can be seen at all three fields of view in Figure 2.10; as with the differentiation in 35 mm^2 dishes, it is assumed that these cell-free areas are the result of dADSCs that have lifted off the bottom of the flask.

At 4x magnification, the way in which the lipid vesicles refracted the light caused the dADSCs to appear as darker areas of cells, while less or un-differentiated cells were seen as lighter areas (in boxes marked in Figure 2.10a). Looking at the area of cell-free space, together with the differentiation ADSCs still adhered, indicated that over 75 % of cells had differentiated. Imaging at 10x magnification indicated that only about half of the cells that still adhered had begun to

differentiate but had not reached the morphology of a mature adipocyte. Differentiated cells can be seen as a spherical shape as they are no longer fully adhered to the flask and have increased lipid (marked with arrows in Figure 2.10b). Overall, based on the qualitative observations dADSCs did not seem as fully differentiated as those at the 30-day time point in 35 mm^2 dishes. The two images shown in Figure 2.10c and 2.10d taken at 40x magnification look most similar to the degree of differentiation seen at Day 19 in the 35 mm^2 dishes. Again, some of the remaining loosely-adhered cells have formed spherical shapes with unilocular lipid vesicles like that of a mature adipocyte and are marked with arrows.

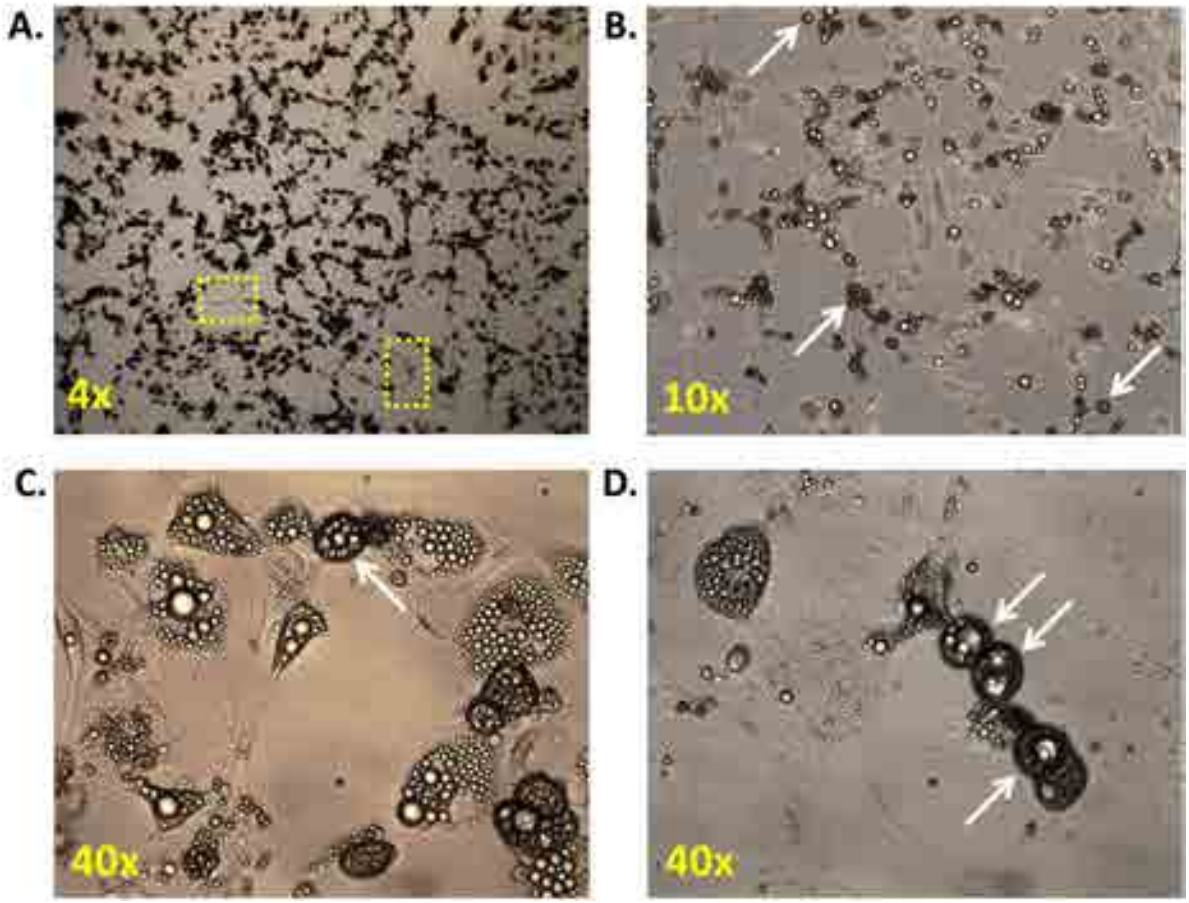


Figure 2.10 Unstained dADSCs after 30 days *in vitro* adipogenic differentiation in 75 cm^2 flasks imaged using phase contrast microscopy. Cells were imaged at 4x (**A**), 10x (**B**), and 40x (**C** & **D**) magnification. Loosely-adhered dADSCs that have attained the morphology of a mature adipocyte with a large central lipid vesicle are marked with arrows.

In addition to this, 30 day differentiated dADSCs were also stained with Oil Red O to confirm their differentiation status in line with the literature standard. This was done on differentiated ADSCs from a different individual to that imaged in Figure 2.10. There were no loosely-adhered spherical cells with unilocular lipid vesicles typical of a mature adipocyte. These cells have presumably been lost in the multiple washing steps in the staining process. As staining is not required before conducting

glycomic and proteomic analysis, the loss of these more differentiated cells was considered not to be a problem. The amount of red staining in the cells (Figure 2.11a) remaining indicated that approximately half the cells had differentiated to some degree, in line with the observation seen in the 10x image in Figure 2.10b.

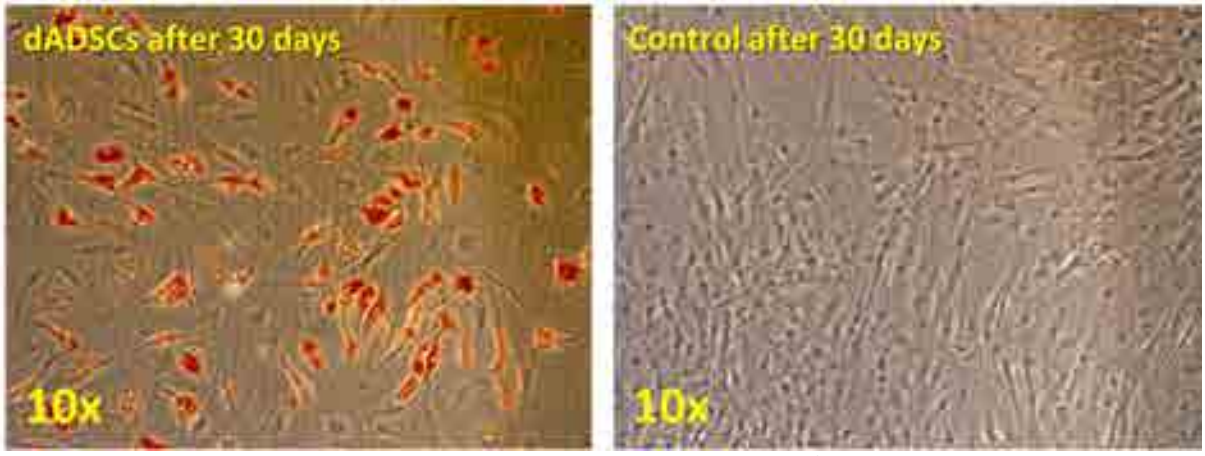


Figure 2.11 Oil Red O staining of ADSCs differentiated for 30 days in 75 cm^2 flasks (**A**) and control ADSCs (**B**). Images were taken at 10x magnification.

The degree of cell-free space is assumed to be created from cells that have fully or near-fully differentiated and subsequently lifted off the flask. These cell-free spaces cannot be attributed to dead cells lifting off as dead cells have a distinct morphology and remain adhered to the plate. Some cells do not survive after the Adipogenic Differentiation Media is added and die, after which they no longer exhibit a fibroblastic ADSC shape, and are clearly identifiable by microscopy (indicated with arrows; Figure 2.12). While some red lipid vesicles can be seen in these senescent cells this is most likely the same small vesicles seen in the control ADSCs prior to induction. As phase contrast is not as sensitive as fluorescence microscopy in detecting the stained vesicles, a smaller number of such vesicles are apparent compared to the control ADSCs visualised by Oil Red O.

For the purposes of large-scale differentiation, the results indicate phase contrast microscopy to be a suitable method for confirming and estimating ADSC differentiation. After 30 days' differentiation at the time of harvest for downstream glycomic and proteomic analysis, approximately 75 % of ADSCs would have differentiated with varying degrees of differentiation. Overall however, the large-scale differentiation of ADSCs in 75 cm^2 flasks was not as effective compared with 35 mm^2 dishes.

Figure 2.12 Oil Red O staining of ADSCs differentiated for 30 days in 75 cm^2 flasks imaged at 20x magnification. Dead cells that do not survive after the addition of Adipogenic Differentiation Media are marked by arrows. Images were taken at 20x magnification.

2.3.5 Comparing in vivo and in vitro *products of adipogenic differentiation*

In order to visually compare the products of *in vitro* and *in vivo* adipogenesis, adipose tissue, dADSCs and their progenitor ADSCs imaged using DIC and fluorescent microscopy are shown in Figure 2.13. As discussed in Section 2.3.4, imaging adipocytes does not require the additional use of DIC however DIC imaging of ADSCs and dADSCs was required to observe cell size and shape. Shown together and imaged at the same magnification, the significant changes in size and morphology can be appreciated between the three cell types. In particular, it can be seen that the differentiated ADSCs, although changed in morphology and accumulating lipid droplets, do not resemble the fully differentiated native adipocytes.

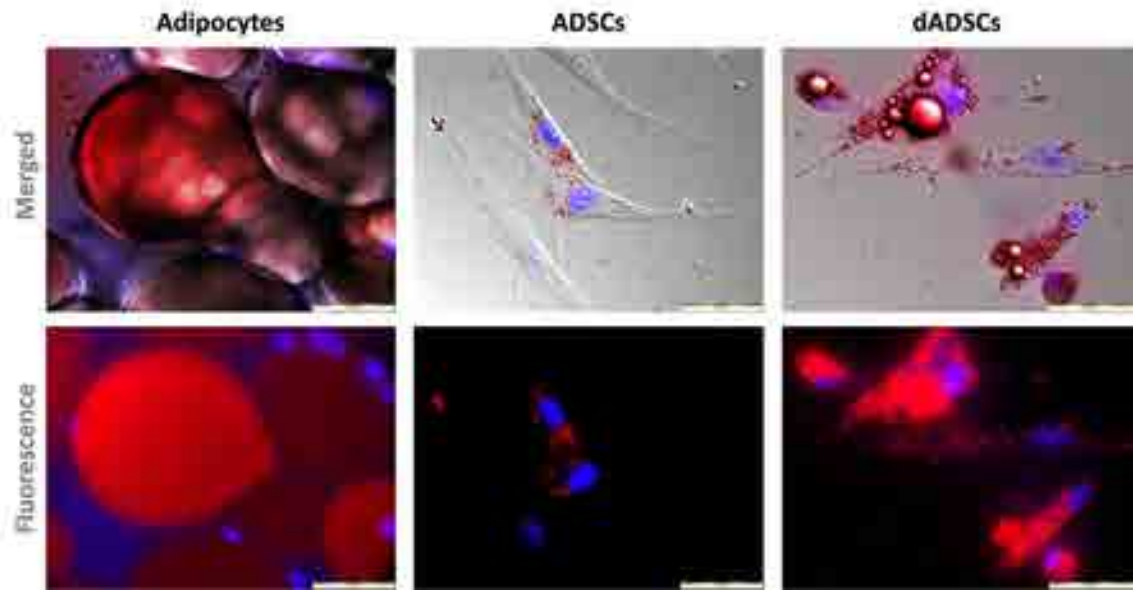


Figure 2.13 Adipocytes, ADSCs and dADSCs were imaged using DIC and fluorescence microscopy. Cells were stained with Oil Red O for lipid and nuclei stained with Hoechst. The top panel of images are an overlay of DIC and fluorescence images while the bottom panel is the fluorescence image alone. All images were taken at 40x magnification and the scale bar denotes 50 µm.

2.3.6 mRNA expression of cell-marker genes by qRT-PCR

As demonstrated by Oil Red O staining, ADSCs differentiated for 30 days in 75cm^2 flasks had entered the adipogenic lineage. To further investigate the adipogenic differentiation stage qRT-PCR was used to measure expression levels of adipocyte marker genes PPARγ2, PLIN-1, PLIN-4 and FABP4 in ADSCs, dADSCs and adipocytes and the stem cell marker gene CD73 was also measured. Total RNA extraction from ADSCs yielded an average of 359.9 ng/µl ± 73.7 after DNase treatment, while dADSCs yielded an average of 165.2 ng/µl ± 39.2 despite using the same starting material amount (cells from 1 x 75 cm^2 flask). Greater variability was seen in the yields of RNA collected from adipocyte samples, with an average of 765.1 ng/µl ± 458.7 collected. Sample purity was determined using A260/280 absorbance ratios and ranged from 1.99 2.07 across the nine samples.

Of the three housekeeping genes tested (Figure 2.14), TBP displayed the most stable expression across cell types with the highest *P*-value as determined by ANOVA (*P*<0.05).

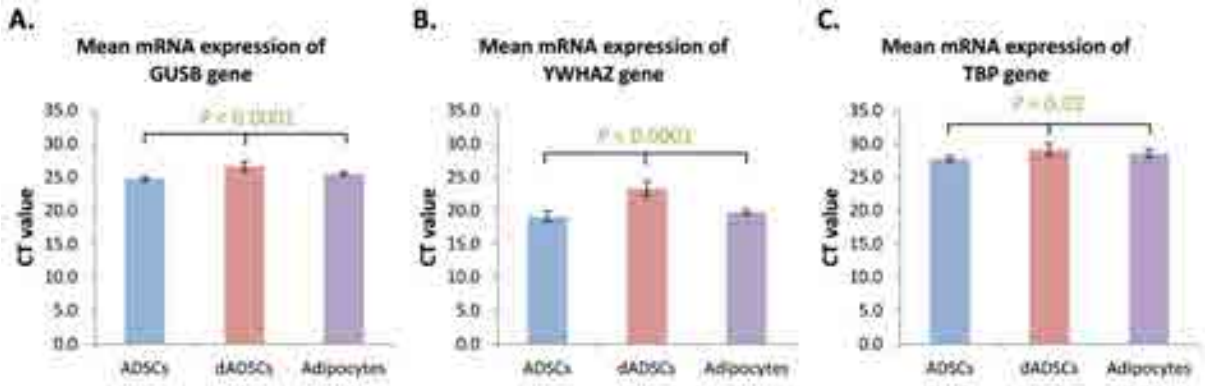


Figure 2.14 Raw qRT-PCR cycle threshold (CT) values for the housekeeping gene GUSB, YWHAZ and TBP. One-way ANOVA testing showed significant differences in expression across ADSCs, dADSCs and adipocytes for all genes. Of the three, TBP varied the least between cell types and was used for normalising genes of interest.

Expression levels of adipocyte marker genes were all found to be significantly higher in dADSCs and adipocytes when compared with their progenitor ADSCs, although at varying degrees of increase for each gene (Figure 2.15a-d). Of these four adipogenic marker genes, PPARγ2, Perilipin-1 and Perilipin-4 displayed lower expression in dADSCs compared with native adipocytes. In the case of Perilipin-1 and Perilipin-4 this difference was significant (*P* = 0.0007 and *P* = 0.0005 respectively). The expression levels of FABP4 however were higher, although not significantly, in dADSCs than native adipocytes. Overall these results confirm that dADSCs have entered the adipogenic lineage. This finding was further supported in the analysis of stem cell marker gene CD73, which was significantly decreased in dADSCs and adipocytes in accordance with a departure from the mesenchymal stem cell lineage (Figure 2.15e).

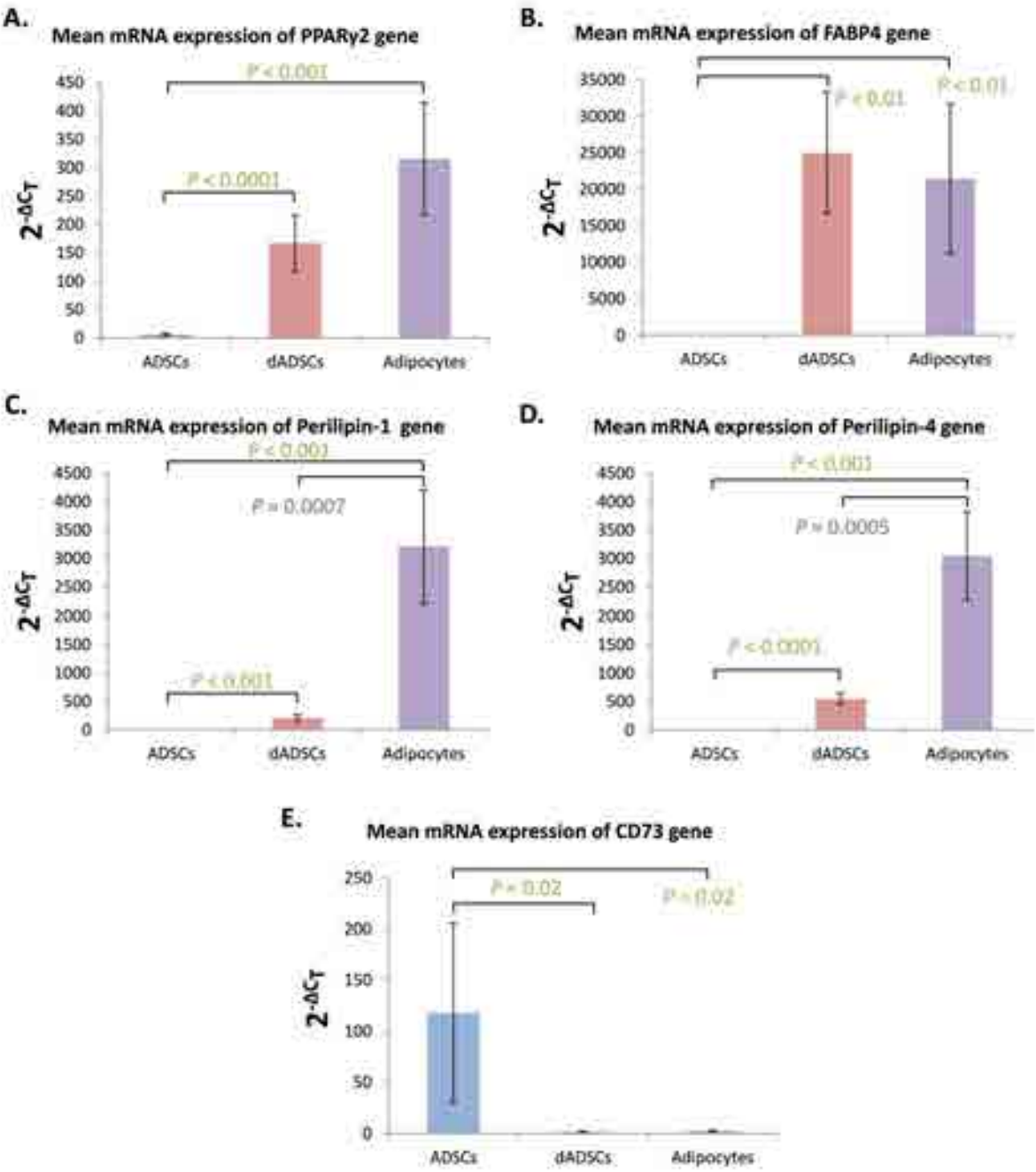


Figure 2.15 Mean mRNA expression levels of adipocyte marker genes (**A-D**) and the stem cell CD73 marker gene (**E**) were measured using qRT-PCR and compared between ADSCs, dADSCs and adipocytes. Expression levels were normalised to the TBP housekeeping gene and shown as $2^{-\Delta CT}$ mean ± standard deviation. *P*-values from student's *t*-test comparison of ADSCs and dADSCs, and ADSCs and adipocytes are shown where significant.

2.4 Discussion

In this chapter, various microscopic techniques were employed to image ADSCs and their *in vivo* and *in vitro* adipogenic progeny in order to determine the progression of differentiation of the stem cells into adipocytes. Adipose tissue and adipocytes from collagenase-digested adipose were stained for lipid and nuclei counter stained to aid visualisation of cell components and imaged with DIC and fluorescence microscopy. Time course imaging of *in vitro* adipogenic differentiation of ADSCs was also performed. Optimisation of ADSC differentiation in larger culture flasks was conducted to obtain sufficient material for subsequent analysis and the differentiation was compared to that in smaller flasks. The differentiation process was further evaluated by measuring adipocyte marker and stem cell marker gene expression changes using qRT-PCR.

Imaging adipose tissue using DIC and fluorescence microscopy revealed fluorescence microscopy alone provides the best tissue structure detail. While DIC can be utilised to enhance contrast of unstained or transparent samples, adipose tissue does not contain such regions. Because the lipid content of an adipocyte can be anywhere up to 95 % of cell volume (Bashkatov, Genina *et al.*, 2005), Oil Red O staining covers much of the cell area, allowing it to be imaged easily. Many of the protocols for Oil Red O staining of lipid are directed to *in vitro* differentiated cells which are already adhered to a culture vessel. Staining adipose tissue in this study required a modification of this protocol, with staining conducted directly on the microscope slide and a longer staining time to allow the Oil Red O to penetrate the sample. The FluoView confocal microscope provided sharper images than the Cell R inverted microscope through its ability to more precisely focus on fluorescence signals (Shen and Chen, 2008). Imaging a less dense area of adipose tissue at 20x magnification allowed a closer view of adipocytes but the characteristic adipocyte morphology where the nucleus is pushed to the edge of the cell by the central lipid vesicle was not able to be seen because of the presence of the nuclear staining of SVF cells.

The imaging of adipocytes from collagenase-digested adipose tissue helped to circumvent the problem of contaminating SVF nuclei. However, this presented different challenges, namely the fragile nature of adipocytes causing them to easily burst, thereby decreasing the number of cells suitable for imaging. The examination of cells isolated on the periphery of the sample resulted in the imaging of an individual adipocyte and its nucleus on both the FluoView and Cell-R. The aim of imaging a mature adipocyte and its nucleus to demonstrate the characteristic adipocyte morphology was achieved in both imaging systems with the FluoView confocal microscope providing a sharper image.

To compare the morphology of *in vitro* differentiated adipogenic products with the mature adipocytes, time course imaging of ADSC differentiation was performed. The adipogenic differentiation protocol was based on an established method in the literature (Zuk, Zhu *et al.*, 2001) that was used to demonstrate the multi-lineage differentiation capacity of ADSCs. The method used in this study differs in that penicillin/streptomycin solution was used in place of penicillin/streptomycin/anti-mycotics as anti-mycotics are not necessary for culture of human cells (S. Blaber, personal communication), and the use of DMEM-HG medium as opposed to plain DMEM which produces a higher percentage of differentiated cells (Master's thesis; H. Hinneburg, 2010).

Time course imaging of *in vitro* differentiated ADSCs portrayed the overall changes in cell shape from a fibroblastic shaped stem cell to rounded, more differentiated cells. The DIC overlay in this instance at both 20x and 40x magnification provided information about overall percentage of differentiation and also a closer view of the specific morphological changes occurring. The major limitation in determining overall percentage of differentiation is that many cells are lost during the staining process. As the dADSCs further proceed down the adipogenic lineage, they start to lift off the bottom of the well, most likely due to the accumulation of lipid, causing them to float. This has also been reported by Zuk *et al*. (2001), and importantly demonstrates that the dADSCs are adopting the physical characteristics of mature adipocytes that float in culture media. By assuming that cell-free areas are created by dADSCs lifting off and combining this with the portion of remaining cells staining positively with Oil Red O, it is possible to estimate the overall percentage of differentiated cells: given that ADSCs are at least 90 % confluent at induction of differentiation (Figure 5) and approximately a dozen remain adhered by Day 30, this indicates that approximately 90 % of cells have differentiated. Since Zuk *et al*. (2001) reported an average differentiation level of 42.4 % (by calculating the number of Oil Red O positive cells as a percentage of the total cells), this is an encouraging result. In a potential therapeutic context, this high percentage is also advantageous.

Adipogenically differentiated ADSCs offer a promising alternative to the transplantation of autologous adipose tissue for injection into areas requiring soft tissue augmentation due to injury or for cosmetic purposes (Kim, Kim *et al.*, 2015; Scott, Nguyen *et al.*, 2011). In a recent report, the use of a densitometry based measurement of Oil Red O staining is proposed as a valuable method to quantify degree of staining (Mehlem, Hagberg *et al.*, 2013). The presence of small Oil Red O stained vesicles in the ADSC controls indicates that the ADSCs are perhaps already in a 'pre-adipocyte' state, but when kept in standard cell culture conditions do not progress down the adipogenic lineage. The percentage of the ADSC-containing human SVF that are pre-adipocytes capable of adipogenic differentiation however has been reported to be as low as 0.02 % (Pettersson, Cigolini *et al.*, 1984).

Scaling up the differentiation of ADSCs slowed the overall differentiation rate. After 30 days, the overall differentiation rate (approximately 75 % of cells) equalled that seen after 19 days of differentiation in smaller wells where the lipid droplets were still small and many, with few cells exhibiting a unilocular appearance. This could be due to the cells being cultured and differentiated on different surfaces, with a preference for the glass-bottomed well. The amount of media used per surface area in the small wells (ml/cm^2) was also higher than in flasks. In both differentiation environments there was also approximately 5-10 % ADSCs which were dead and still adhered. This is important to note as these cells would be a small contaminant in downstream analysis of the ADSCs differentiated in flasks. In contrast, for the majority ADSCs that entered the adipogenic lineage there were varying degrees of differentiation, with an estimated 50 % acquiring the shape of a mature unilocular adipocyte and lifting off the culture flask, whilst approximately 25 % did not. It would however be very challenging to try separating the more differentiated ADSCs from those with less differentiation or ADSCs that have not differentiated at all. Like mature adipocytes, dADSCs were also observed to be fragile cells with many lipid vesicles bursting after media removal for cell collection.

In order to further monitor stem cell differentiation in addition to staining, cell-specific marker gene/s transcript expression levels can be measured in the differentiated cell products (Hamouda, Ullah *et al.*, 2013; Wall, Bernacki *et al.*, 2007; Yu and Thomson, 2008). RNA was successfully extracted from ADSCs and dADSCs and yielded appropriate concentrations of RNA for downstream cDNA synthesis. The same protocol could not be used for adipocytes due to their high lipid content. Because of this characteristic, there are several commercial products available that have been developed specifically for samples with high lipid content. A protocol for RNA extraction from 250 µl adipocytes reported by Guan and Yang (2008) avoids the need for such products which are often expensive, and instead offers an optimised procedure based on the standard extraction method where lipid removal is incorporated. The yield of RNA obtained from this procedure with a starting volume of 250 µl adipocytes was however insufficient in this study and of poor quality. Therefore, the protocol was up-scaled to accommodate a starting volume of 12 ml adipocytes which ultimately provided enough high quality RNA. The variation seen in RNA concentrations from the three adipocyte samples may have been due to different lipid content levels or density of each sample. The liposuction method and physical disruption used to collect adipose tissue may also result in different adipocyte-to-lipid ratios which would affect the concentration of adipocytes collected after collagenase digestion. Another challenge encountered in the isolation of RNA from all samples related to inability to normalise RNA amounts due to difficulties in enumerating the three cell types. To best combat this, ADSCs and dADSCs were taken from the same number of culture flasks of the same size, with cells at approximately 90 % confluency, and adipocytes normalised by cell suspension

volume. Overall, this work further highlights the unique nature of adipogenic products and the need for optimised handling methods, as seen in the staining and imaging of adipocytes.

For the relative quantitation of gene transcripts, qRT-PCR is the most reliable and sensitive detection method available (Guan and Yang, 2008). When measuring gene expression levels during ADSC differentiation, choosing a stable housekeeping gene is important to obtain reliable results. TBP, YWHAZ and GUSB have previously been demonstrated as the most stable housekeeping genes when investigating mesenchymal stem cell differentiation (Fink, Lund *et al.*, 2008; Ragni, Vigano *et al.*, 2012). Of these three, TBP and YWHAZ have been reported as the most stable in ADSC differentiation and in adipogenesis regardless of the progenitor stem cell source (Ragni, Vigano *et al.*, 2012). In the present work, TBP was the most stable across ADSCS, dADSCs and adipocytes, however was still found to have varied expression between the differentiated and undifferentiated cells. Expression levels of the adipocyte marker genes, PPARγ2, FABP4, PLIN1 and PLIN4 were significantly higher in dADSCs than ADSCs, confirming their adipogenic lineage. The products of these genes play integral roles in induction of adipogenesis (PPARγ2 and FABP4; De Gemmis, Lapucci *et al.*, 2006; Fruhbeck, 2008; Rosen and MacDougald, 2006) or are involved in the unique lipid vesicle structure in adipocytes (Perilipin-1 and Perilipin-4; Fruhbeck, 2008) making them ideal markers to test for adipogenic differentiation. Adipocyte marker genes were expressed at higher levels in native adipocytes than dADSCs with the exception of FABP4 which was higher in expression than native adipocytes and is considered a marker of the later stages of *in vitro* adipogenesis (Galateanu, Dinescu *et al.*, 2012). Since the expression of FABP4 revealed no significant difference ($P > 0.05$) between dADSCs and adipocytes, this indicates that FABP4 expression in dADSCs largely mirrors that of native adipocytes.

Correspondingly, the expression of a stem cell marker, CD73 (encoded by NT5E gene) was significantly lower in dADSCs than ADSCs. While there is no specific marker of ADSCs or mesenchymal stem cells, CD73 forms part of a panel of CD markers used to characterise ADSCs (Blaber, Webster *et al.*, 2012; Dominici, Le Blanc *et al.*, 2006; Hamouda, Ullah *et al.*, 2013). While many studies do not include measurement of stem cell marker genes changes in differentiation, the observed down-regulation of CD73 in adipocytes and dADSCs compared to the stem cells in particular further confirms their commitment to the adipogenic lineage. In the present work however, CD73 showed varied expression, due possibly to biological variation, or perhaps indicates it to be an inappropriate marker for 'stemness' in the current model. Future work might consider the inclusion of other 'stemness' markers such as CD90 or CD105. However, overall it can be concluded that the ADSCs are being differentiated successfully with regard to some of the genes being expressed. Although

however, the gene expression and microscopy indicates that although the ADSCs have entered the adipogenic lineage, they are not fully differentiated when compared with native mature adipocytes.

2.5 Concluding remarks

In this study, adipogenic differentiation of ADSCs *in vitro* was monitored microscopically and by measuring up-regulation of adipocyte gene markers using qRT-PCR. Together, the results confirmed adipogenic differentiation occurs *in vitro* using the proscribed initiators. *In vivo* adipocytes were compared with the differentiated stem cells, offering a morphological comparison to the products of the *in vitro* adipogenesis. There were large phenotypic differences between the stem cells, the *in vitro* adipogenic products and native mature adipocytes. Overall, while the results confirm adipogenic differentiation in line with established criteria in the literature, the ability to compare the data with native adipocytes provided a new level with which to evaluate differentiation *in vitro*. This indicated that while *in vitro* adipogenic products have phenotypic and gene expression characteristics that are similar to mature adipocytes as expected, these characteristics were in fact not expressed to the same degree in this study. In a broader sense, in order to ensure quality assurance of both ADSCs and dADSCs it may be pertinent to include the measurement of stem cell marker gene expression in addition to gene markers of the differentiation lineage. In addition to this it is recommended to undertake further orthogonal analyses to that carried out here (morphological assessment and gene marker transcripts) such as measurement of gene product expression to assess the degree of differentiation or the stem cell 'stemness' state.

Importantly however, the successful large scale optimisation of *in vitro* adipogenic differentiation conducted in this study provided the large sample amounts required to investigate these differences in adipocytes differentiation by glycomic and proteomic analyses (discussed in Chapters 3 and 4 respectively).

Chapter 3:

Characterisation and comparison of *N*- and *O*-linked glycosylation during *in vivo* and *in vitro* adipogenic differentiation

3.1 Introduction

Adipose-derived stem cells (ADSCs) are a popular source of stem cells for both research and regenerative medicine (Zuk, 2013), with their differentiation ability a key characteristic in their potential therapeutic use. Many glycoproteins have been demonstrated to play key roles in the control of lineage determination and commitment (Bowers and Lane, 2008; Christodoulides, Lagathu *et al.*, 2009; James, 2013; Lanctot, Gage *et al.*, 2007) and many glycan stem cell markers have also been reported (Capela and Temple, 2006; Hennen and Faissner, 2012; Muramatsu and Muramatsu, 2004)

Currently, no investigation into the glycosylation of ADSCs and their differentiated progeny has been undertaken to our knowledge in the literature. A glycomic analysis of ADSCs and their differentiated products therefore provides an opportunity to characterise glycan epitopes present at the cellular surface that respond to external adipogenic differentiation signals and potentially identify glycan cell-specific markers. Bringing together native adipocytes, ADSCs and ADSCs adipogenically differentiated *in vitro*, creates a robust model to study glycosylation during *in vivo* and *in vitro* adipogenic differentiation.

PGC-LC coupled with ESI-MS/MS is a robust and reproducible method for the mass spectrometric analysis of released glycans. The combination of sample separation using LC coupled with MS detection allows greater coverage of the sample and provides detailed information for structural analysis (Wuhrer, Deelder *et al.*, 2005). The use of PGC enables the separation of isomeric structures and allows for the unobstructed interrogation of MS/MS fragmentation data for each isomer which is carried out using diagnostic ions for different glycan structural characteristics (Everest-Dass, Abrahams *et al.*, 2013; Harvey, Royle *et al.*, 2008).

In this chapter, *N*- and *O*-linked glycans were released from the membrane glycoproteins of ADSCs, *in vitro* differentiated ADSCs (dADSCs) and native adipocytes collected from the same individuals. dADSC samples were obtained using the adipogenic differentiation method established in Chapter 2 for the large scale differentiation of ADSCs. Released glycans were subjected to glycomics analyses using PGC-LC coupled with negative ion ESI-MS/MS. This was conducted in biological triplicate for in-depth *N*- and *O*-linked glycan structural elucidation using MS/MS based spectral interrogation and relative quantitation.

To our knowledge the work presented in this chapter describes the first glycomic investigation conducted on human ADSCs and their *in vitro* and *in vivo* progeny. The novelty of the model used in this study lies in its inclusion of *in vivo* differentiated progeny, differing from other models of differentiation which have incorporated only the undifferentiated precursors and *in vitro*

differentiated states. As glycosylation can be cell-type specific, it was investigated whether different glycomic profiles were expressed on the *in vivo* adipogenic cell proteins and on the *in vitro* differentiated products. The main objectives of this study were two-fold:

1) Primarily, to characterise the glycans present on the membrane glycoproteins of ADSCs and their *in vivo* and *in vitro* adipogenically differentiated progeny and determine whether glycosylation changes occur during adipogenesis, and

2) To determine whether *in vitro* differentiated stem cells (dADSCs) reflect the membrane glycosylation of *in vivo* differentiated stem cells (adipocytes).

3.2 Materials and Methods

3.2.1 Materials

Tris, sodium chloride (NaCl), Triton X-114, direct blue 71, trifluoroacetic acid, sodium borohydride ($NaBH_4$), potassium hydroxide (KOH), bovine serum albumin (BSA), human IgG, bovine fetuin, Tween-20, acetone and acetonitrile were purchased from Sigma (St Louis, MO, USA). 10 % (w/v) NuPAGE gel were purchased from Invitrogen (USA). *N*-Glycosidase F (PNGase F; recombinant cloned from *Flavobacterium meningosepticum* and expressed in *Escherichia coli*) lyophilized in 10 mM sodium phosphate buffer, pH 7.2, was purchased from Roche Diagnostics (Basel, Switzerland). AG50W-X8 cation-exchange resin (H^+ form), Flamingo fluorescent gel stain and gel markers were purchased from BioRad (Hercules, CA, USA). Immobilon-P Polyvinylidene Fluoride (PVDF) membranes were purchased from Millipore (MA, USA). HPLC-grade water was obtained using a Milli-Q® Synthesis water purification system also from Millipore. Methanol and urea were purchased from Fluka Biochemica (Germany). Acetic acid was purchased from LabServ (Victoria, Australia). All exoglycosidases and PGC solid-phase extraction cartridges were purchased from Prozyme (CA, USA). HyperCarb PGC column was purchased from Thermo Hypersil (Runcorn, UK). PerfectPure C18 Zip Tips were from Eppendorf (Hamburg, Germany). PHA-E lectin and Streptavidin-Alexa 488 were purchased from Vector Laboratories (Burlingame, CA).

3.2.2 Lipo-aspirate processing, and adipocyte and ADSC isolation

Human abdominal lipo-aspirates were obtained from three patients undergoing elective liposuction procedures for cosmetic reasons. The processing of lipo-aspirate material for the collection of adipocytes and ADSCs contained in the SVF, and expansion of ADSCs was conducted as described in Chapter 2, Section 2.2.2. ADSCs were collected at each passage (as described in Chapter 2, Section 2.2.9), with the exception of passage zero to determine whether glycosylation changes occur as a

result of cell culture alone. Once the cells reached approximately 90% confluency at the third passage,

a) ADSCs from half the flasks were collected for analysis using a cell scraper, and
b) The other half were used for induction of adipogenic differentiation as described in Chapter 2, Section 2.2.6.

3.2.3 Optimisation of cell membrane preparation

Adipocytes are difficult to accurately count using conventional cell counting methods (Gorzelniak, Engeli *et al.*, 1998). Floating adipocytes collected from digested adipose tissue can be transferred to a haemocytometer counting chamber, but it is then difficult to distinguish between adipocytes and morphologically similar lipid droplets in order to gain a reliable cell count. In the present work, adipocytes were approximately normalised between biological replicates by using a starting volume of 5 ml of suspended cells for all samples (Blaber, Webster *et al.*, 2012). In order to isolate membrane proteins, the intracellular lipid needed to first be removed. To burst the adipocytes, the adipocyte suspension was treated to water bath sonication (Nielsen and Mandrup, 2014) for 20 min at ambient water temperature and then centrifuged at 2000 *g* for 15 min to allow the lipid-containing supernatant to be removed. These steps were repeated as required until no visible lipid remained. Five ml of adipocytes were prepared as above, and the resulting lipid-free cell pellets were processed according to the membrane preparation based on Nakano, Saldanha *et al*. (2011). Cell pellets were homogenised on ice in 2 ml of lysis buffer (taken from a 50 ml stock containing 0.2 M Tris-HCl (pH 7.4), 0.1 M NaCl, 1 mM EDTA and one cOmplete Protease Inhibitor Cocktail tablet) using a polytron homogeniser. Samples were centrifuged at 2000 *g* for 20 min at 4 °C to pellet nuclei and any un-lysed cells. Two ml of Tris buffer (20 mM Tris-HCl (pH 7.4), 0.1 M NaCl) was then added to the supernatant and membrane proteins were pelleted by ultracentrifugation at 120, 000 *g* for 80 min at 4 °C. The supernatant was discarded and the cell membrane protein pellets resuspended in 100 µl Tris-buffer, before the addition of 400 µl of 1 % Triton X-114 (v/v) (Bordier, 1981) in Tris-buffer to one of the samples. This mixture was homogenised by pipetting and then chilled on ice for 10 min, followed by 20 min at 37 °C. Phase partitioning was carried out by centrifugation at 1000 *g* for 3 min and the upper aqueous phase was discarded. One ml of ice-cold acetone was then mixed with the lower detergent phase and stored at -20 °C overnight. One ml of ice-cold acetone was also added to the sample not treated with Triton X-114 and stored at -20 °C overnight. For both samples, precipitated proteins were pelleted by centrifugation at 1000 *g* for 3 min and stored at -20 °C and prepared for SDS-PAGE and dot blotting in Section 3.2.5.

3.2.4 Comparison of membrane proteins and N-glycosylation with and without the use of Triton X-114 membrane protein enrichment

Two adipocyte protein pellets were re-solubilised in 10 µl 8 M urea and water bath sonicated at 50 °C for 30 min. Samples were centrifuged at 14, 000 g for 5 min before protein concentration was determined using Bradford assay. 10 µg of each sample was loaded onto a 10 % (w/v) NuPAGE gel for electrophoresis. The gel was fixed and stained using Flamingo Fluorescent gel stain overnight before imaging on the Typhoon TRIO Variable Mode Imager (GE Healthcare). Ten µg of each sample was also spotted onto PVDF membrane for *N*-linked glycan release as described in the following section.

3.2.5 N- *and* O-*linked glycan release from cell membrane glycoproteins*

For glycan release, membrane protein pellets were first re-solubilised in 10 µl 8 M urea following membrane preparation and acetone precipitation. Sample amounts of 2.5 µl were spotted onto PVDF until the entire sample had been applied. *N*- and *O*-linked glycans were released from cell membrane glycoproteins and analysed according to Jensen, Karlsson *et al*. (2012). *N*-linked glycans were released by incubation with PNGase F (2.5 units) overnight at 37 ℃. Released *N*-linked glycans were treated with 100 mM ammonium acetate pH 5 (final concentration 15 mM) at room temperature for 1 hr to de-aminate the reducing terminus. This was followed by reduction with 1 M $NaBH_4$ in 50 mM KOH for 3 h at 50 ℃, and the reaction was subsequently quenched with 1 µl glacial acetic acid.

Following the PNGase F release of *N*-linked glycans from the protein spots, *O*-linked glycans were released from the same samples by reductive β-elimination. Spots were incubated overnight in 20 µl 0.5 M $NaBH_4$ in 50 mM KOH at 50 ℃. The reaction was quenched with 2 µl acetic acid and both reduced *N*- and *O*-linked glycan alditols were cleaned using µC_8 ZipTips packed with 50 µl of AG50W-X8 cation-exchange resin (H^+ form). Samples were eluted with water and borate was removed by the addition of 100 µl methanol and drying under vacuum three times.

Samples that had been reduced and alkylated prior to spotting onto PVDF were further purified using a carbon solid-phase extraction (SPE) tip made by adding slurry of carbon material in 50 % (v/v) methanol to µC_8 ZipTips at a volume equal to the bed volume of the C18. Glycans were eluted with 40 µl acetonitrile containing 0.05 % (v/v) trifluoroacetic acid and then dried under vacuum. All samples were re-suspended in 10 µl water for PGC LC ESI MS/MS separation and analysis.

3.2.6 Exoglycosidase digestion of reduced N-*linked glycans*

To further confirm terminal monosaccharide linkages (Royle, Radcliffe *et al.*, 2006) that could not be elucidated by PGC retention time patterns alone, *N*-linked glycans were digested with α2-3,6 neuraminidase from *Arthrobacter ureafaciens* (ABS), α2-3 neuraminidase from *Streptococcus pneumonia* (NAN1) and β1-3,4 galactosidase from bovine testis (BTG) according to the

manufacturer's instructions. For example, glycans digested with NAN1 to determine sialic linkage types were incubated with 2 µl NAN1 enzyme, 2 µl NAN1 Reaction Buffer (50 mM sodium phosphate, pH 6) and 6 µl water for 37 °C overnight. Digested glycans were subsequently purified using the SPE cleanup step described above.

3.2.7 PGC-LC ESI-MS/MS and Data Analysis

N- and *O*-linked glycan alditols were analysed on a capillary LC MS/MS using an Agilent MSD ion-trap XCT Plus mass spectrometer coupled to an Agilent 1100 capillary LC (Agilent Technologies, CA, USA), as outlined by Everest-Dass, Abrahams *et al*. (2013). A single technical replicate was analysed for each biological sample (n=3). At a flow rate of 2 µl/min, *N*- and *O*-linked glycans were separated on a 5 µm Hypercarb PGC column (180 µm inner diameter x 100 mm length) and eluted directly into the ESI source. *N*-linked glycans were separated over an 85 min gradient of 0 45 % (v/v) acetonitrile in 10 mM ammonium bicarbonate. *O*-glycans were separated over a 45 min gradient of 0 90 % (v/v) acetonitrile in 10 mM ammonium bicarbonate. The voltage of the MS capillary outlet was set at 3 kV, and the temperature of the transfer capillary kept constant at 300 ℃. Spectra were acquired in negative mode with two scan events: full MS scan over a range of *m/z* 200-2200 and maximum accumulation time of 200 ms, and data-dependent MS/MS scan of the top two most abundant ions for *N*-linked glycans, and top three most abundant for *O*-linked glycans.

Glycan compositions were predicted by searching the monoisotopic masses of detected ions appearing in the base peak chromatogram using GlycoMod (Cooper, Gasteiger *et al.*, 2001) (available at http://web.expasy.org/glycomod/) using a mass tolerance of ± 0.6 Daltons. These compositions were then manually confirmed using the retention times of the extracted ion chromatograph (EIC) peaks and MS/MS fragmentation data. GlycoWorkbench was used to draw glycan structures and produce *in silico* theoretical MS/MS fragments (Ceroni, Maass *et al.*, 2008). Relative quantitation to compare glycan abundances across samples was performed by integrating the EIC peak area of each individual glycan, taking into account multiple charge states. Each peak area was then expressed as a percentage of the total area of all glycans in the sample.

Qualitative and quantitative analysis of the elucidated structures and their relative abundance was performed. Statistical *t*-tests (paired Student *t*-test) analyses of relative percentage abundances were performed using GraphPad Prism 6 with a 95% confidence level.

3.2.8 Lectin dot blotting

Bovine serum albumin (BSA), bovine fetuin and human IgG were spotted onto pre-wet PVDF membranes. The volume of fetuin and IgG was normalised to the equivalent of approximately 300 pmol of glycans (spotted in triplicate), while 20 µl of BSA was spotted in duplicate. The membrane

was blocked with 0.01 % (v/v) Tween-20 in PBS (PBST) for 2 hr before washing three times with PBST. Biotinylated PHA-E lectin (5µg/ml in PBST) was applied to the membrane overnight at 4 °C. The lectin solution was removed before the membrane was washed thrice again in PBS. Streptavidin-Alexa488 (diluted 1:2000 in PBS) was applied to the membrane for 2 hr, before washing again in PBS. The membrane was imaged using a Fuji LAS 3000 camera and densitometry analysis performed using ImageJ software (Schneider, Rasband *et al.*, 2012).

3.2.9 qRT-PCR to measure MGAT3 and Fut8 mRNA expression levels

RNA extraction, cDNA synthesis and qRT-PCR of adipocytes, ADSCs and dADSCs was performed as described in Chapter 2, Sections 2.2.9 - 2.2.11. As outlined in Chapter 2, this work was performed in biological triplicate, and with ADSCs and dADSCs originating from the same three individuals. One adipocyte sample was also from one of these individuals, and the other samples from two other separate individuals. These samples were analysed in technical duplicates resulting in n=6 per sample.

Mannosyl (β-1,4-)-glycoprotein β-1,4-*N*-acetylglucosaminyltransferase (MGAT3) and Fucosyltransferase 8 (α(1,6) fucosyltransferase or Fut8) expression levels were normalised to TATA-box binding protein (TBP) expression in the same sample, which was established as the most stable housekeeping in Chapter 2, Section 2.3.6. Both Fut8 and MGAT3 primers were designed using in Primer-BLAST (Ye, Coulouris *et al.*, 2012), and MGAT3 primers were kindly donated by Jenny Chik, and used previously in Chik *et al.* (2013). Primer sequences are listed in Table 3.1.

Table 3.1 qRT-PCR primer pair sequences

Gene Name	Symbol	Forward (5'-3')	Reverse (5'-3')	Primer Target	Product Size (bp)
Glycosyltransferases:					
Mannosyl (β-1,4-)-glycoprotein β-1,4-*N*-acetylglucosaminyl transferase	MGAT3	CGGGACCTGAACTACATCCG	GTCCAGCAGGTAGTGGAACC	Specific to MGAT3	150
Fucosyltransferase 8 (α(1,6) fucosyltransferase)	Fut8	TATGCTCACCAACCCCGAAC	TATAGGCCCGTCCTTCCCAA	transcript variant 1	136
Housekeeping gene:					
TATA-box binding protein	TBP	GCACCACTCCACTGTATCCC	GCTGCGGTACAATCCCAGAA	transcript variant 1	150

3.3 Results

The membrane glycosylation of adipocytes was compared to that of ADSCs from the same individuals (in three biological replicates) as a model to explore glycosylation changes during *in vivo* adipogenesis. As a model of *in vitro* adipogenesis, ADSCs were adipogenically differentiated in cell culture and the membrane glycome of the resulting cells (dADSCs) were compared with mature adipocytes and the progenitor ADSCs.

3.3.1 Evaluating the use of Triton X-114 for membrane protein enrichment

Adipocytes from one individual were used as a representative sample to evaluate the use of Triton X-114 in the enrichment of membrane proteins in comparison with the initial ultracentrifugation to pelleted membranes. With the same starting volume of 5 ml adipocytes, quantitation of protein concentration by Bradford assay indicated more protein was recovered in the final acetone precipitated Triton X-114 treated sample (6.3 µg/µl compared to 3.3 µg/µl in the ultracentrifuged membranes). SDS-PAGE separation of equal protein amounts revealed differences in the banding patterns of the two samples, indicating an enrichment of different proteins with Triton X-114 phase separation (Figure 3.1a). Although the gel staining indicated un-equal loading of protein amount as calculated by Bradford Assay, 10 µg of each sample was spotted onto PVDF membrane and direct blue 71 staining intensity indicated equal protein amounts. The *N*-linked glycan total ion mass profiles from each spot however showed up to five times more glycans in the Triton X-114 fraction (Figure 3.1b). Qualitatively, the same ion signals are present in both profiles however the samples differ quantitatively in the relative ratios of each signal. Overall the results from this comparison indicate the inclusion of Triton X-114 in the membrane preparation procedure enriches glycoproteins from the crude membrane preparation. The following data in this chapter were therefore generated from Triton X-114 enriched membrane proteins.

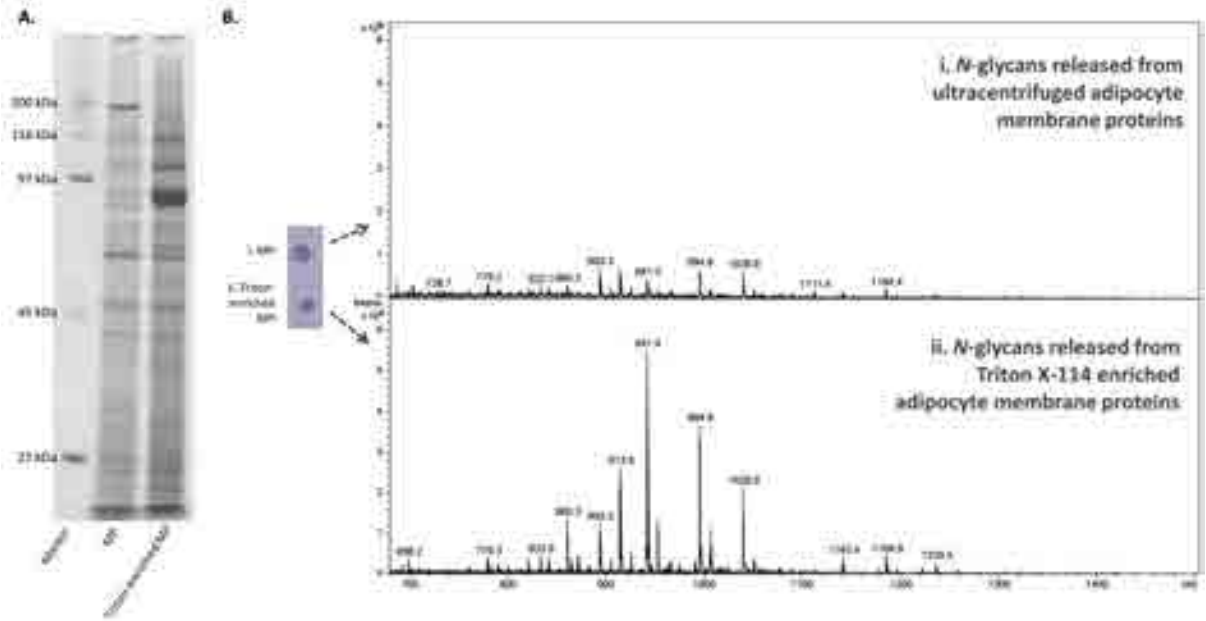


Figure 3.1 Membrane proteins (MP) extracted with and without Triton X-114 were compared visually by staining SDS-PAGE separated proteins (**A**), and *N*-linked glycan average mass profiling of membrane glycoproteins (**B**).

3.3.2 N- *and* O-*glycomic profiling of ADSCs over cell culture passages*

To determine whether cell culturing conditions alter cell surface glycosylation, the *N*- and *O*-glycomes of ADSCs were profiled from passage one through to six, with the original sample SVF cells collected from the lipo-aspirate compared. The SVF is a heterogeneous mixture of cells collected following the adipose digestion procedure and showed a very heterogeneous profile of *N*-linked structures (Figure 3.2). Several of these glycans decrease after the SVF cells were placed in culture and a homogeneous ADSC population was established (marked by shaded panels in Figure 3.2). The decreased ion signals corresponded to glycan compositions with more than one fucose residue, indicating the presence of outer *N*-linked glycan arm fucosylation which is characteristic of glycosylation patterns in red blood cells, suggesting that these cells are present in the SVF and do not adhere to the culture flask as are typical of the stem cells. The average mass profiles from passages one, three and six of ADSCs indicate that the *N*-glycans on the membrane proteins are qualitatively, but not quantitatively, consistent over cell culture passaging.

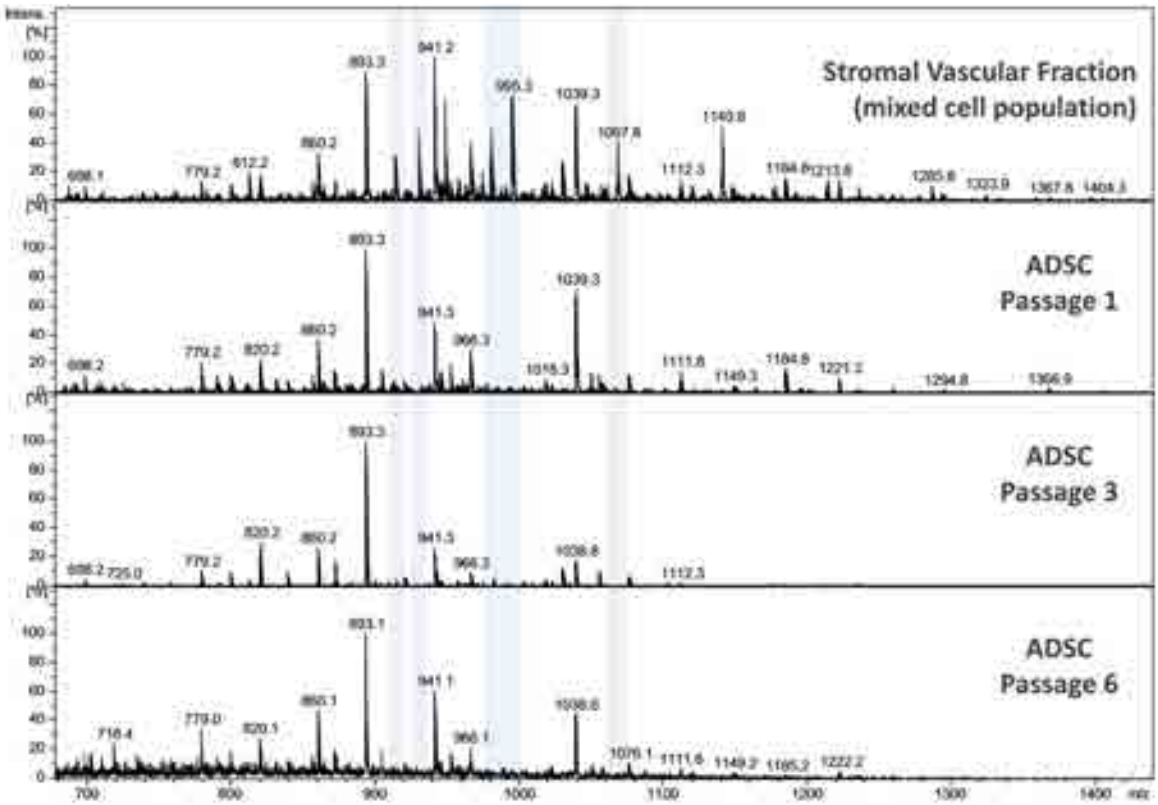


Figure 3.2 *N*-linked glycan profiles of SVF cells and ADSCs from passage one, three and six (*m/z* 680 to 1500). Shaded panels highlight ion signals that decrease in intensity after a homogeneous cell population is established following the seeding of the SVF.

In the *O*-linked glycan profiles, a limited number of ion signals corresponding to glycan compositions were observed and are marked with an arrow in Figure 3.3. *O*-linked glycans were also found to be qualitatively, but not quantitatively consistent over cell culture passages, with the ratio of structures with *m/z* 675.0 and 966.1 appearing similar over passages but changing in relative abundance comparison to the major structure with *m/z* 471.9.

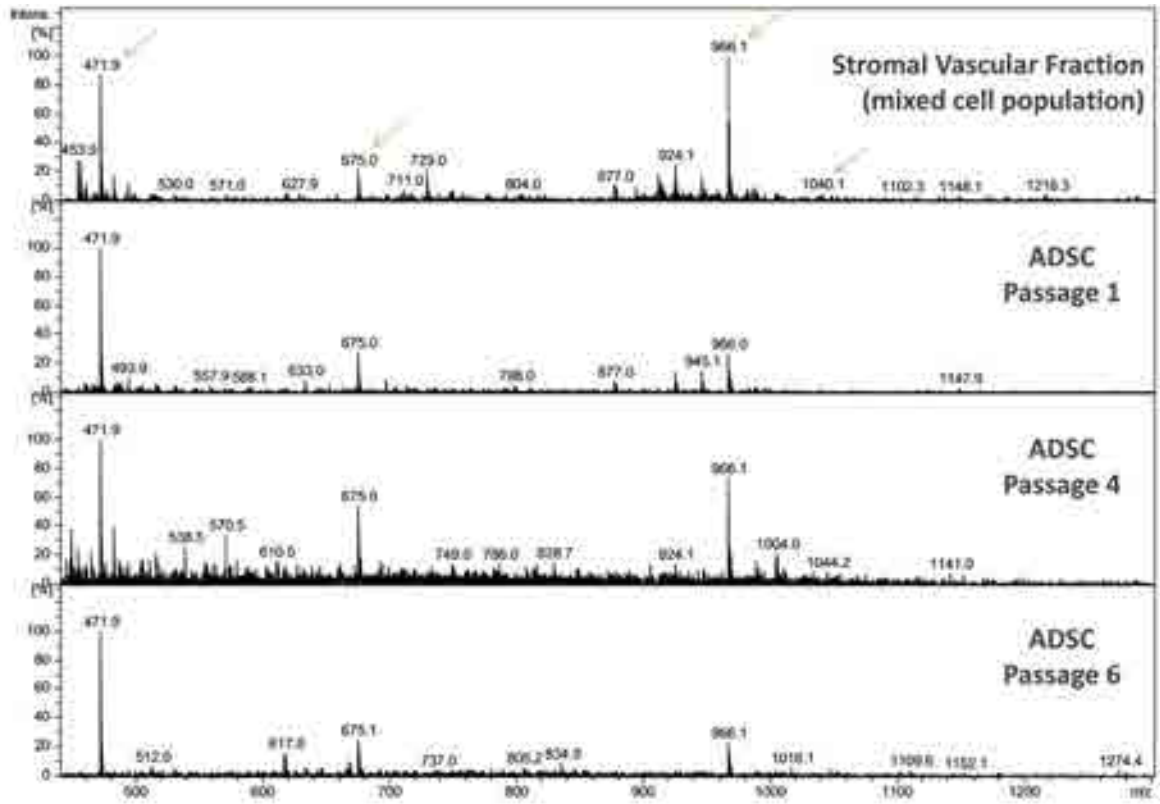


Figure 3.3 *O*-linked glycan profiles of SVF cells and ADSCs from passage one, four and six (*m/z* 440 - 1300). *O*-linked glycans were found to be qualitatively consistent over cell culture passages. Ion signals corresponding to true structures are marked with an arrow.

3.3.3 Optimisation of glycan release method

Bovine fetuin was used as a standard glycoprotein to determine whether reduction and alkylation prior to spotting on PVDF membrane resulted in a higher yield of glycans. The principle of spotting glycoproteins to PVDF is to take advantage of the hydrophobic proteins binding to the membrane and exposing the hydrophilic glycans to enzyme digestion and release. The average MS profile of fetuin that was reduced and alkylated prior to spotting to PVDF indicates that the extra reduction and alkylation steps exposed more glycans to PNGase F digestion (Figure 3.4), with the signal intensity is almost doubled for the glycans released from reduced and alkylated fetuin. In terms of downstream MS and data analysis, this has two implications: the increase in signal intensity of glycans from reduced and alkylated glycoproteins increased the signal to noise ratio (for example, *m/z* $[1148.4]^{2-}$ is newly identified in Figure 3.4b) and this resulted in more informative MS/MS fragmentation of these ions to aid in manual structural characterisation (discussed further in Section 3.3.4) and area-under-curve quantitation. Secondly, a smaller amount of starting material could be used while still acquiring relatively informative MS spectra. An extra cleanup step using a carbon SPE tip was also included to ensure an even better MS signal to noise ratio. The results described in the following sections were therefore obtained from *N*- and *O*-linked glycans released from reduced and alkylated adipocyte, ADSC and dADSC membrane glycoproteins.

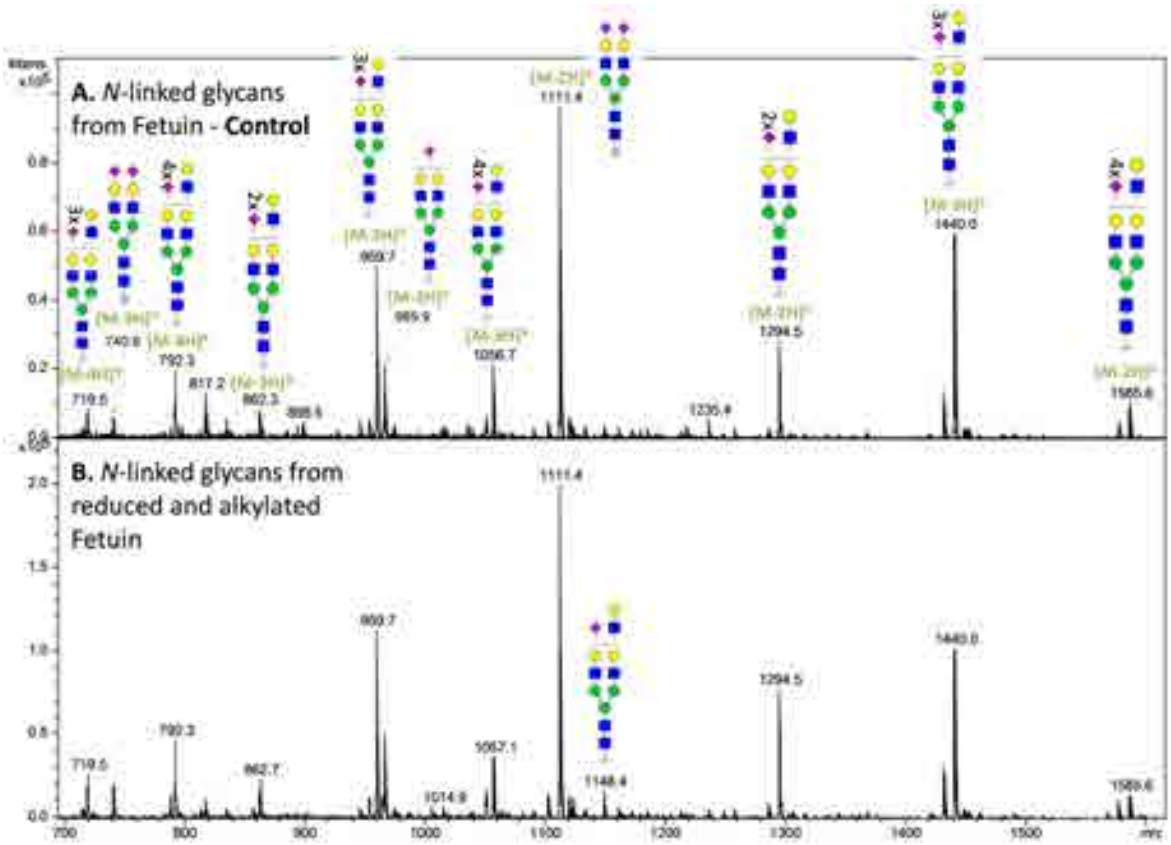


Figure 3.4 Representative average MS profiles (*m/z* 700 - 1600) of *N*-linked glycans released from fetuin (**A**) and fetuin glycoproteins reduced and alkylated prior to spotting on PVDF membrane (**B**). Ion charge states are labelled along with ion signals corresponding to glycan structures. Monosaccharide names corresponding to symbolic notation can be found in Figure 1.2.

3.3.4 PGC-LC ESI-MS/MS separation and identification of isomeric structures

The use of PGC chromatography facilitates the reproducible separation of isomeric structures and can provide important information about differential expression of isomers. As parent ion masses increase, the complexity of potential corresponding glycans also increases, along with the number of different compositional and sequence possibilities. However with the presence of different MS/MS diagnostic ions (Everest-Dass, Abrahams *et al.*, 2013; Harvey, Royle *et al.*, 2008) coupled with PGC retention time patterns (Pabst, Bondili *et al.*, 2007), it is possible to determine the topology of the structure. Two examples of this type of manual structural determination will be discussed in the next two sections. Structure numbers correspond to the total list of characterised glycans in Table S3.1 (Supplementary Data).

3.3.4.1 Identification of isomers with bisecting GlcNAc and lactosamine repeats

Four structures that correspond to *m/z* $[1177.5]^{2-}$ and have a composition of (Hex)3 (HexNAc)4 (Fuc)1 + (Man)3(GlcNAc)2 were detected only in the adipocyte glycome and eluted at different time points (Figure 3.5a). The MS/MS fragmentation profiles of each structure are shown in Figure 3.5b. The use of diagnostic ions, in particular the D ion, allowed the structural details of each isomer to be

determined. The D ion fragment provides information about the composition of the 6'arm of *N*-linked glycans. In bisecting GlcNAc structures, the D ion fragment is missing. Instead, there is a D-221 ion created from the loss of the bisecting GlcNAc from the D ion (Harvey, Royle *et al.*, 2008). For example, the D-221 ions in the MS/MS profiles of structures #87 and 88 (*m/z* [670.2] $^-$ and *m/z* [1035.5] $^-$ respectively; Figure 3.5b-1 & 3.5b-2) allowed the assignment of a bisecting GlcNAc at the inner most mannose of the *N*-linked glycan core. The D ion of structure #88 is different to #87, and the difference in mass corresponds to a composition of (Hex-HexNAc)2, which is used to discriminate between this structure branching on the 3'arm (structure #87) compared to the 6'arm (structure #88). The MS/MS profile of structure #88 also contained a dominant [M-221-18] ion (*m/z* [1017.4] $^-$ in Figure 3.5b-2), which in addition to the presence of a weak D-221 ion confirmed the presence of a bisecting GlcNAc (Harvey, 2005; Harvey, Royle *et al.*, 2008) and 6'arm branching. Of course, the composition of (Hex-HexNAc)2 can also correspond to poly-*N*-acetyllactosamine repeat. In instances where this composition corresponds to elongation and not branching, it can be differentiated by the presence of a relatively prominent F ion with *m/z* [789] $^-$ indicating the presence of a Hexose-HexNAc-Hexose-HexNAc chain (Harvey, Royle *et al.*, 2008). This ion was not present in the MS/MS of structures #87 and 88, but was present in the MS/MS of structures #89 and 90 (Figure 3.5b-3 & 3.5b-4). Together with the D ion, a poly-*N*-acetyllactosamine repeat could be assigned to the 3'arm in structure #89 and the 6'arm in structure #90. In all the isomers shown in Figure 3.5, the fucose residue could easily be assigned to the chitobiose core in all instances because of the presence of the prominent *m/z* [571] $^-$ ion corresponding to a reducing core GlcNAc with an additional GlcNAc and Fuc residue.

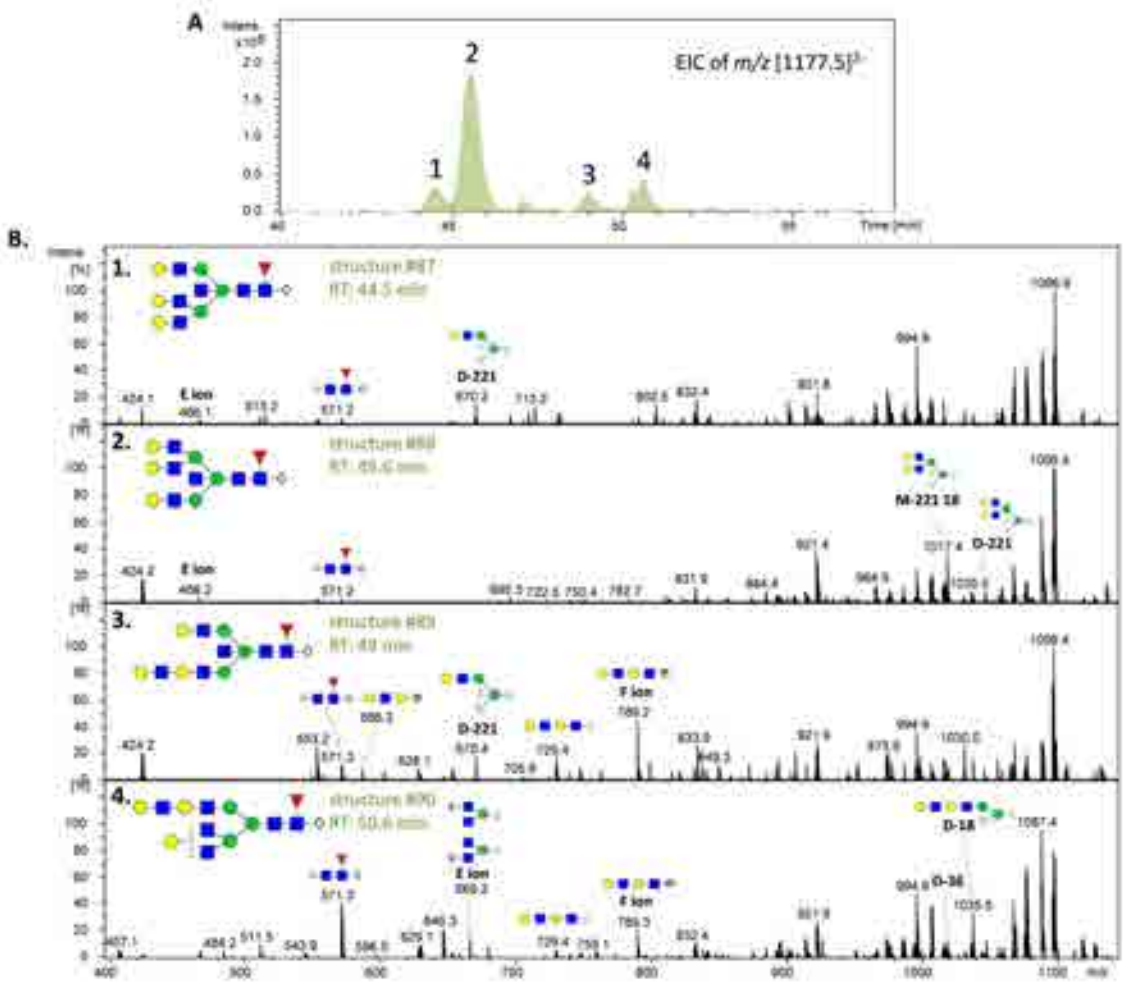


Figure 3.5 Four isomeric structures with *m/z* $[1177.5]^{2-}$ were separated using PGC-LC (**A**). A representative MS/MS fragmentation profile of each isomer is shown (**B**). The *m/z* range is zoomed to *m/z* 400 - 1140 as it contains the main diagnostic ions needed for structural characterisation of these structures. 'RT' = retention time.

3.3.4.2 Identification of sialic acid linkage isomers

The second example of using PGC separation and MS/MS fragmentation information to differentiate isomeric structures are three EIC peaks with *m/z* $[1038.9]^{2-}$ corresponding to a composition of (Hex)2 (HexNAc)2 (Fuc)1 (NeuAc)1 + (Man)3(GlcNAc)2 (Figure 3.6a) which were found in all three cell types. All three isomers are biantennary, mono-sialylated structures, readily distinguishable from one another by information provided by the D ion. For example, the prominent D ion with *m/z* $[979.4]^{-}$ seen in the MS/MS of structure #58 (shaded panel in Figure 3.6b-2) indicates that the sialic acid residue is on the 6'arm antennae. Conversely, the absence of this mass in the MS/MS of structures #57 and 59 indicates that the sialic acid is on the 3'arm in these structures (highlighted by broken arrows). The specific linkage type of the sialic acid residues was determined based on known PGC retention order of sialylated glycans where α2-6 linked sialylated structures elute approximately

eight minutes before α2-3 linked structures (Anugraham, Jacob *et al.*, 2014; Nakano, Saldanha *et al.*, 2011).

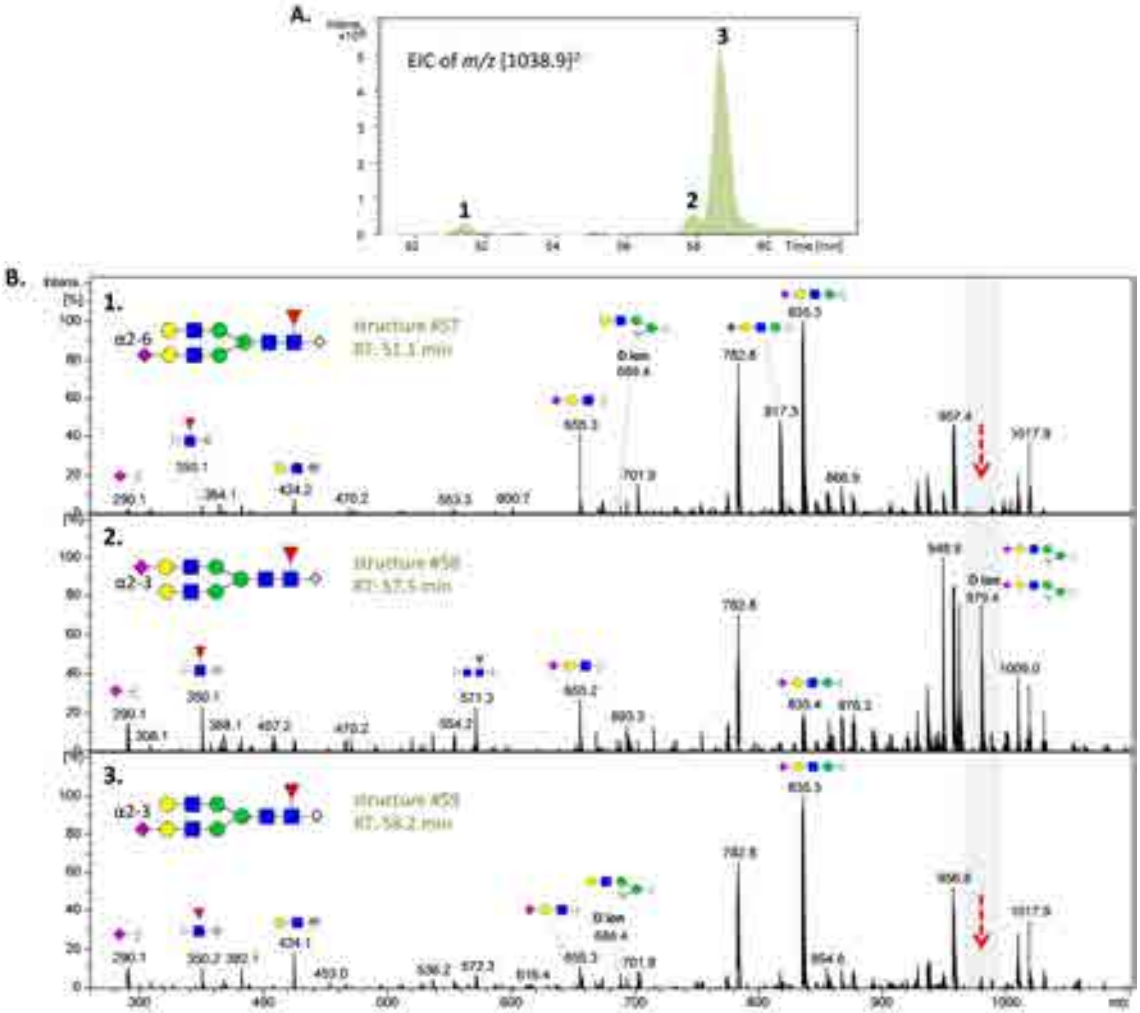


Figure 3.6 Representative EIC of *m/z* $[1038.9]^{2-}$ (from adipocytes) corresponding to a composition of (Hex)2 (HexNAc)2 (Fuc)1 (NeuAc)1 + (Man)3(GlcNAc)2 revealed three isomers corresponding to this mass (**A**). Representative MS/MS fragmentation spectra of the three isomers contain diagnostic ions that can be used to determine the topology of each structure (**B**). The shaded panel around *m/z* 979.4 highlights the presence of this diagnostic D-type fragmentation ion in structure #58, whilst the broken arrows indicate its absence in structures #57 and 59. 'RT' = retention time.

3.3.4.3 Identification of structures containing a terminal Gal-Gal epitope

Three structures with *m/z* $[974.9]^{2-}$, corresponding to the composition (Hex)3(HexNAc)2 (Fuc)1 + (Man)3(GlcNAc)2, were identified (structures #47, 48 and 49). Inspection of the MS/MS fragmentation data for the three isomers (Figure 3.7) revealed the first eluting structure to be a hybrid-type (structure #47; Figure 3.7a), however further structural details could not be determined for this structure due to the presence of ions indicating the possible, but not conclusive presence of a bisecting GlcNAc residue. For example, the fragment ion with *m/z* $[628.7]^{-}$ in Figure 3.7 could be a D-18 type ion (indicating the absence of a bisecting GlcNAc residue) or a D-221 type ion (indicating the

loss of a bisecting GlcNAc residue during MS/MS). The position and identity of the *N*-acetylhexosamine residue was therefore not assigned. MS/MS data for the second and third eluting isomers (structures #48 and 49; Figure 3.7b & c) contained evidence for the presence of a terminal Gal-Gal epitope in both. Using the D-ion in the MS/MS of structure #48 (*m/z* [850.1] $^-$ in Figure 3.7b), the Gal-Gal repeat was assigned to the 6'arm. The D and D-18 ions in structure #49 (*m/z* [688.2] $^-$ and [670.2] $^-$ in Figure 3.7c respectively) and also the B_2-ion (*m/z* [323.1] $^-$ in Figure 3.7c) confirmed the presence of the Gal-Gal repeat and its location on the 3'arm. Structures #48 and #50 were identified in adipocytes, ADSCs and dADSCs, while structure #49 was identified only in ADSCs and dADSCs.

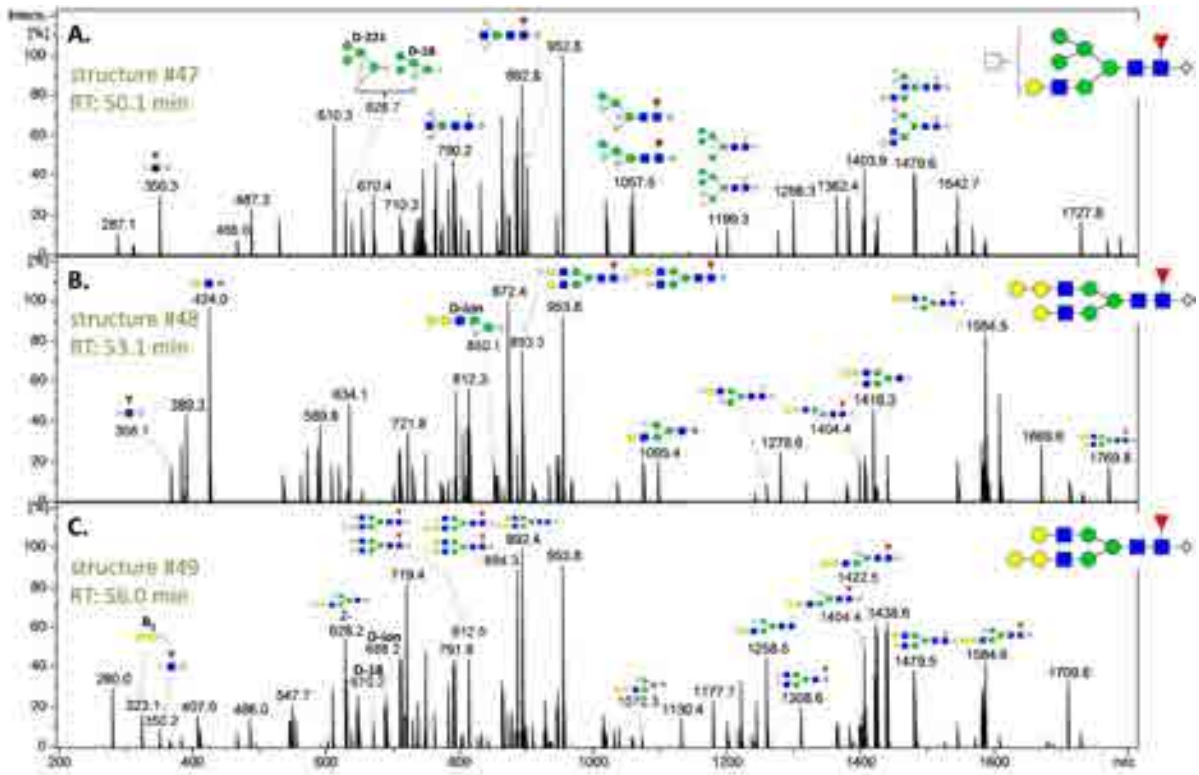


Figure 3.7 Three structures with *m/z* $[974.3]^{2-}$, corresponding to the composition (Hex)3 (HexNAc)2 (Fuc)1 + (Man)3(GlcNAc)2, were identified. The first isomer, structure #47 was identified as a hybrid structure; however the MS/MS fragmentation for this structure was inconclusive for full structural characterisation (**A**). The MS/MS fragmentation for structures #48 and 49 **(B & C)** however contained fragments that indicate they are arm isomers of a Gal-Gal epitope-containing structure. A representative MS/MS fragmentation profile is shown for each isomer. 'RT' = retention time; White square monosaccharide symbol = *N*-acetylhexosamine residue not able to be distinguished as *N*-acetylglucosamine or *N*-acetylgalactosamine.

As Gal-Gal epitopes typically occur with an α1-3 linked terminal galactose that is antigenic in humans (Galili, 2013) their presence in this study is noteworthy. However, the absence of an F ion of *m/z* [586.0] $^-$ (Figure 3.8b-c) which corresponds to an antennae composition of αGal-Gal (Harvey, Royle *et al.*, 2008) indicates that the terminal galactose residue may not be α-linked. To determine whether the terminal galactose of the Gal-Gal epitope on structures #48 and 49 was α- or β-linked, the

released *N*-linked glycans from adipocytes and ADSCs were digested with β1-3,4 galactosidase (BTG). Due to the low abundance of the structures, only the first isomer (structure #48) was visible in the undigested sample (Figure 3.8a; eluting at 51 min). The MS/MS fragmentation of this peak was confirmed as structure #48 by the MS/MS spectrum (data not shown). After BTG digestion, the peak corresponding to structure #48 was absent, with the fully digested product *m/z* [731.3] $^{-}$ increased in intensity (Figure 3.8b). The loss of the three galactoses indicated that the terminal galactose was most probably β-linked in either β1-3 or β1-4 linkage. The expression of a Gal-β1-4-gal terminal has only recently been reported in human melanoma cells, with nothing previously known about its possible expression in humans (Hoja-Lukowicz, Link-Lenczowski et al., 2013).

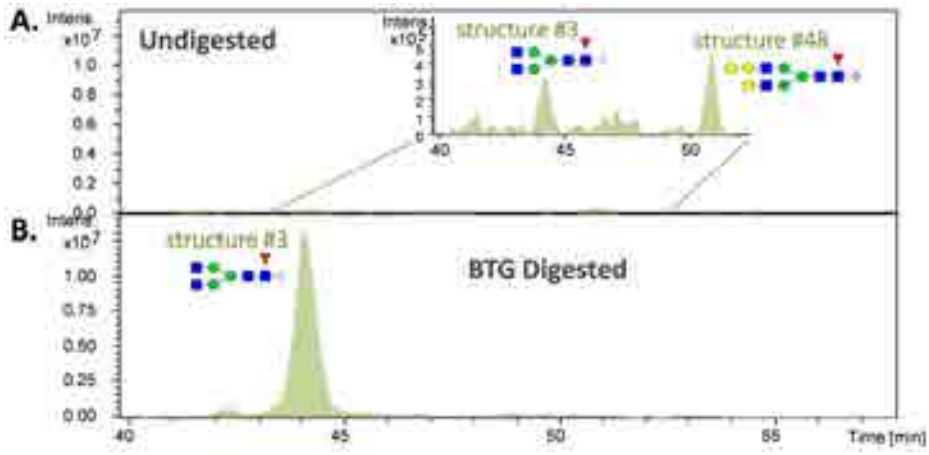


Figure 3.8 EIC of *m/z* $[731.3]^{1-}$ and $[974.9]^{2-}$ in undigested (**A**) and BTG digested *N*-linked glycans (**B**). Shown in the top panel is structure #48, eluting at 51 min. The lower panel shows the profile of the BTG treated sample with structure #48 no longer present, and with the fully digested product *m/z* $[731.3]^{1-}$ increased in intensity.

3.3.5 Comparison of ADSCs and adipocytes from the same individual: a model of in vivo *differentiation*

N- and *O*-linked glycans were released from ADSCs and adipocytes in order to compare *in vivo* adipogenesis products with their progenitor stem cells. This analysis was carried out in biological triplicate. Qualitative comparison of *N*-linked glycan profiles from ADSCs and adipocytes revealed a number of structural types present in both samples, with the exception of three ion signals with *m/z* $[913.9]^{2-}$, $[994.9]^{2-}$ and $[1140.9]^{2-}$ which dominate the adipocyte profile (marked by shaded panels in Figure 3.9) and are not seen in the average MS profile in the stem cells (highlighted by broken arrows in Figure 3.9). MS/MS interrogation revealed that these masses correspond to the five different bisecting GlcNAc-type structures (structure #34, #35, #50, #83 and #84) shown in Figure 3.9 and also listed in Table S3.1 (Supplementary Data).

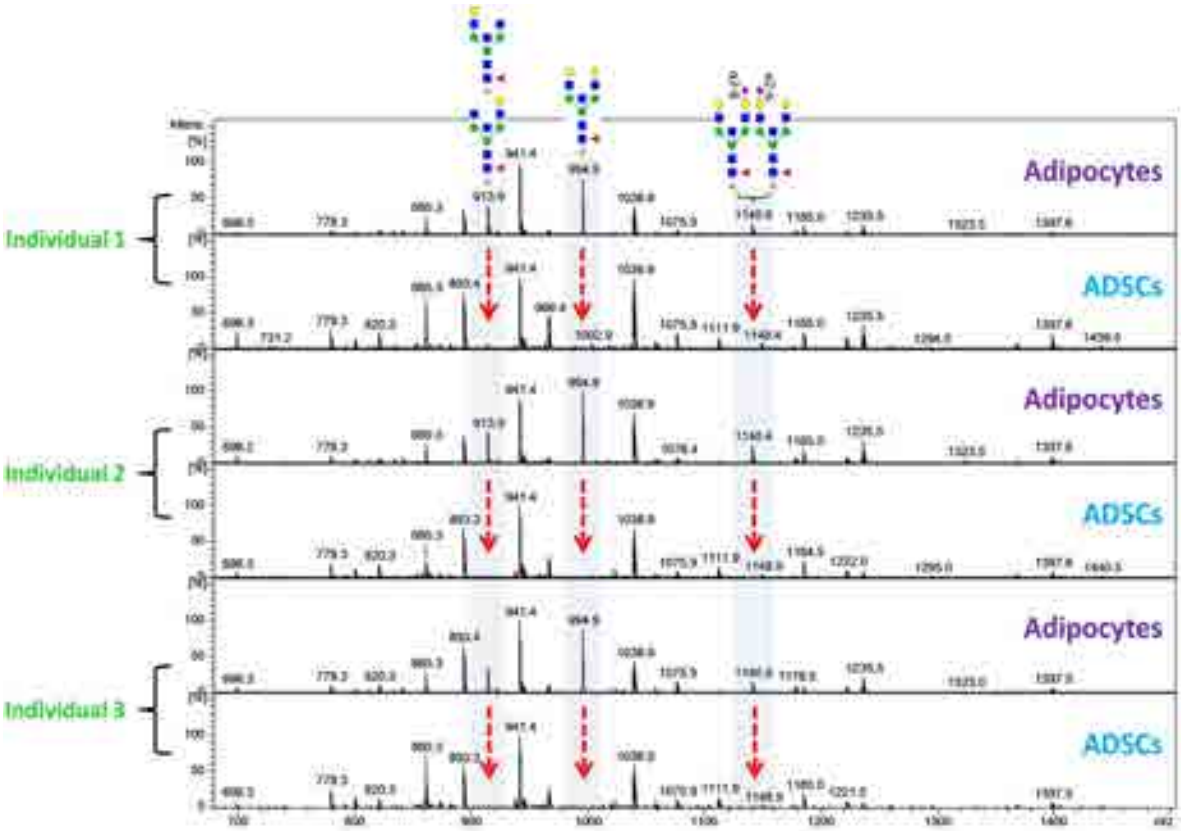


Figure 3.9 Average MS profiles of released *N*-linked glycans from adipocytes and ADSCs (*m/z* 680 to 1500). *N*-linked glycan profiling of paired adipocytes and ADSCs from the same individual was conducted on three biological replicates. The major differences seen in the shaded panels contain the three ion signals with *m/z* $[913.9]^{2-}$, $[994.9]^{2-}$ and $[1140.9]^{2-}$, corresponding to five bisecting GlcNAc-type structures that dominate the adipocyte profile. The absence of these structures in ADSCs at the average MS profile view is marked by broken arrows.

Comparison of *O*-linked glycans between ADSCs and adipocytes revealed similarities in the ion signals present (Figure 3.10). Differences in the intensities of these signals however indicate quantitative differences in the *O*-glycomes. Further in-depth analysis of *N*- and *O*-linked glycans from these samples will be further discussed in Sections 3.3.7 and 3.3.10.

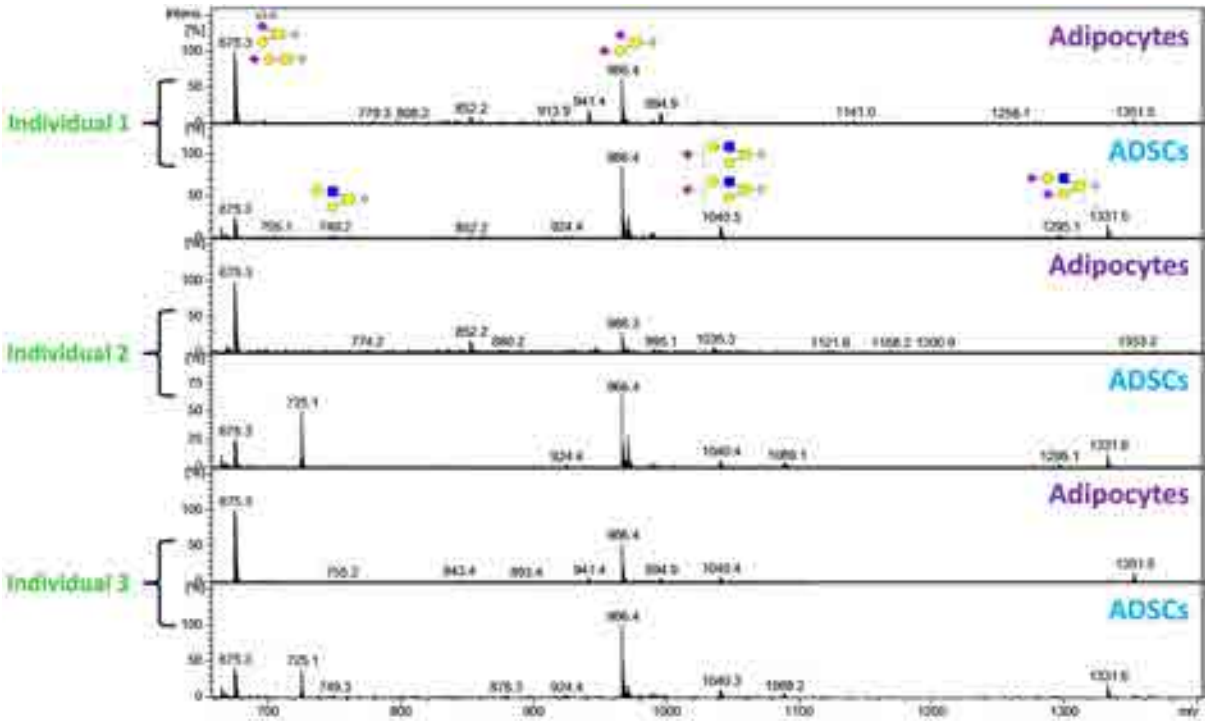


Figure 3.10 Average MS profiles of released *O*-linked glycans from adipocytes and ADSCs (*m/z* 650 - 1400). Paired *O*-linked glycan profiling of adipocytes and ADSCs from the same individual was conducted on three biological replicates. Ion signals corresponding to *O*-linked glycans are labelled with the corresponding structure/s.

3.3.6 Comparison of ADSCs and dADSCs from the same individual: a model of in vitro *differentiation*

N- and *O*-linked glycans were released from ADSCs and 30 day differentiated dADSCs (based on optimised large-scale differentiation conditions as described in Chapter 2, Section 2.2.6) in order to compare *in vitro* adipogenesis products with their progenitor stem cells. The MS average mass profiles of *N*-linked glycans from ADSCs and dADSCs are shown in Figure 3.11a & c, and appear qualitatively similar to each other. The three ion signals with *m/z* $[913.9]^{2-}$, $[994.9]^{2-}$ and $[1140.9]^{2-}$ that strongly dominate the adipocyte *N*-linked glycan profile (Figure 3.11c and first seen in Figure 3.9), and that are missing in the ADSCs (Figure 3.11a), are absent in the dADSC profile (Figure 3.11b) when viewed at the summed MS level. Also evident in the average MS is the relatively high abundance of high mannose structures in all samples, especially the ion signal with *m/z* $[941.4]^{2-}$, corresponding to a composition of (Hex)6 + (Man)3(GlcNAc)2 (Figure 3.11). More structures that are not seen in the summed MS profiles were identified and are shown in Table S3.1 (Supplementary data). A further in-depth qualitative and quantitative analysis of these samples will be further discussed in Sections 3.3.7.

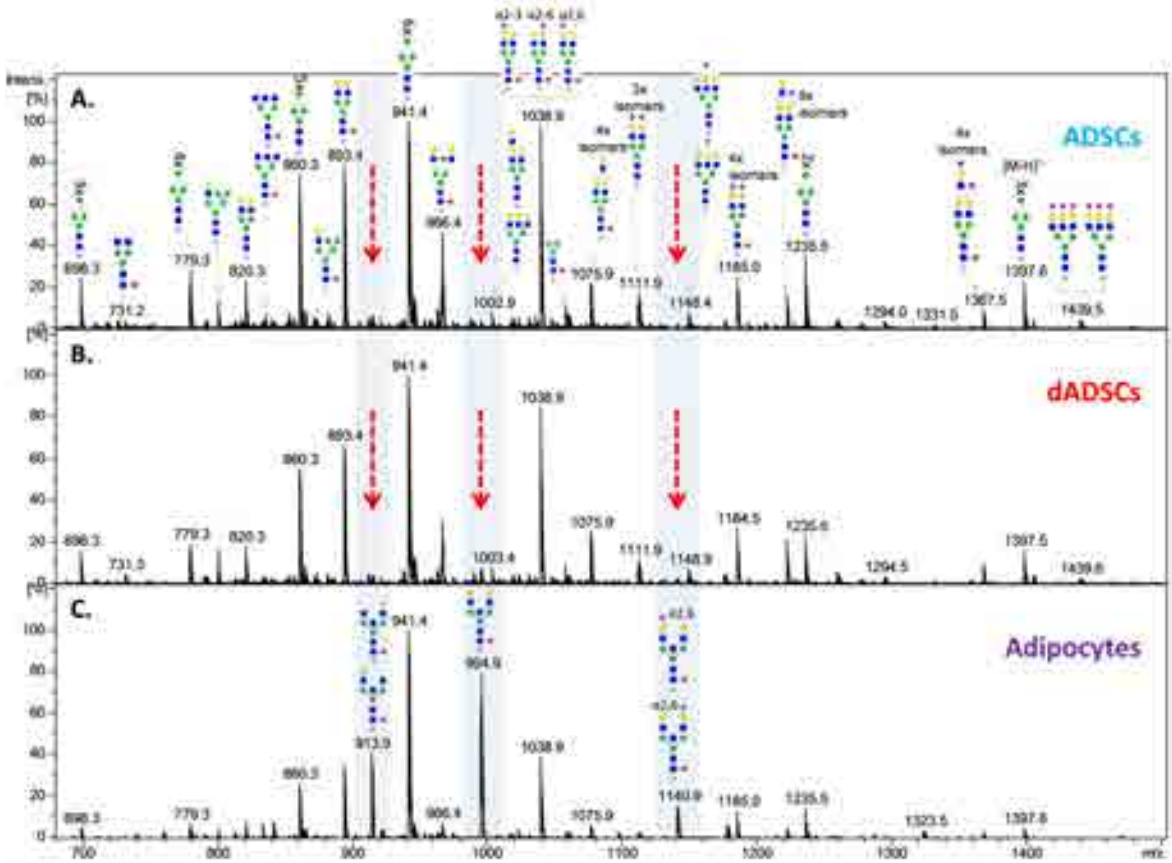


Figure 3.11 Representative average *N*-linked glycan MS profiles ADSCs (**A**), dADSCs (**B**) and adipocytes (**C**) from the same individual *(m/z* 680 - 1500). The shaded panels highlight three ion signals with *m/z* $[913.9]^{2-}$, $[994.9]^{2-}$ and $[1140.9]^{2-}$ present in high abundance in adipocytes, and present in very low levels not visible in the average MS in the ADSCs and dADSCs (highlighted by the broken arrows). In some instances, more than one structure has been assigned to a mass due to the presence of isomeric structures.

3.3.7 N-*glycome of ADSCs, dADSCs and adipocytes:* in vitro *vs.* in vivo *adipogenesis*

To determine the exact structures that are changing between ADSCs, dADSCs and adipocytes, a detailed and comprehensive interrogation of the *N*-glycomic profiles of these cells from three individuals was conducted using diagnostic ions from MS/MS fragmentation data and PGC retention patterns as described in the Section 3.3.4, in addition to knowledge of potential glycan biosynthetic pathway products. The analysis identified a total of 130 *N*-linked glycan structures across the three cell types. These structures are summarised in Table S3.1 (Supplementary data), together with their relative abundances as calculated by area under the curve from the EIC of each mass. The 130 structures were elucidated from 61 separate compositions, due to the number of isomeric structures. Twenty three further potential glycan compositions were indicated (see Supplementary Table S3.2) but these were too low in abundance to integrate peak areas or obtain fragmentation data and were not included in the analysis.

A qualitative analysis of *N*-linked structures (*n* = 130) identified an average of 110.3 ± 4.2 structures in adipocytes, 84.7 ± 3.5 in ADSCs and 75.3 ± 12.3 in dADSCs. Comparison using a proportional Venn diagram (Figure 3.12) revealed 79 structures common to the three cell types, with 31 structures found only in adipocytes. The high similarity between ADSCs and dADSCs was clearly visualised in the Venn diagram where the two samples are displayed as overlapping (dashed) circles. Both ADSCs and dADSCs contained ten hybrid and complex-type structures not seen in adipocytes. It is important to note here that the higher number of structures in adipocytes may be due to a larger sample amount used for analysis.

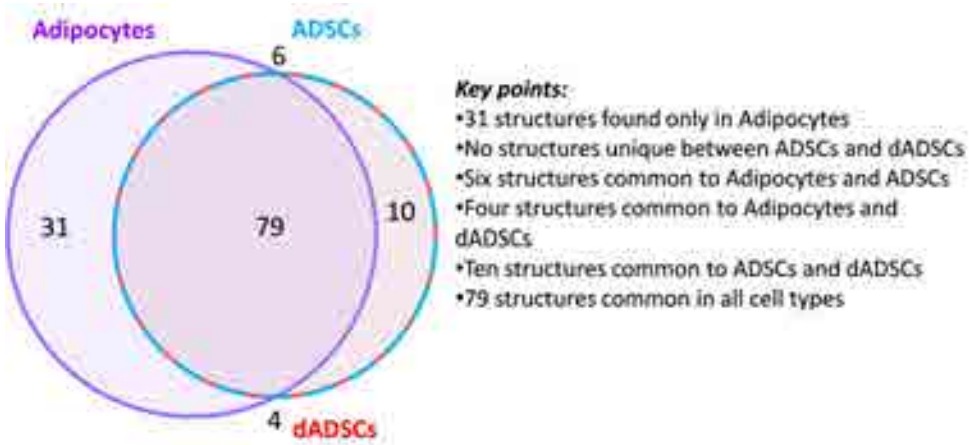


Figure 3.12 Proportional Venn diagram comparison of *N*-linked glycans in ADSCs, dADSCs and adipocytes. ADSCs and dADSCs show high similarity in the *N*-linked glycan constituents with the two samples in fact overlapping and each containing no unique structures.

For relative quantitative analysis, the 130 identified *N*-linked glycans were grouped according to broad structural class types (Figure 3.13). The overall relative percentage of each structural type was compared between ADSCs, dADSCs and adipocytes. High Mannose and complex-type structures were the major structures with some hybrid (<10 %) and pauci mannose-types (<5 %) (Figure 3.13a).

3.3.7.1 High Mannose Structures

High mannose-type structures had the lowest average percentage abundance in adipocytes (29.6 % ± 1.8), followed by ADSCs (36.9 % ± 3.1), and dADSCs (38.7 % ± 7.5) (Figure 3.13a). Despite these differences in the overall percentage, not all high mannose structures individually follow this trend. Expression levels of the most abundant Man9 structure with *m/z* $[941.3]^{2-}$ (structure #39) were constant between cell types at an average of 15.3 % ± 2.5 of the total *N*-glycan structures (Figure 3.13b). With the exception of two Man4 isomers (structures #63 & 64, with *m/z* $[1073.5]^{-}$) all high mannose structure were common to all three cell types.

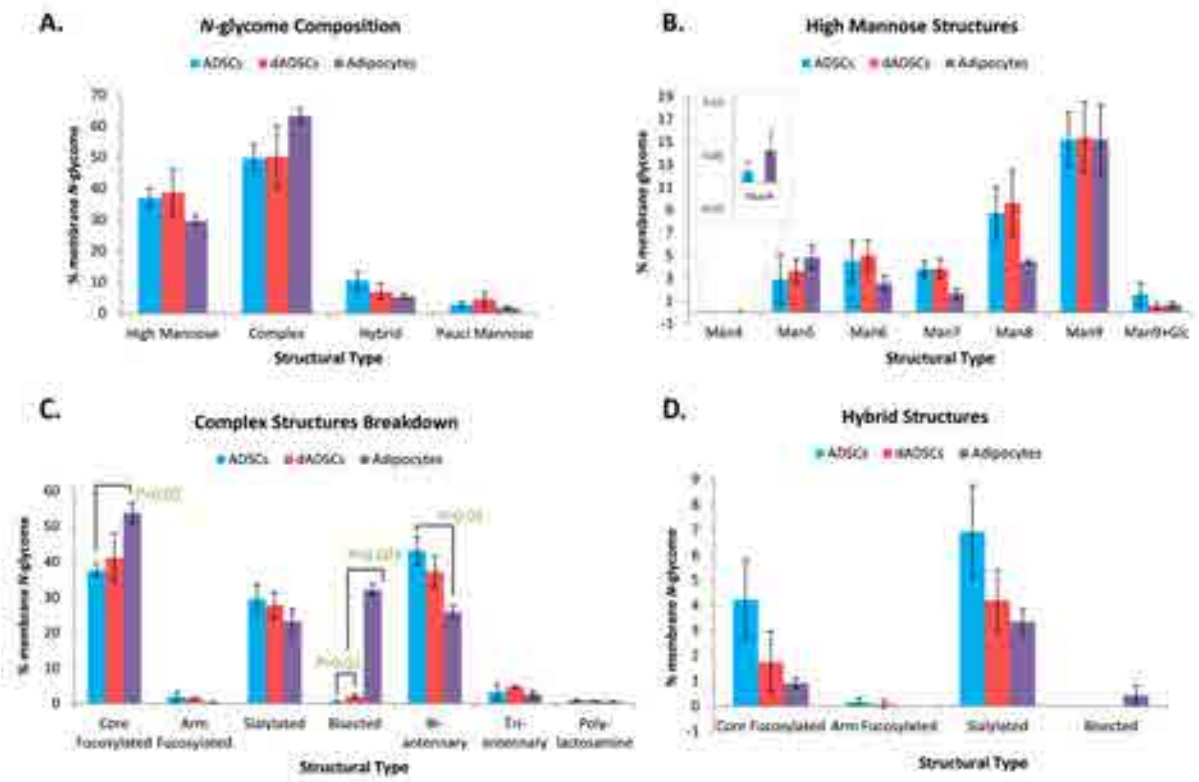


Figure 3.13 A comparison of structural classes in the *N*-glycomes of ADSCs, dADSCs and adipocytes. *N*-linked glycans were grouped according to broad structural classes; high mannose, complex, hybrid, and pauci mannose-type structures (**A**) and the percentage composition of each structural class was statistically compared. Each class type was further classified according to various structural characteristics (**B** - **D**). It is important to note that some structural types ay be present across several categories. Student's *t*-tests were performed comparing dADSCs and adipocytes with ADSCs and those with $P < 0.05$ are shown.

3.3.7.2 Complex Structures

Classification into different complex structural sub-classes (Figure 3.13c), revealed that core fucosylated ($P = 0.02$) and bisecting GlcNAc structures ($P < 0.001$) were significantly up-regulated in adipocytes compared to the stem cells and are mainly responsible of the overall increase of complex-type structures in this cell type. In dADSCs, the increases seen in core fucosylation compared to ADSCs were not found to be significant. The increase in expression of bisecting GlcNAc structures in dADSCs however was significantly higher than in their ADSC progenitors, although still at low abundance.

The further analysis of complex-type structures (Figure 3.13c) also revealed expression levels of biantennary structures were significantly higher in ADSCs compared to adipocytes ($P = 0.03$). The main structure responsible for the overall expression of biantennary structures was a biantennary, core fucosylated, mono-sialylated glycan with *m/z* $[1038.9]^{2-}$ (structure #59). The elucidation of this abundant isomeric structure as described in Section 3.3.4.2, characterised the structure as having α2-3 linked sialic acid on the 3'arm.

Poly-lactosamine-containing structures (structures #52, 68, 69, 74, 89 and 90) were also identified amongst the complex-type structures. These structures can be identified by the presence of an F ion with *m/z* [789] $^{-}$ corresponding to the composition (Hex-HexNAc)2 as described in Section 3.3.4.1. The abundance of poly-*N*-acetyllactosamine-containing structures was highest in ADSCs (0.9 ± 0.2 %), followed by dADSCs (0.7 ± 0.1 %) and adipocytes (0.6 ± 0.1 %) (Figure 3.13c).

3.3.7.3 Hybrid and Pauci Mannose Structures

Hybrid-type structures were found present in all cells, with ADSCs expressing the highest average percentage (10.5 ± 2.8 %), followed by dADSCs (6.9 ± 2.5 %) and adipocytes (5.3 ± 0.9 %) (Figure 3.13a). Hybrid structures containing a bisecting GlcNAc were found only in adipocytes at a relative abundance of 0.4 ± 0.4 %.

Pauci mannose structures were observed in all cell types, with highest expression in dADSCs at 4.6 ± 2.4 %, and lowest in adipocytes (1.8 ± 0.4 %) (Figure 3.13a). These expression differences were not found to be significant.

3.3.8 Bisecting GlcNAc structures are characteristic of adipogenesis

Overall, bisecting GlcNAc complex-type structures make up approximately 32.1 ± 1.5 % of the adipocyte *N*-glycome; a stark difference in expression when compared to the ADSCs and dADSCs (0.4 ± 0.4 % and 1.9 ± 0.7 % respectively). The increased expression of bisecting GlcNAc structures compared to their progenitor ADSCs was found to be significant for both adipocytes ($P < 0.001$) and although, much lower, for the dADSCs ($P = 0.02$). The difference in bisecting GlcNAc abundance is exemplified by the expression pattern of structure #50, a di-galactosylated, bisecting GlcNAc structure with *m/z* $[994.9]^{2-}$. When normalised to structure #39 with *m/z* $[941.4]^{2-}$ (Figure 3.14a) whose expression is constant between cell types, the expression of structure #50 clearly decreases in the ADSCs and dADSCs. The structure was characterised by the loss of the bisecting GlcNAc residue from the D ion as seen in the MS/MS fragmentation spectra of this structure (Ion signal with *m/z* [670.3] $^{-}$ in Figure 3.14b). The core fucose residue can also be assigned due to the presence of two diagnostic ions: *m/z* [350.1] $^{-}$ and [571.3] $^{-}$

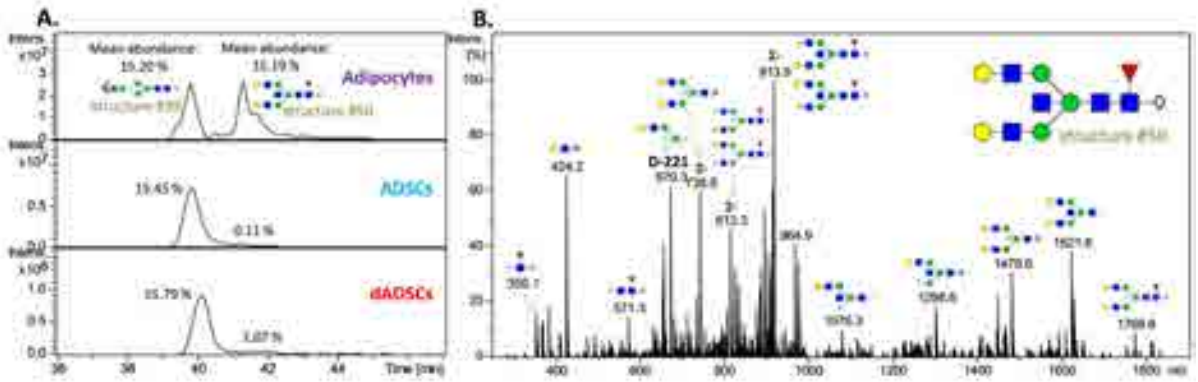


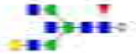
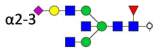
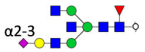
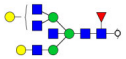
Figure 3.14 Representative expression profile (**A**) and MS/MS of structure #50 with *m/z* $[994.9]^{2-}$ (**B**). The expression profile of structure #50 is normalised to the high mannose structure with *m/z* $[941.4]^{2-}$ (structure #39). Doubly charged ions are marked with "2-".

Investigating bisecting GlcNAc structures more specifically, among the 31 structures that were only detected in adipocytes, 13 were bisecting GlcNAc-type structures (Table 3.2a). These 13 bisecting structures account for 3.8 ± 0.3 % of the adipocyte *N*-glycome, and include among them one of the isomers with *m/z* $[913.9]^{2-}$ (structure #34) that dominates the adipocyte average mass spectrum (Figure 3.9). Three bisecting GlcNAc structures were also found common between adipocytes and dADSCs (structures #20, 83 and 35; Table 3.2b). These structures were not detected in the same number of cells of the progenitor ADSCs, indicating that their expression is induced by *in vitro* differentiation. Amongst them is the second isomer corresponding to *m/z* $[913.9]^{2-}$ (structure #35) that is seen in both adipocytes and dADSCs.

Four related biantennary bisecting GlcNAc structures (structures #23, 36, 50 and 84; Table 3.2c) were also found within the 79 structures common to all three cell types. The average relative abundance of these structures was very low in ADSCs, and their expression was increased in adipocytes and dADSCs, indicating the expression of bisecting GlcNAc structures may be characteristic of adipogenesis. The fold-changes in expression for adipogenic products compared to ADSCs for these structures are detailed in Table 3.2c and show the increase in bisecting structures upon *in vitro* differentiation of stem cells, although to a much lesser extent than that present in mature adipocytes.

Table 3.2 Bisecting GlcNAc structures identified in the present study

A. Bisecting GlcNAc structures detected only in mature adipocytes:

Structure #	Detected *m/z*	Theoretical Mass (Da)	Retention time (min)	Composition	Relative abundance (Mean % ± SD)			Proposed Structure
					ADSCs	dADSCs	Adipocytes	
4	739 3]^{2}	1479 5	30 6	(Hex)1 (HexNAc)2 + (Man)3(G cNAc)2			0 06 ± 0 04	
7	759 8]^{2}	1520 6	31 9	(HexNAc)3 + (Man)3(G cNAc)2			0 39 ± 0 16	
17	820 3]^{2}	1787 7	33 0	(Hex)2 (HexNAc)2 + (Man)3(G cNAc)2			0 22 ± 0 10	
18	820 3]^{2}	1787 7	36 4	(Hex)2 (HexNAc)2 + (Man)3(G cNAc)2			0 58 ± 0 28	
34	913 8]^{2}	1828 7	37 7	(Hex)1 (HexNAc)3 (Fuc)1 + (Man)3(G cNAc)2			0 05 ± 0 04	
61	1059 4]^{2}	2119 8	46 1	(Hex)1 (HexNAc)3 (Fuc)1 (NeuAc)1 + (Man)3(G cNAc)2			0 36 ± 0 12	
62	1059 4]^{2}	2119 8	47 4	(Hex)1 (HexNAc)3 (Fuc)1 (NeuAc)1 + (Man)3(G cNAc)2			0 09 ± 0 04	
72	1096 4]^{2}	2396 9	43 2	(Hex)2 (HexNAc)4 (Fuc)1 + (Man)3(G cNAc)2			0 14 ± 0 03	
73	1096 4]^{2}	2396 9	45 1	(Hex)2 (HexNAc)4 (Fuc)1 + (Man)3(G cNAc)2			0 19 ± 0 03	
87	1177 4]^{2}	2355 9	44 5	(Hex)3 (HexNAc)4 (Fuc)1 + (Man)3(G cNAc)2			0 17 ± 0 03	
88	1177 4]^{2}	2355 9	45 6	(Hex)3 (HexNAc)4 (Fuc)1 + (Man)3(G cNAc)2			1 194 ± 0 053	
89	1177 4]^{2}	2355 9	49 0	(Hex)3 (HexNAc)4 (Fuc)1 + (Man)3(G cNAc)2			0 12 ± 0 03	

113	1286 1]^{2}	2572 9	56 7	(Hex)2 (HexNAc)3 (Fuc)1 (NeuAc)2 + (Man)3(G cNAc)2			0 28 ± 0 06	

B. Bisecting GlcNAc structures newly expressed upon *in vitro* differentiation – detected in both dADSCs and mature adipocytes:

Structure #	Detected *m/z*	Theoretical Mass (Da)	Retention time (min)	Composition	Relative abundance (Mean % ± SD)			Proposed Structure
					ADSCs	dADSCs	Adipocytes	
20	832 8]^{2}	1666 6	38 4	(HexNAc)3 (Fuc)1 + (Man)3(G cNAc)2		0 09 ± 0 06	1 09 ± 0 33	
83	1140 5]^{2}	2281 8	48 2	(Hex)2 (HexNAc)3 (Fuc)1 (NeuAc)1 + (Man)3(G cNAc)2		0 08 ± 0 09	1 18 ± 0 04	
35	913 8]^{2}	1828 7	40 1	(Hex)1 (HexNAc)3 (Fuc)1 + (Man)3(G cNAc)2		0 32 ± 0 27	6 51 ± 1 03	

C. Bisecting GlcNAc structures detected in ADSCs, dADSCs and mature adipocytes:

Structure #	Detected *m/z*	Theoretical Mass (Da)	Retention time (min)	Composition	Relative abundance (Mean % ± SD)			Proposed Structure	dADSCs fold change	Adipocytes fold change
					ADSCs	dADSCs	Adipocytes			
23	840 8]^{2}	1828 7	34 6	(Hex)1 (HexNAc)3 + (Man)3(G cNAc)2	0 02 ± 0 04	0 08 ± 0 06	1 63 ± 0 29		**3.5**	**74.77 ***
36	921 5]^{2}	1844 7	36 7	(Hex)2 (HexNAc)3 + (Man)3(G cNAc)2	0 03 ± 0 01	0 15 ± 0 03	0 95 ± 0 27		**4.42 ***	**27.72 ***
50	994 9]^{2}	1990 7	41 2	(Hex)2 (HexNAc)3 (Fuc)1 + (Man)3(G cNAc)2	0 10 ± 0 11	1 04 ± 0 22	15 23 ± 0 01		**9.78 ***	**142.9 ***
84	1140 5]^{2}	2281 8	48 9	(Hex)2 (HexNAc)3 (Fuc)1 (NeuAc)1 + (Man)3(G cNAc)2	0 02 ± 0 03	0 14 ± 0 04	2 45 ± 0 70		**9.29 ***	**166.31 ***

* = $P < 0.05$

- = below detection

3.3.9 Cell-specific markers

Several studies investigating the membrane glycome of stem cells and their differentiated progeny have done so with the aim of elucidating a specific or panel of stem cell glycan markers. With this in mind, statistical analysis (paired *t*-tests) of those structures which are shared between ADSCs and adipocytes, and between ADSCs and dADSCs were used to determine whether any glycan structures could be a marker of stem cell differentiation. Structures with $P<0.05$ and a fold change of greater than two-fold between samples were used in this analysis.

In the comparison of ADSCs and adipocytes, seven *N*-linked glycan structures were found to be significantly increased in adipocytes (Figure 3.15), while 18 structures were significantly decreased in adipocytes (Figure 3.15). Four of the seven structures up-regulated in adipocytes were bisecting GlcNAc-type structures, substantiating that this class of glycan type is characteristic of primary adipocytes. Differences in sialic acid linkages were also seen between native adipocytes and ADSCs, namely in the hybrid-type structures (structures #40 and 54) that both contain an α2-3 sialic acid residue and differ only in the presence of a core fucose residue in structure #54.

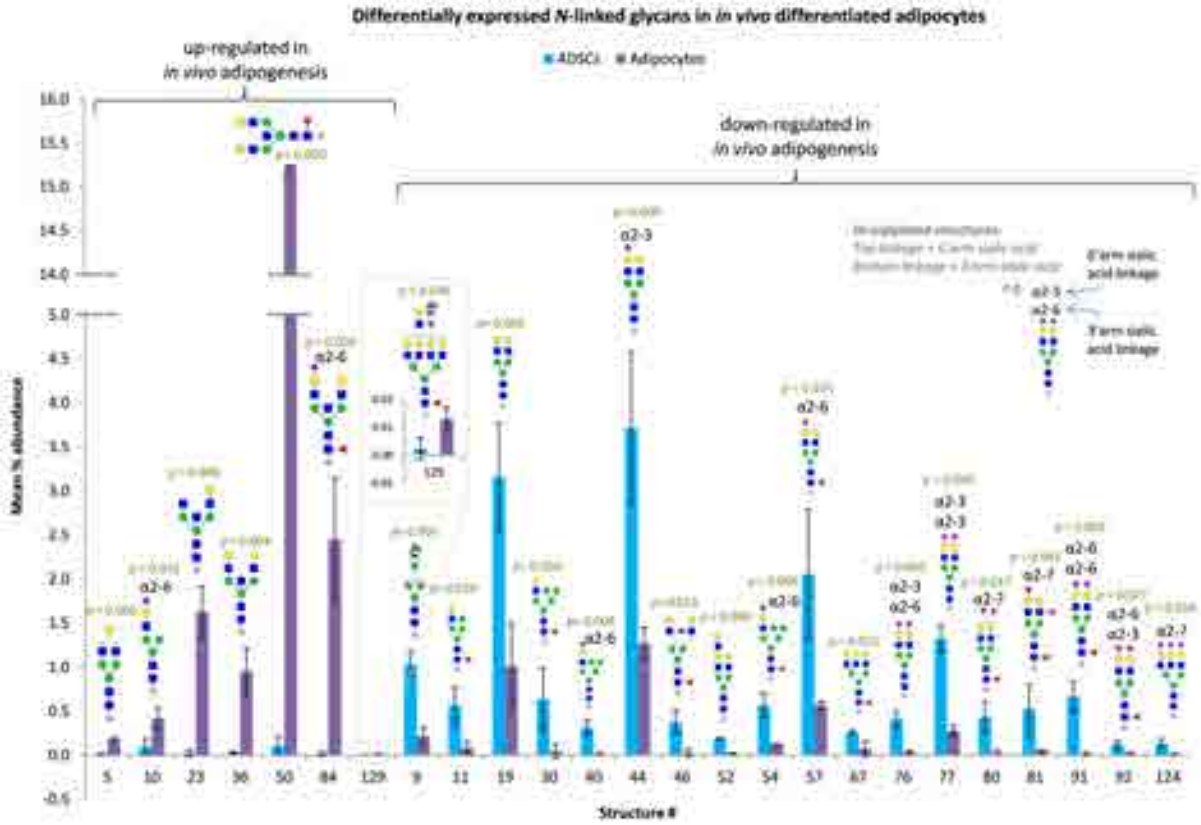


Figure 3.15 Glycan structures with significantly increased and decreased expression ($P < 0.05$) in adipocytes compared to stem cells with a fold change greater than two. Seven structures were up-regulated in expression in adipocytes, and 18 down-regulated.

On the other hand, five structures were found to be significantly up-regulated in dADSCs compared to ADSCs and four structures significantly down-regulated (Figure 3.16) with a fold change of greater than two-fold. Three of the five structures (structures #36, 50 and 84) were bisecting GlcNAc structures and are the same as those significantly increased in the native adipocytes. The bisecting GlcNAc structures therefore may present a panel of biosynthetically related markers with which to assess the degree of differentiation of ADSCs. These structures are present at very low abundances in the differentiated *in vitro* dADSCs (0.1-1%) compared to native adipocytes (1-15%) of total structures.

Again, sialic acid differences were seen between ADSCs and dADSCs, namely the down-regulation of α2-3 linked sialic acids in the di-sialylated glycans (structures #77 and 94). As with structures #40 and 54, these two structures are identical with the exception of the core fucose residue in structure #94. Both structures carry a α2-3 linked sialic acid residue on each arm and indicate di-sialylated glycans are characteristic of ADSCs.

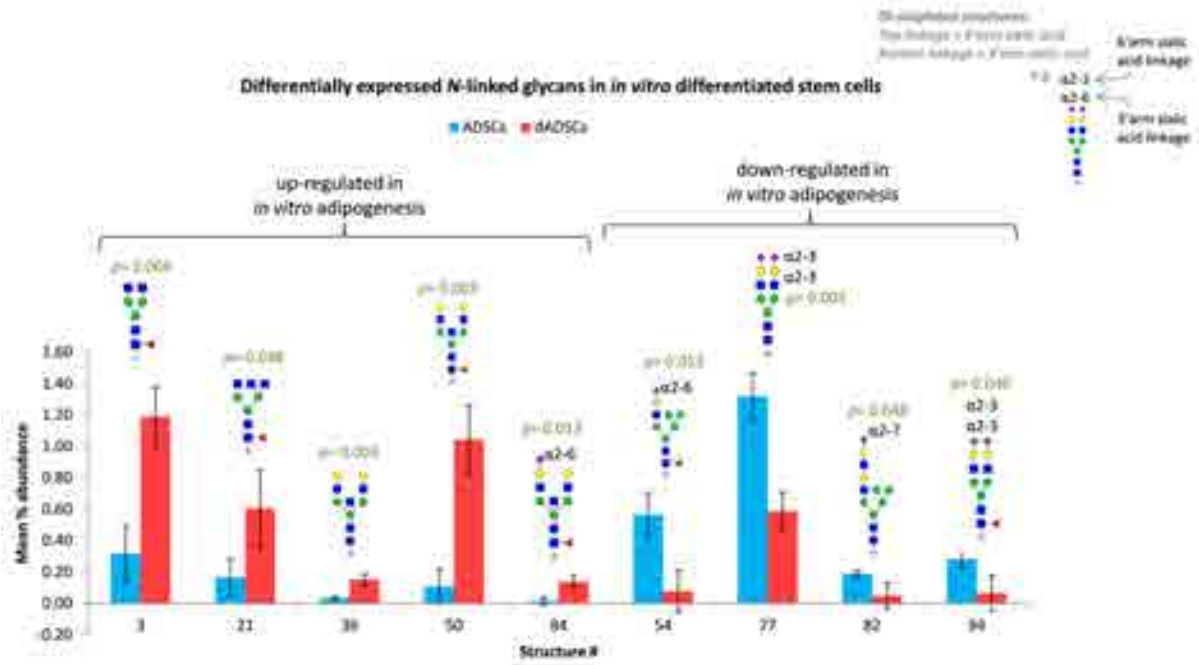


Figure 3.16 Glycan structures with significantly increased and decreased expression ($P < 0.05$) in dADSCs compared to ADSCs with a fold change greater than two-fold. Five structures were up-regulated in expression in dADSCs, and four down-regulated in dADSCs.

3.3.10 O*-glycome of ADSCs, dADSCs and adipocytes*

O-linked glycans were released from the membrane proteins of ADSCs, dADSCs and adipocytes using reductive beta elimination before analysis using PGC-LC ESI-MS/MS. Structural analysis of the released oligosaccharides resulted in the identification of seven *O*-linked glycans from five different compositions (Table S3.3, Supplementary data).

Qualitatively, the *O*-glycome did not differ between ADSCs, dADSCs and adipocytes, with the exception of one of the structural isomers with *m/z* [675.3] $^{-}$ (structure O1), which was only present in adipocytes. Comparison of the average MS profiles (Figure 3.17) also displayed a mono-sialylated, linear Core 1 structure with *m/z* [675.3] $^{-}$ (structure O3) as the most dominant ion observed in the adipocyte average *O*-linked glycan MS profile, followed by the di-sialylated counterpart with *m/z* [966.4] $^{-}$ (structure O4). The reverse percentage abundance trend was observed in the ADSC and dADSC profiles.

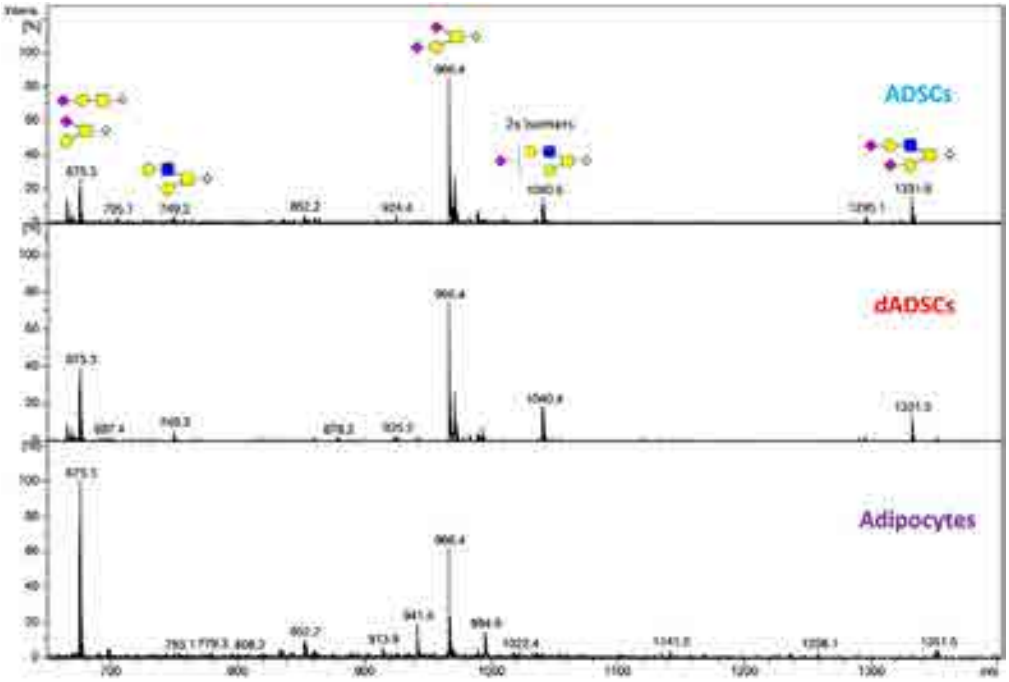


Figure 3.17 Representative average *O*-linked glycan MS spectra of ADSCs, dADSCs and adipocytes. Structures have been assigned to their corresponding ion signals.

Overall, mono-sialylated structures were significantly increased in adipocytes (P = 0.02) and dADSCs (P < 0.01), while di-sialylated structures were significantly decreased in adipocytes (P = 0.03) and dADSCs (P = 0.04) when compared to ADSCs. No statistical differences were observed in the neutral structures in either model of adipogenesis. These changes are displayed in Figure 3.18.

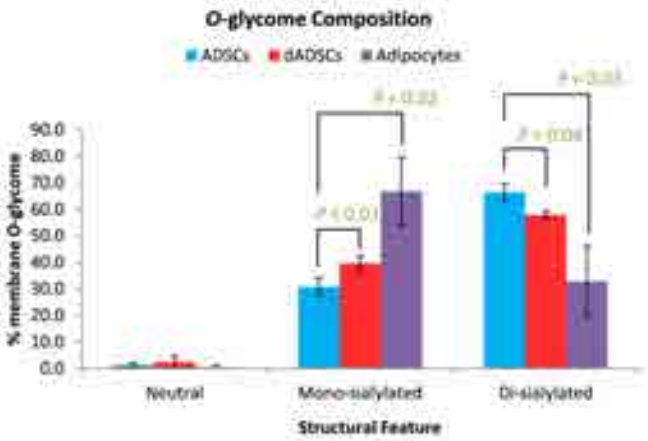


Figure 3.18 In total, seven *O*-linked glycans were identified in ADSCs, dADSCs and adipocytes. These structures were compared according to structural features, namely neutral and degree of sialylation.

3.3.11 Validation of bisecting GlcNAc and core fucosylation expression in differentiation of ADSCs

The results of glycomics analyses can be validated using various orthogonal techniques. Targeted techniques such as lectin blotting or antibody-based platforms such as Western blotting can also be used to identify glycan cell markers (Christiansen, Chik *et al.*, 2014) and used to verify specific structural features identified in MS data. To validate the significant increase in expression of bisecting GlcNAc structures in both *in vivo* and *in vitro* differentiation, lectin blotting was evaluated as a potential orthogonal technique. mRNA transcript expression was used to measure expression levels of the genes responsible for the synthesis of bisecting GlcNAc-type (MGAT3) and core fucosylated (Fut8) structures.

3.3.11.1 Lectin blotting to confirm bisecting GlcNAc expression

Lectin dot blotting using the PHA-E (Phaseolus vulgaris erthyroagglutinin) lectin which recognises bisecting GlcNAc residues was evaluated by testing PHA-E specificity using bovine fetuin and bovine serum albumin (BSA) as negative controls, and human IgG as a positive control. PHA-E did not bind to BSA (Figure 3.19), a non-glycosylated protein (Cao, Shen *et al.*, 2009), but did bind to the second negative control, bovine fetuin which contains no bisecting GlcNAc structures as previously confirmed by PGC-LC ESI-MS (Figure 3.4). As expected, binding was seen in the positive control human IgG, which has been demonstrated to contain bisecting GlcNAc structures (Wuhrer, Stam *et al.*, 2007). Densitometry analysis using ImageJ software (Schneider, Rasband *et al.*, 2012) indicated that PHA-E binding was greater in IgG (8.87E+06 units) compared to fetuin (5.30E+06 units), however binding to fetuin indicated that the lectin had unspecific binding which may be from binding to GlcNAc residues in general. Lectin blotting using PHA-E was therefore not further pursued.

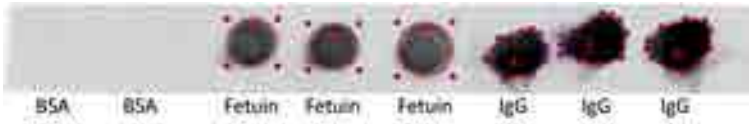


Figure 3.19 PHA-E lectin dot blot of standard glycoproteins bovine fetuin and human IgG, and bovine serum albumin.

3.3.11.2 mRNA expression levels of MGAT3 and Fut8 during adipogenesis

To validate the significant increase in expression of bisecting GlcNAc structures in both *in vivo* and *in vitro* differentiation, and the significant increase in core fucosylated structures in native adipocytes compared to ADSCs and core fucosylated structures, expression of mRNA gene transcripts for the corresponding MGAT3 and Fut8 genes was conducted using qRT-PCR. Expression levels of MGAT3 were significantly higher in both differentiated ADSCs (149-fold increase) and mature adipocytes (279-fold increase) compared to ADSCs (Figure 3.20c), closely mirroring the trends seen in expression of bisecting GlcNAc structures (Figure 3.20a). However, expression levels of Fut8 (Figure 3.20d) was significantly higher in ADSCs compared to both adipocytes and dADSCs, which contrasts with the increased trends seen in core fucosylated product structures (Figure 3.20b). It is thus clear that expression of the glycosyltransferase genes does not always reflect the product level of their activities.

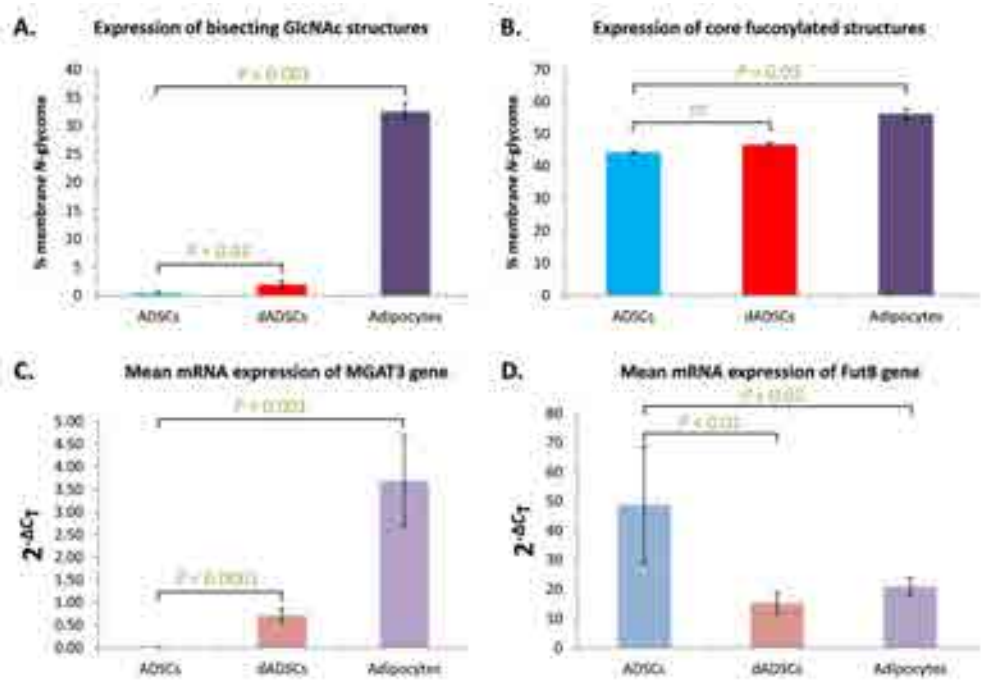


Figure 3.20 The overall abundance of bisecting GlcNAc-type structures **(A)** and total core fucosylated structures (**B**) in the *N*-glycomes of ADSCs, dADSCs and adipocytes were compared to the average mRNA expression levels of MGAT3 (**C**) and Fut8 (**D**). mRNA expression levels were normalised to the TBP housekeeping gene and shown as mean ± SD $2^{-\Delta CT}$ (n = 6). *P*-values from student *t*-tests comparing expression levels between dADSCs and adipocytes with ADSCs are shown.

Therefore, the most striking difference observed in differentiation of adipocyte-derived stem cells was the expression of bisecting GlcNAc structures on the membrane proteins, which is seen here to correlate with the increased expression of the biosynthetic gene MGAT3.

3.4 Discussion

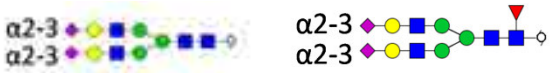
In this chapter, the released *N*- and *O*-linked glycans from the membrane glycoproteins of ADSCs and their *in vivo* (adipocytes) and *in vitro* (dADSCs) adipogenically differentiated products were extensively characterised and compared. Using PGC-LC coupled to negative ion ESI-MS/MS, 130 *N*-linked and seven *O*-linked glycan structures were characterised. A further 23 possible *N*-linked structures were detected but could not be verified with MS/MS fragmentation data and were therefore not included in the analysis. The major findings of this chapter are discussed below and summarised in Table 3.3.

The use of Triton X-114 phase partitioning as an extra enrichment step of membrane proteins was found to be an important and crucial step, with approximately five times enrichment of *N*-linked glycans detected when phase partitioning was carried out. A visual comparison of the proteins present also revealed more protein bands present when Triton X-114 was used. Since the majority of glycoproteins are membrane proteins which require detergents for solubilisation due to their hydrophobicity (Abdul Rahman, Bergstrom *et al.*, 2014), this is the most likely explanation for the higher yield of glycans seen in phase-partitioned samples. Differences in the degree of gel staining indicated difficulty in obtaining equal sample loading despite protein quantitation using Bradford assay which may be due to incompatibility of sample solvent and the Bradford assay reagent, despite decreasing the concentration of the solvent components to within the recommended range.

For unambiguous assignment of structures, the mass spectrometric analysis of un-labelled glycans relies on a relatively high amount of starting material. The reduction and alkylation of fetuin prior to de-glycosylation was evaluated and found to double the signal intensity of the spectra. In conjunction with PGC-clean up prior to MS analysis, the generated data had a lower signal to noise ratio and also provided more informative MS/MS fragmentation data with which to assign structures thereby circumventing the need for more starting sample. To avoid day-to-day MS and particularly column retention time variation, samples from different biological replicates were run in succession on the same column before glycomics analysis. For glycan analysis, this allowed confident and reproducible isomer retention time matching between samples which was important when good quality MS/MS fragmentation data was available in one sample and not another. Retention time patterns were also relied upon along with the precursor mass, diagnostic MS/MS ions, knowledge of

the mammalian glycosylation pathway and exoglycosidase digestions for assignment of structure topology and linkage assignment.

Table 3.3 Summary of key findings

Significant changes (*P* < 0.05) upon adipogenic differentiation of ADSCs *in vivo* and *in vitro*:	
- ↑ Complex-type, bisecting GlcNAc *N*-linked glycans (0.4 to 1.9 % *in vitro*, and 0.4 – 32.1% *in vivo*) ↑ Corresponding mRNA expression of MGAT3 (149-fold higher *in vitro* and 279-fold higher *in vivo*) - ↑ Mono-sialylated *O*-linked glycans (30.8 to 39.7 % *in vitro*, and 30.8 to 66.8 %) - ↓ Di-sialylated *O*-linked glycans (66.5 to 58.0 % in vitro, and 66.5 to 32.9 % *in vivo*)	
Significant changes (*P* < 0.05) upon adipogenic differentiation of ADSCs *in vivo*:	**Percentage abundance of corresponding feature in *in vitro* differentiated ADSCs:**
- ↑ Complex-type, α1,6 core fucosylated structures (38 to 54 %)	41%
- ↓ Complex-type biantennary *N*-linked glycans (43 to 26 %)	37 %
Biosynthetically related structural markers characteristic of *in vivo* mature adipocytes with potential future use in assessing degree of *in vitro* adipogenic differentiation:	
 structure #36 structure #50 structure #84	
Biosynthetically related structural markers characteristic of ADSCs (significantly higher in ADSCs compared to both adipogenic products):	
 structure #77 structure #94	

Comparison of the *N*- and *O*-glycomic profiles over different cell culture passages confirmed that the expression profiles were not overly influenced by the artificial growth environment. A large variety of *N*-linked glycans were identified, ranging from pauci mannose to larger structures containing poly-*N*-acetyllactosamine repeats. The greatest variation in glycan expression was observed between dADSC samples, indicating varying responses to the differentiation of cells from different individuals, which has been reported in the literature (Schipper, Marra *et al.*, 2008).

Significant up-regulation of bisecting GlcNAc-type structures was evident in both models of adipogenesis compared to the progenitor ADSCs. This structure type is formed by the addition of an *N*-acetylglucosamine residue in a β1-4 linkage to the inner most mannose residue of the *N*-linked glycan core (Takahashi, Kuroki *et al.*, 2009). Bisecting GlcNAc complex-type structures were found to be characteristic of *in vivo* adipogenesis comprising 32.1 ± 1.5 % of the adipocyte *N*-glycome. The adipocytes contained 13 bisecting GlcNAc structures that were not detected in ADSCs and dADSCs. *In*

vitro adipogenesis of the ADSCs resulted in the expression of three bisecting GlcNAc structures in the dADSCs which were the same as those in the native adipocytes. This increase in bisecting GlcNAc structures was the only significant *N*-glycosylation change seen during *in vitro* adipogenesis, but only comprised 1.9 ± 0.7 % of the dADSC *N*-glycome. While some bisecting GlcNAc structures were shared by all three cell types, the majority of these had significantly higher expression in both adipogenic products compared to ADSCs (0.4 ± 0.4 %). mRNA expression profiling of the glycosyltransferase responsible for catalysing the addition of a bisecting GlcNAc, β1-4 *N*-acetylglucosaminyltransferase (coded by the MGAT3 gene) revealed transcript expression levels reflecting the overall expression of bisecting GlcNAc *N*-linked glycans (bisecting complex- and hybrid-type structures combined) in both adipocytes and dADSCs. Significantly elevated MGAT3 transcripts were observed in the adipocytes, and a lesser but still significant increase seen in the *in vitro* differentiated stem cell products. This is thus a close correlation between bisecting glycan structures being expressed on the membrane proteins and the transcript levels of the glycosyltransferase responsible for their production. There have been no reports in the literature of the functional role bisecting GlcNAc structures play during adipogenesis or in mature adipocytes. This functional importance of this structural type however has been demonstrated in various proteins from other biological systems such as epidermal growth factor receptor in HeLaS3 cells (Sato, Takahashi *et al.*, 2001) and integrin-α5β1 in melanoma cells (Isaji, Gu *et al.*, 2004).

A recent study by Hamouda and colleagues also investigated *N*-glycomic changes upon *in vitro* adipogenic differentiation (Hamouda, Ullah *et al.*, 2013), using matrix assisted laser desorption ionisation time-of-flight mass spectrometry analysis of permethylated glycans. The mesenchymal stem cells (MSCs) used in their study were derived from bone marrow, and while a large number of *N*-linked glycans were identified, the authors did not report the presence of any bisecting GlcNAc structures. Indeed, in the present study, complex-type bisecting GlcNAc structures in the differentiated ADSCs were only at low abundance (0.4 % and 1.9 % of the ADSC and dADSC *N*-glycomes respectively). It is possible therefore that any bisecting structures in the *in vitro* system used may have been below the limit of detection in the study by Hamouda *et al*. Apart from this, the main global observation shared between both studies was the dominance of core fucosylated structures in both the bone marrow MSCs and adipogenically differentiated MSCs. A comparison of the average MS profile of PNGase-F released structures from bone marrow derived stem cells analysed by Hamouda *et al*. (Figure 4; Hamouda, Ullah *et al*., 2013) with that of the average ADSC and dADSC MS profiles in the present study also revealed that two most abundant *N*-linked glycan structures were the same complex-type structures: (Hex)2 (HexNAc)2 (Fuc)1 + (Man)3(GlcNAc)2, and the mono-sialylated counterpart with composition of (Hex)2 (HexNAc)2 (Fuc)1 (NeuAc)1 +

(Man)3(GlcNAc)2. The present study identified three structural isomers of this composition, with the isomer with an α2-3 linked sialic acid on the 3'arm primarily responsible for the high intensity of the MS signal ion. This highlights the advantages of using PGC-LC for exposing differential expression of isomeric structures. Hamouda *et al*. (2013) also reported the presence of neutral and sialylated structures containing Gal-Gal epitopes as seen in this study, although the authors did not report the linkage type of the terminal galactose residue.

In addition to a significant increase in percentage abundance of bisecting GlcNAc structures in the differentiated ADSCs, the second significant change was the high abundance of complex-type core fucosylated structures in adipocytes compared with ADSCs with an increase indicated also upon *in vitro* differentiation of the ADSCs. In mammalian systems, core fucosylation is a process catalysed by α-1,6 fucosyltransferase (Fut8). Fut8 alone is responsible for core fucosylation, with no known homologous genes documented (Wang, Wang *et al.*, 2005). Fut8 gene expression however did not correlate with the overall expression of core fucosylated structures on membrane proteins (including core fucosylated pauci mannose, hybrid and complex structures). In fact, while there was a significant Fut8 transcript expression difference in the *in vivo* differentiated adipocytes, this was not evident in the product glycan structures. Fut8 mRNA expression was significantly lower, while core fucosylated structures were significantly higher in adipocytes compared to ADSCs. Because glycosylation is a post-translational modification, it is not necessarily always driven by gene transcripts (Nairn, Aoki *et al.*, 2012; Nairn, York *et al.*, 2008). In the case of Fut8 expression here, the results indicate that control of core fucosylation in adipogenesis indeed is not regulated at the transcript level. The biological importance and function of core fucosylation is still not fully understood, however many studies have demonstrated its role in different systems. Perhaps the most poignant example of the importance of core fucosylation is in the work of Wang and co-workers, who created Fut8-/- mice (Wang, Wang *et al.*, 2005). Early death occurred in 70 % of mice, and growth retardation and emphysema symptoms presented in the remaining 30 % emphasising the important role of Fut8 in development. Amongst other examples, the importance of core fucosylation has also been reported in epidermal growth factor receptor ligand binding with subsequent impact on cell differentiation and growth (Wang, Gu *et al.*, 2006). Undoubtedly core fucosylation is important but its role is in adipogenesis in the context of this study is not clear at present.

Conversely, a high abundance of biantennary-type structures was found in ADSCs which was significantly higher than in adipocytes. This may correlate with these structures not being further processed by the β-1,4 *N*-acetylglucosaminyltransferase to become bisecting GlcNAc structures as seen in the native adipocytes. A large percentage of high mannose-type structures were found

present in all three cell types, which is most likely due to the presence of Golgi and endoplasmic reticulum (ER) membrane proteins also enriched with the plasma membranes in the sample processing. Intracellular glycans tend to be immature, unprocessed glycans and likely contributed to the overall *N*-glycan profile in all samples. This observation has also been reported by Heiskanen and co-workers in their characterisation of osteogenic differentiation of bone marrow-derived MSCs (Heiskanen, Hirvonen *et al.*, 2009).

Expression of linear poly-*N*-acetyllactosamine-containing structures was detected in all samples. Poly-*N*-acetyllactosamine antennae are generally hard to detect in routine glycomics analyses, primarily due to their large size and low abundance in the glycome (Bern, Brito *et al.*, 2013). In this study, seven poly-*N*-acetyllactosamine-containing structures were identified with the highest average percentage abundance in adipocytes, comprising 0.6 ± 0.1 % of the *N*-glycome. Higher branching type *N*-linked glycans with compositions containing four or more *N*-acetylhexosamine residues and three or more *N*-acetylneuraminic acid residues indicated the presence of large highly branched structures that are also generally hard to detect in routine glycomics analyses.

Interestingly, no structures contained the *N*-glycolylneuraminic acid (NeuGc) xenoantigen were detected in this study. NeuGc is a common contaminant in fetal bovine serum supplemented cultures and there are various proposals for the mechanism of incorporation into the glycome (Lanctot, Gage *et al.*, 2007). The culturing of human ADSCs with heat-inactivated FBS, as was the case in the present study, has previously been demonstrated to reduce NeuGc incorporation and prevent immune response (Komoda, Okura *et al.*, 2010). When proposing cultured cells as a source of therapy, eliminating NeuGc is an important consideration to avoid immune system response and their absence in the present cell differentiation study is encouraging for future therapeutic applications.

Previous studies have used mass spectrometry to characterise the surface glycome of MSCs during different types of differentiation, with the specific aim to identify MSC-specific glycan markers to aid in their detection and isolation from complex tissues. While different differentiation protocols were used in these studies, the glycome of the progenitor MSCs can be compared with the ADSCs in the present study as ADSCs are in fact classified as mesenchymal stem cells. Similar to the present study, Heiskanen and colleagues (2009) reported high mannose-type structures, linear poly-*N*-acetyllactosamine repeats and α2-3 sialylation as being amongst the characteristic glycosylation features in bone marrow derived MSCs. Satomaa *et al*. (2009) reported sialyl Lewis x and H type 2 antennae as the most abundant complex-type structures in the *N*-glycome of human embryonic stem cells. In comparison, very few Lewis x-containing structures were identified in the present study, with core fucose residues mainly found, and only one possible structure containing a H-type 2 epitope

(structure #79). Hamouda *et al*. (2013) also reported fucosylation occurring predominantly at the reducing terminus. So far, the work of Hamouda *et al*. has been the only other report to propose specific structures as markers of MSCs and their *in vitro* adipogenically differentiated products. One of the proposed MSC markers was a core fucosylated, biantennary structure which was also identified in the present work as higher in ADSCs when compared to dADSCs (structure #3). While encouraging, the overall variation in conclusions drawn from the different studies suggests that efforts to identify a single or set of glycan markers specific to MSCs should be refocused to finding tissue specific-derived MSC markers.

The small number of *O*-linked glycans identified in this study is not surprising, as *O*-linked glycans tend to be largely present on secreted glycoproteins such as mucins (Thaysen-Andersen and Packer, 2014; Tran and Ten Hagen, 2013). Despite the small number of *O*-linked glycans identified, the percentage of mono-sialylated structures was significantly higher in both *in vivo* and *in vitro* adipogenesis, while di-sialylated structures were significantly decreased compared to the progenitor stem cells.

The fragmentation type used throughout this thesis was collision induced dissociation (CID) which has a well documented fragmentation pathway and produces glycosidic and information-rich cross ring fragments (Domon and Costello, 1988). Recent advances in mass spectrometry have seen the emergence of other fragmentation methods such as high energy C-trap dissociation (HCD) which has been demonstrated as a viable alternative to CID and avoids problems such as neutral loss and fragment rearrangement events during MS/MS (Kenny, A. Hayes *et al.*, 2011; Kenny, Issa *et al.*, 2011). HCD could therefore be considered for complementary analyses in future glycoprofiling experiments.

Using a targeted approach to investigate specific glycosyltransferase genes, a clear correlation between MGAT3 transcript expression with bisecting GlcNAc glycan type expression was evident. Measurement of Fut8 transcript levels however did not reflect the overall expression of core fucosylated structures. Taken together, the results indicate that control of specific glycosylation features is not always determined at the transcript level in adipogenesis. Recent work by Kurimoto *et al.* (2014) also suggested that expression of glycosylation characteristics was rather dependent on the type of glycoprotein. Ultimately, the expression of particular glycosylation structural features is determined by many components; enzymes, co-factors, transporters and activated sugar donors that are collectively referred to as the 'glycosylation machinery' and encompass several cellular organelles (van Kooyk, Kalay *et al.*, 2013). Recent work has also demonstrated the role DNA methylation status plays on glycan expression (Anugraham, Jacob *et al.*, 2014), and miRNA regulation of the glycome has

also been reported (Kasper, Koppolu *et al.*, 2014). Overall the results presented here further highlight the complex nature of glycosylation and the still preliminary understanding of how glycan biosynthesis is controlled.

3.5 Concluding remarks

The use of PGC-LC coupled with ESI-MS/MS proved to be a reproducible and reliable analytical method, and allowed the identification of a large number of membrane protein derived *N*-linked glycan structures. The detailed structural characterisation of the membrane glycosylation of ADSCs and their *in vivo* and *in vitro* adipogenically differentiated products revealed bisecting GlcNAc *N*-linked glycan structures on the membrane proteins to be characteristic of adipogenesis. This structural type most likely plays a critical role in the functioning of adipocytes and their unique role in the body. *In vitro* differentiated adipose derived stem cells however, had only about 10% of the bisecting GlcNAc structures of those seen to be abundant in the *in vivo* native adipocytes. This finding was mirrored in the mRNA expression of the enzyme responsible for the synthesis of bisecting GlcNAc, MGAT3 (β-1,4 *N*-acetylglucosaminyltransferase), transcripts. Overall, the results in this chapter indicate that the membrane glycosylation of *in vitro* differentiated adipose-derived stem cells do not reflect that of *in vivo* differentiated stem cells (adipocytes). The glycome differences seen in native adipocytes compared to their progenitor ADSCs demonstrates the alteration of specific glycan modifications in cell differentiation, confirming and offering a valuable insight into the role of glycosylation during adipogenesis, and points to the analysis of the membrane glycome as a tool to evaluate adipocyte differentiation stage.

Chapter 4:

Characterisation and comparison of the total membrane proteome during *in vivo* and *in vitro* adipogenic differentiation

4.1 Introduction

The cell membrane contains a myriad of proteins that initiate a variety of intracellular signalling cascades (Jeong, Ko *et al.*, 2007). This protein profile is constantly changing and adapting as part of response and interaction with the external environment (Gundry, Boheler *et al.*, 2008). Stem cell differentiation potentially is one such process that influences protein changes on the cell surface as the cell responds to signals from within the stem cell niche. By characterising and quantitating these protein changes during stem cell differentiation, the processes, pathways and networks underlying differentiation may be further understood.

Some studies have investigated the whole proteome during stem cell differentiation (Jeong, Ko *et al.*, 2007; Melo-Braga, Schulz *et al.*, 2014), but few have specifically targeted the plasma membrane (PM) proteome (Dormeyer, van Hoof *et al.*, 2008a; Hao, Li *et al.*, 2013). Much of this effort has been with the aim of identifying new stem cell markers, as one of the current criterion used to identify stem cells is a panel of membrane proteins called 'clusters of differentiation' or 'CD markers' (Dominici, Le Blanc *et al.*, 2006).

Proteomic investigations of adipogenic differentiation in human stem cells in particular have included whole-cell protein analysis of adipogenically differentiated bone marrow-derived mesenchymal stem cells (Jeong, Ko *et al.*, 2007) and of adipose-derived stem cell (ADSCs) (DeLany, Floyd *et al.*, 2005). Adipogenic differentiation of ADSCs has also been investigated at the transcriptome level (Urs, Smith *et al.*, 2004) and by membrane CD marker profiling using flow cytometry (Gronthos, Franklin *et al.*, 2001). However, to our knowledge, such an analysis has not been attempted by a targeted membrane proteomics approach.

Mass spectrometry (MS) is largely the method of choice for identifying and characterising proteins (Luque-Garcia and Neubert, 2009), and when paired with various techniques such as sample-specific preparation, pre-MS sample fractionation methodologies and post-MS bioinformatics tools (van Hoof, Krijgsveld *et al.*, 2012), MS offers a powerful tool to probe the depths of the proteome. Following from the analysis of membrane glycosylation conducted in Chapter 3 and the observation of glycosylation changes during adipogenic differentiation, an analysis of the membrane proteome was undertaken. The aim was to investigate other cell surface molecular changes occurring during adipogenic differentiation and, as an orthogonal analysis of protein glycosylation, to characterise the population of proteins from which the glycans analysed in Chapter 3 were released. Cell membrane enrichment using ultracentrifugation and phase-partitioning of hydrophobic proteins was followed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) separation of membrane proteins. By enriching for hydrophobic membrane proteins, this enriched a population of proteins

which also included plasma membrane (PM) proteins, allowing changes at the cell surface to be targeted in addition to the global membrane protein analysis. Using a label-free quantitative shotgun proteomics workflow, samples were subjected to nanoLC/MS-MS to characterise and compare membrane protein changes during *in vitro* and *in vivo* adipogenic differentiation. Given cell membrane CD markers change upon induction of stem cell differentiation (Gronthos, Franklin *et al.*, 2001), together with the changes in glycosylation observed in Chapter 3, broader membrane protein changes were explored in cells which had undergone *in vivo* and *in vitro* adipogenesis.

4.2 Materials and Methods

4.2.1 Materials

Lithium dodecyl sulphate (LDS) sample buffer, 3-(N-morpholino)propansulfonic acid (MOPS) buffer, NuPAGE Novex 4-12 % (w/v) polyacrylamide Bis-Tris precast gels were purchased from Invitrogen (Life Technologies, Australia). Bradford assay, dithiothreitol (DTT), ammonium bicarbonate (NH_4HCO_3), formic acid, acetonitrile (ACN) and iodoacetamide (IAA) were from Sigma (St Louis, MO, USA). µC $_8$ ZipTips were from Eppendorf (Hamburg, Germany), and Coomassie Brilliant Blue G-250 dye was from Biorad (NSW, Australia). HPLC-grade water was obtained using a Milli-Q synthesis water purification system (Millipore, MA, USA). Magic C18 resin was from Michrom Bioresources (Auburn, CA) and sequencing grade modified Trypsin from Promega (WI, USA).

4.2.2 Sample collection and membrane protein enrichment

The human abdominal lipo-aspirates used to derive adipose-derived cells came from the same three individuals used in the glycomic analysis carried out in Chapter 3. Samples were collected and prepared as described in Chapter 2, Section 2.2.2. Briefly, adipocytes from collagenase-digested adipose tissue were collected in 5 ml amounts for analysis, and the stromal vascular fraction introduced to cell culture to isolate plastic adherent ADSCs. At passage three, half of the culture flasks were harvested for proteomic analysis of the ADSCs while the remaining cells were used for induction of adipogenic differentiation. The cell membrane proteins were isolated from all samples as described in Chapter 3, Section 3.2.3 using Triton X-114 phase partitioning to enrich for hydrophobic membrane proteins.

4.2.3 SDS-PAGE fractionation of membrane proteins and in-gel digestion

Proteins were quantified using Bradford assay (Bradford, 1976) and accordingly approximately 25 µg of protein was reduced with 1 µl 100 mM DTT at 70 °C for 10 min in LDS sample buffer diluted with the appropriate volume of water for each sample for a total volume of 14 µl. IAA (1 µl 500 mM) was added to each sample and proteins alkylated at room temperature for 30 min. Samples were then

loaded onto a NuPAGE Novex 4-12% (w/v) polyacrylamide Bis-Tris precast gel and electrophoresed for 2 h at 100 V using MOPS buffer. The gel was fixed with 10 % (v/v) methanol, 1 % (v/v) acetic acid before overnight staining with Coomassie G-250. Stained proteins were visualised using an Odyssey Infrared Imaging System (LiCor).

Each gel lane was excised from the gel and cut into six equal fractions, and then each chopped into smaller pieces which were transferred into eppendorf tubes for each fraction. Gel pieces were washed twice with 50 % (v/v) ACN and then dehydrated with 100 % (v/v) ACN. This was followed by rehydration with 100 mM NH_4HCO_3 for five min before dehydration again with 100 % (v/v) ACN. This was repeated until all pieces were destained. For protein digestion, 1 µl trypsin (50 mg/µl in 1 mM HCl) was diluted in 2 µl 25 mM NH_4HCO_3 and applied to dehydrated gel pieces for 10 min at 37 ℃. Approximately 50 µl 100 mM NH_4HCO_3 was added to each sample to ensure gel pieces were covered for overnight incubation at 37 ℃. Peptides were extracted with 30 µl 25 mM NH_4HCO_3 followed by 15 min sonication in a water bath and collected into a fresh eppendorf tube. Gel pieces were dehydrated with 100 % (v/v) ACN followed by 15 min sonication and the supernatant pooled with the initial extracts. Gel pieces were then washed twice with 5 % (v/v) formic acid followed by sonication and the supernatants collected and pooled at each step. Peptide extracts were then dried using a vacuum centrifuge and resuspended in 10 µl 0.1 % (v/v) formic acid. For desalting of samples, peptides were applied to a μC_8 ZipTip and the flow through reloaded to maximise recovery. Peptides were eluted in 90 % (v/v) ACN, 0.1 % (v/v) formic acid and dried under vacuum.

4.2.4 Nanoflow RPLC-MS/MS

Each sample was resuspended in 10 µl 2 % (v/v) formic acid and the peptides separated and analysed using nanoflow LC-MS/MS with a LTQ-Velos Pro linear ion trap mass spectrometer (Thermo, San Jose, CA). Peptides were separated on a reversed-phase column using an Easy nLC II system (Thermo, San Jose, CA). The reversed-phase column was made in-house by packing a fused silica capillary with an integrated electrospray tip with 100 Å, 5 µm Magic C18 resin to approximately 10 cm (with 100 µm inner diameter). A vented pre-column of 3 cm x 100 µm inner diameter packed with the same material was used upstream of the analytical column for sample loading. An electrospray voltage of 1.8 kV was applied via a liquid junction upstream of the C18 column. Samples were injected onto the column followed by an initial wash step with buffer A (5 % (v/v) ACN, 0.1 % (v/v) formic acid) for two minutes at 550 nl/min. At a flow rate of 500 nl/min, peptides were eluted with 0 50 % buffer B (95 % (v/v) ACN, 0.1 % (v/v) formic acid) for 38 min at 500 nl/min followed by 9 min at 800 nl/min. The column eluant was directed into a nanospray ionisation source of the mass spectrometer. Spectra in positive ion mode were scanned over the range *m/z* 400 1500 and, using Xcalibur software (Version 2.06, Thermo), automated peak recognition, dynamic exclusion of 90 s

and MS/MS of the top nine most intense precursor ions at 35 % normalisation collision energy were performed.

4.2.5 Database search for peptide and protein identification

The LTQ-Velos Pro .raw output files were converted into mzXML format and searched against the curated human Uniprot/Swissprot database (20, 264 proteins as of 14 March 2014). This was performed using the Global Proteome Machine (GPM) software (version 2.1.1) using the X!Tandem algorithm (Craig and Beavis, 2004). The six fractions of each gel-separated membrane protein sample were processed sequentially, generating output files for each fraction along with a merged non-redundant output file for the sample. Peptide identification was determined using a 2 Da parent ion tolerance, 0.4 Da fragment ion tolerance, carbamidomethyl was considered as a complete modification, while oxidation of methionine and threonine, and carbamylation of N-termini and lysine were considered as partial modifications. Reverse database searching was also conducted for estimating false discovery rates (FDRs).

4.2.6 Data processing and quantitation

The lists of proteins identified at low stringency for three biological replicates of each cell sample were exported from GPM and contaminants of bovine serum albumin, porcine trypsin and Lys C hits were removed. Protein identifications were further processed using the Spectral Counting Reporting and Analysis Program (Scrappy) (Neilson, Keighley *et al.*, 2013). For each respective set of three biological replicates for adipocytes, ADSCs and dADSCs, the data were combined to produce a single list of very high stringency identified proteins. These were generated using the following criteria:

a) The protein must be present in all three replicates and,
b) Have a minimum total of five or more peptides across all three replicates in at least one sample type

For example, for a given protein, if the ADSCs had ≥ five total peptides and the dADSCs only three, with no missing values in any of the replicates, it was still considered present in the dADSCs and used for pair-wise comparison. For proteins where the ADSCs had ≥ five peptides and dADSCs had none or missing values in a replicate, it was treated as present in the former and absent in the latter.

Peptide FDR was calculated by 2 x (total number of peptides identified for reversed protein hits/total number of peptides identified for all proteins in the list) x 100 (Peng, Elias *et al.*, 2003) and protein FDR calculated by (number of reversed protein hits/total number of proteins in the list) x 100.

For quantitation, normalised spectral abundance factors (NSAF) were calculated using the following formula: NSAF = (SpC/L)/Σ(Spc/L), where SpC refers to spectral count (number of non-redundant

peptide identifications for a given protein) and L refers to the length of the protein (Zybailov, Mosley *et al.*, 2006). Half a spectral count (0.5) was added to all counts to compensate for null values before the natural log of the NSAF values were taken for data normalisation and statistical analysis. Data quality was further examined by generating box plots and kernel density plots of log NSAF values to ensure normal distribution.

4.2.7 Statistical methods

Student's *t*-tests were carried out for pair-wise comparisons (Adipocytes vs. ADSCs, and dADSCs vs. ADSCs) and reproducibly identified proteins with $P < 0.05$ and fold change ≥ 1.5 were considered differentially expressed. Analysis of variance (ANOVA) was conducted on proteins shared between ADSCs, dADSCs and adipocytes. Proteins with significantly changing expression ($P < 0.05$) were clustered on a heatmap using a correlation-based distance.

4.2.8 Gene Ontology annotation

Lists of differentially expressed proteins, proteins that were only identified reproducibly in a given sample, and proteins shared between adipocytes and dADSCs were submitted to the Database for Annotation, Visualisation and Integrated Discovery (DAVID) for functional categorisation (Huang, Sherman *et al.*, 2009a; Huang, Sherman *et al.*, 2009b). Proteins were categorised according to Biological Process at a cut-off of GO level 4, Cellular Component at a cut-off of GO level 3, and using the Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway analysis tool, with a defined minimum of two proteins in a category. To statistically determine over- or under-represented GO terms in each sample (category enrichment), submitted proteins were compared with the background of total human proteins in DAVID (a default feature of DAVID when no background list of proteins has been also submitted). GO categories that satisfied a test *P*-value < 0.05 in addition to a Benjamini-Hochberg value greater than the test *P*-value were then retained as significantly enriched. The default inclusion of a Benjamini-Hochberg test in DAVID corrects for multiple testing error that may increase the number of false positive identifications (Benjamini and Hochberg, 1995). The number of proteins in each category was then calculated as a percentage of the number of proteins successfully categorised by DAVID. Similar GO categories that satisfied the statistical criteria outlined above (enriched categories) were grouped using DAVID's Functional Annotation Clustering analysis at medium level stringency (default level) to reduce redundancy of terms. This analysis uses an EASE score, a modified Fisher Exact *P*-value to determine whether the combined terms in a cluster are enriched with respect to the total human proteome (higher EASE scores denote higher degree of enrichment). In instances where a GO term was not recommended on the European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) website to be used for direct annotation, the term was manually removed from DAVID output. For information on individual

proteins regarding their biological process, cell component or KEGG pathway involvement, DAVID's Functional Annotation Table tool was used.

4.3 Results

Using a label-free shotgun proteomics approach, the membrane proteome of ADSCs was characterised and compared with the corresponding *in vivo* (adipocytes) and *in vitro* (dADSCs) adipogenically differentiated products. The work was carried out in biological triplicate, using liposuction samples from the same three individuals employed for the membrane glycosylation analysis in Chapter 3.

4.3.1 Data quality

Difficulties were encountered in obtaining equal sample loading in each lane. To obtain enough protein for analysis, the lanes on two SDS-PAGE gels were combined, with the exception of lanes containing Adipocyte-1 and Adipocyte-3 membrane proteins that are seen to be heavily loaded (Figure 4.1). Whilst the protein content had been quantified using the Bradford assay (25 µg per sample), the inaccuracy and inconsistency in sample loading, as seen in both gels, may be due to interference in the Bradford assay by the high remaining lipid content of the adipocytes as indicated by the streaking in the gel, problems with protein solubility in 3 M urea, or incompatibility of Triton X-114 with the Bradford reagent.

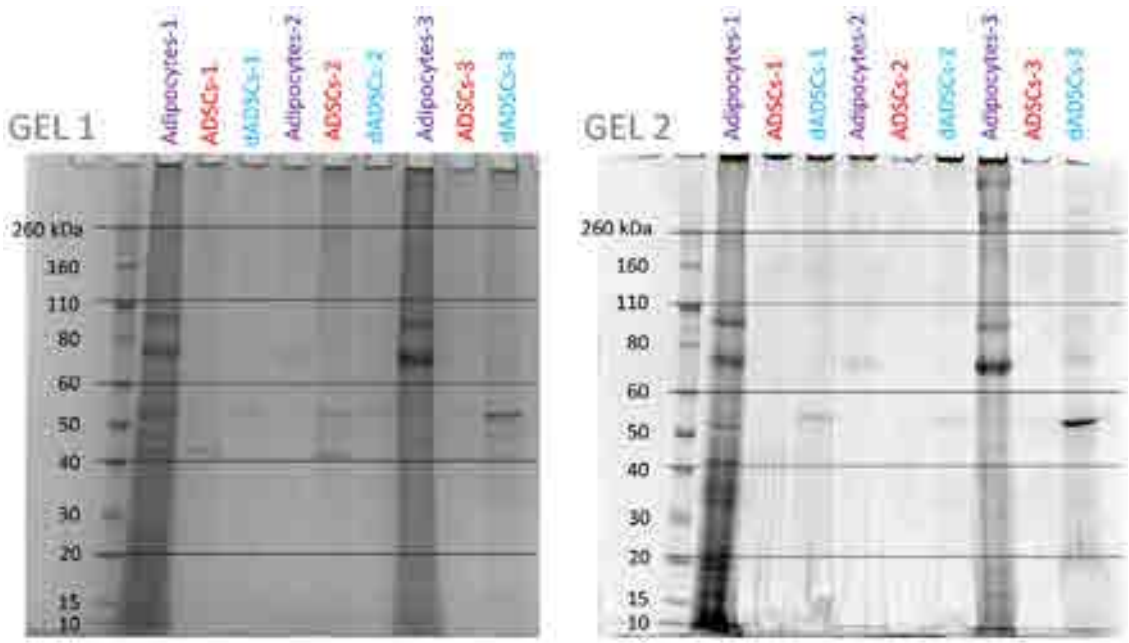


Figure 4.1 SDS-PAGE gels of separated membrane proteins from adipocytes, ADSCs and dADSCs from three individuals. The broken lines indicate where the gels were cut horizontally into molecular mass fractions.

The six molecular mass fractions from each lane were subjected to proteomic analysis. Data quality was first confirmed by plotting log-transformed NSAF values of the identified proteins as box and kernel density plots. The alignment of box plot medians (Figure 4.2a) and histograms of replicates from the same sample tracking close together (Figure 4.2b) indicated that data distribution was approximately normal. The greatest variation was seen in the histograms of dADSC replicates.

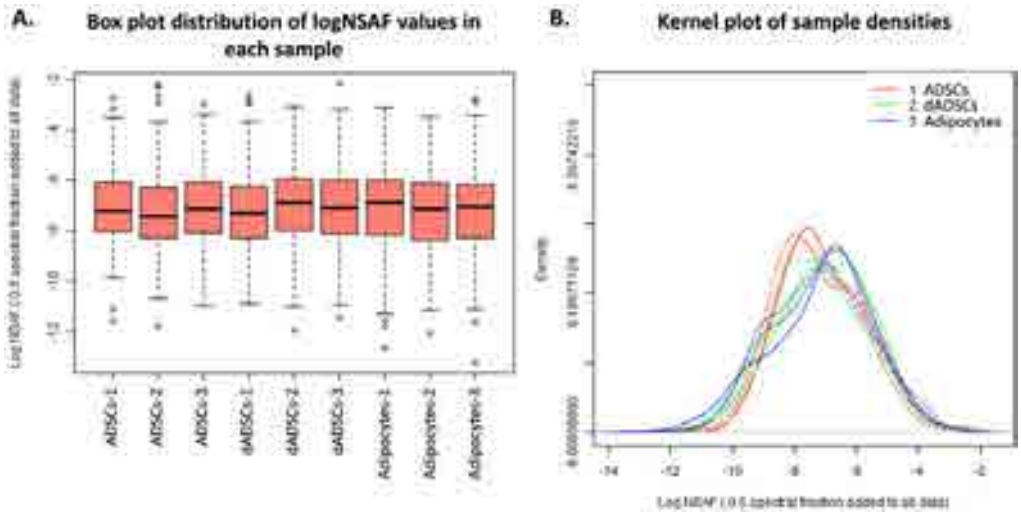


Figure 4.2 Data was normalised by log transformation of NSAF values and plotted in box (**A**) and kernel density plots (**B**) to check the distribution as approximately normal.

4.3.2 Membrane proteomics results

In total, 413 proteins were identified, at an overall protein FDR of 1.21 %. Reversed hits were removed from the list to produce a final list of 408 proteins across the three types of adipose-derived cell samples. Specifically, 175 proteins were identified reproducibly in ADSCs, 237 proteins in dADSCs and 264 proteins in adipocytes. A summary of the number of proteins and peptides identified and false discovery rates is shown in Table 4.1. The full list of proteins identified and their NSAF and SpC information can be found in Table S4.4 and S4.5 respectively (in accompanying DVD).

Table 4.1 Summary of number of proteins and peptides reproducibly identified

Sample	Number of proteins common to the three replicates	Number of peptides identified			Average no Peptides ± SD	Protein FDR (%)	Peptide FDR (%)
		Rep 1	Rep 2	Rep 3			
ADSCs	175	2138	2591	1772	2167 ± 410	0.57	0.17
dADSCs	237	2014	2543	3984	2847 ± 1020	1.66	0.78
Adipocytes	264	6014	2836	3440	4097 ± 1688	0.38	0.14

'Rep' = Replicate

Of the 408 proteins identified, 271 (67.1 %) were categorised as general membrane proteins (all organelle membrane proteins, including the PM) by DAVID at GO level 3 (see Supplementary Table

S4.1 for all Cellular Component GO annotation results). More specifically, 36.4 % of total proteins were categorised as PM proteins, thus confirming their presence by the Triton X-114 membrane protein enrichment.

4.3.3 Distribution of proteins

An analysis of the distribution of reproducibly identified proteins between the three samples is summarised in Figure 4.3a. One hundred and seventeen proteins were detected only in adipocytes, 72 proteins were only in dADSCs and 34 only in ADSCs. The distribution of proteins that were categorised as PM proteins is shown in Figure 4.3b and shows more similarity across the three cell types compared with total proteins. Functional annotation of the proteins uniquely identified in each cell type (from Figure 4.3a) was conducted by submitting protein IDs to DAVID for GO annotation according to Biological Process, Cellular Component and KEGG pathways and the results from these analyses are presented in the following sections. GO categories found to be significantly enriched (over-represented in a sample with respect to the overall human proteome) are expressed as a percentage of the number of proteins in a category from the total number of proteins categorised. It is important to note not all proteins were categorised and some are categorised into multiple categories. In the following sections, the abbreviation 'Uniprot' denotes Uniprot accession number for a given protein, while the prefix 'hsa' denotes a KEGG pathway reference number.

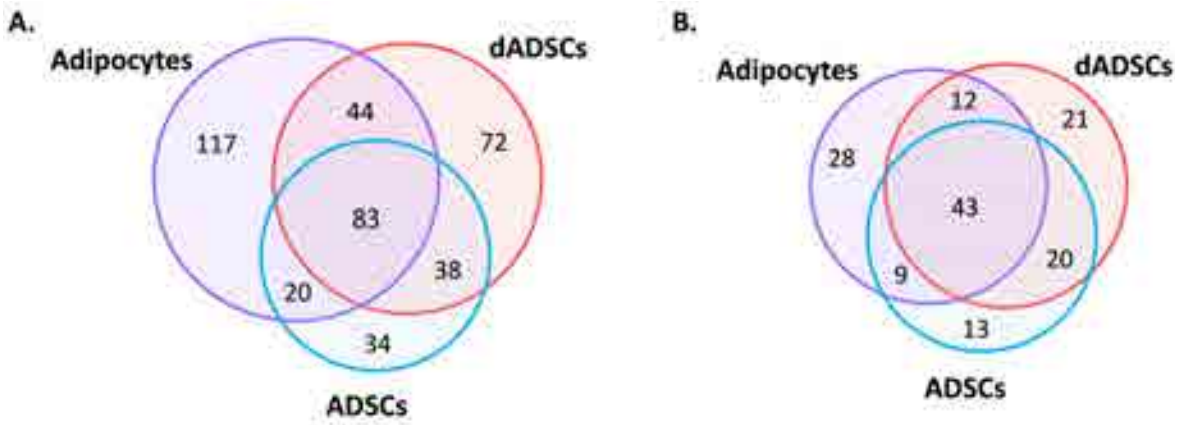


Figure 4.3 Venn diagram distribution of total proteins identified (**A**) and of proteins identified as PM proteins by DAVID GO annotation (**B**) in the three types of cells.

4.3.3.1 Proteins identified reproducibly in adipocytes alone

A total of 117 proteins were identified reproducibly only in adipocytes, and was the largest subset of proteins in the dataset (listed in Table S4.6, accompanying DVD). Eighty four of these proteins were successfully grouped into 51 significantly enriched biological processes and 39 clustered into 11 distinct groups (Figure 4.4a). Amongst these groups were several categories involving metabolic processes, particularly lipid and carbohydrate metabolic processes. One category of interest in the

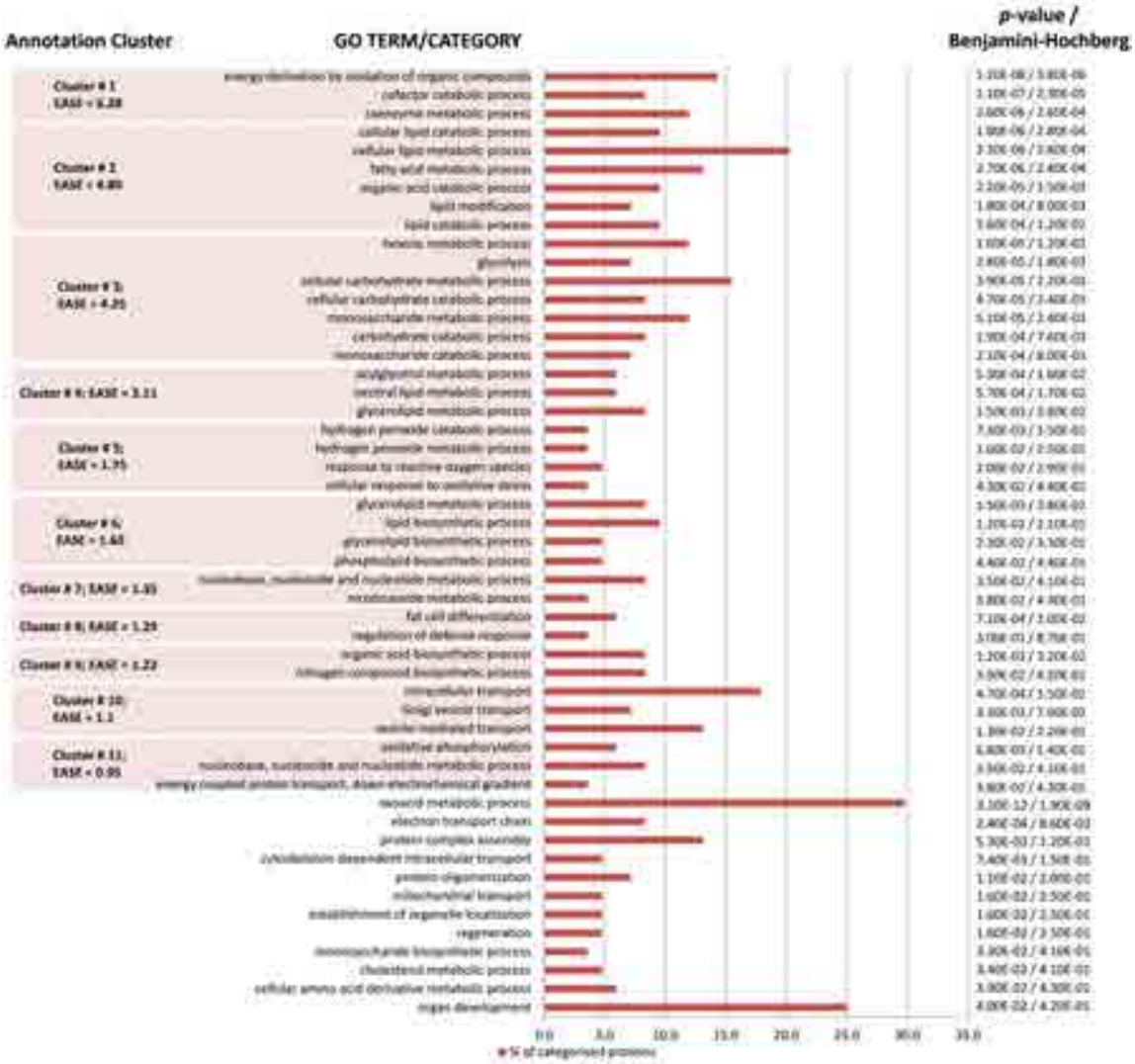
context of adipogenesis was fat cell differentiation. The five proteins in the mature adipocytes that grouped under fat cell differentiation were endoplasmic reticulum aminopeptidase 1/Adipocyte-derived leucine aminopeptidase (Uniprot: Q9NZ08), FABP4 (Uniprot: P15090), Glutathione peroxidase 1 (Uniprot: P07203), integrin alpha-6/CD49f (Uniprot: P23229), and laminin subunit alpha-4 (Uniprot: Q16363). It is interesting to note that FABP4 was amongst the proteins found by the proteomics approach in adipocytes only whilst mRNA transcript analysis detected FABP4 in both dADSCs and adipocytes (Chapter 2, Figure 2.14b). A comparison of protein and mRNA expression levels will be further discussed in Section 4.3.7.

The enrichment of several lipid metabolism-related processes (three out of 11 clusters) is unsurprising in adipocytes. However, several carbohydrate-related categories were also enriched: cellular carbohydrate metabolic process (15.5 % of proteins), cellular carbohydrate catabolic process (8.3 % of proteins), monosaccharide metabolic process (11.9 %), carbohydrate catabolic process (8.3 %), hexose metabolic process (11.9 %) and monosaccharide catabolic process (7.1 %). Enrichment of these processes may be indicative of the increased energy usage (catabolism) of adipocytes compared to ADSCs.

To determine the origin of the 117 proteins unique to adipocytes, proteins were categorised according to cellular component (Figure 4.4b). Of the 117 proteins, 113 were categorised into a total of 24 significantly enriched categories with 14 clustered into three groups. This analysis revealed that the most enriched categories were proteins derived from the mitochondria, which is in accordance with the enrichment of the metabolic processes described above. Furthermore, KEGG pathway analysis of these proteins (49 of the 117 were successfully categorised) indicated the TCA cycle (hsa00020) to be the most significantly enriched pathway with one of the highest percentage of proteins categorised (18.4 %; Figure 4.4c). Several other metabolic pathways were enriched, including fatty acid metabolism (hsa00071; 10.2 % of proteins) and glycolysis/gluconeogenesis (hsa00010; 10.0 % of proteins).

Other enriched pathways of interest in the context of *in vivo* adipogenesis were regulation of the actin cytoskeleton (hsa04810; 16.3 % of proteins), and the related pathways of ECM-receptor interaction (hsa04512; 12.0 % of proteins) and focal adhesion (hsa04510; 18.0 % of proteins). Interestingly, neurological disease pathways were also enriched, namely those associated with Parkinson's, Huntington's and Alzheimer's diseases.

A


Annotation Cluster
GO TERM/CATEGORY
p-value / Benjamini-Hochberg
% of categorised proteins

B

Annotation Cluster
GO TERM/CATEGORY
p-value / Benjamini-Hochberg
organelle inner membrane
mitochondrial envelope
mitochondrial membrane
organelle envelope
organelle membrane
cytoplasm
intracellular
intracellular organelle
intracellular membrane-bounded organelle
laminin complex
basement membrane
basal lamina
proteinaceous extracellular matrix
extracellular matrix
organelle outer membrane
outer membrane
membrane
proton-transporting ATP synthase complex
respiratory chain
cell leading edge
extrinsic to membrane
cell projection
organelle lumen
proton-transporting two-sector ATPase complex
Cluster # 1
EASE = 11.81
Cluster # 2
EASE = 6.4
Cluster # 3
EASE = 2.06
0.0
20.0
40.0
60.0
80.0
100.0
% of categorised proteins

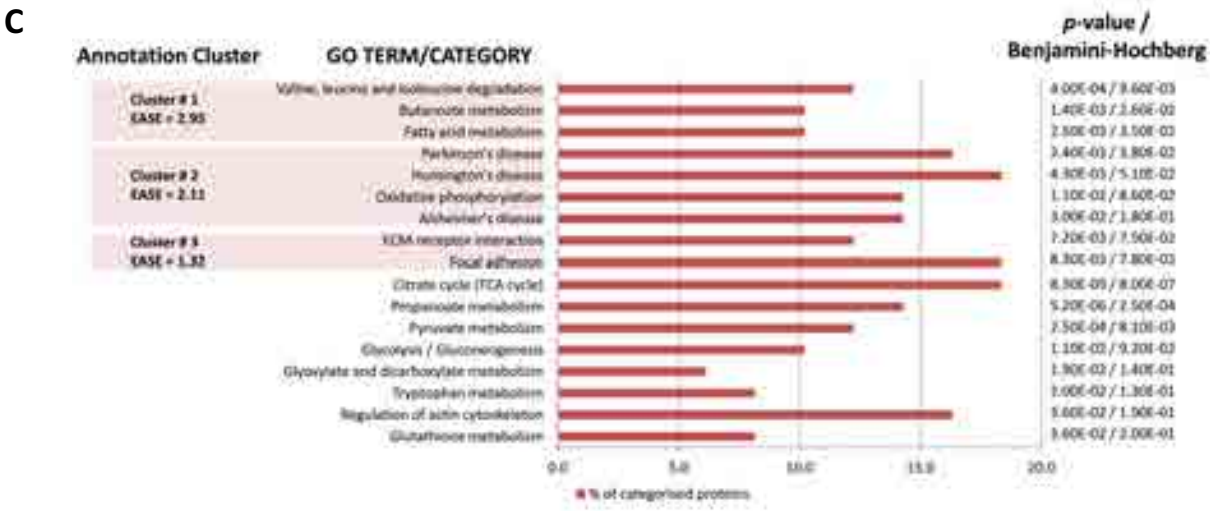


Figure 4.4 Functional categorisation of proteins detected only in mature adipocytes. Proteins were categorised according to biological process (**A**), cellular component (**B**) and KEGG pathway analysis (**C**). Similar GO terms are clustered together and listed in order of their enrichment (EASE score) with respect to the total human proteome. GO terms that were not included in any annotation clusters are listed in order of enrichment (increasing *P*-value).

4.3.3.2 Proteins identified reproducibly in dADSCs alone

A total of 72 proteins were identified reproducibly only in dADSCs (listed in Table S4.6, accompanying DVD), of which 32 were categorised by DAVID according to biological process into nine categories (Figure 4.5a). These categories differed substantially from those identified only in adipocytes, with nuclear processes involving DNA packaging such as nucleosome assembly, protein-DNA assembly, macromolecular assembly and chromatin organisation showing the highest enrichment and clustering together. Unlike the subset of proteins found only in mature adipocytes, no lipid-related processes were significantly enriched in the differentiated stem cells. Only one biological process was shared between proteins found only in dADSCs and those found only in adipocytes, i.e. oxoacid metabolic process (25.0 % of proteins in dADSCs and 29.8 % in adipocytes) which encompasses reactions involving compounds containing oxygen, at least one hydrogen bound to oxygen, and one other element.

The proteins found only in dADSCs were categorised into 22 cellular component categories, with 15 terms clustered into five groups (Figure 4.5b). Nuclear-related proteins were an enriched category in accordance with the results seen above in the biological process functional annotation. The second group of proteins belonging to a specific organelle were mitochondrial protein-related categories which were also enriched in proteins identified only in adipocytes.

In the categorisation of proteins according to KEGG pathways (Figure 4.5c), only 13 of the total 72 proteins were categorised, and like the proteins found only in adipocytes, proteins involved in TCA cycle were significantly enriched (hsa00020; 23.1 % of proteins). There is some indication therefore of increased energy demands in dADSCs similar to that seen in adipocytes.

A second interesting KEGG pathway was significantly enriched in the differentiated ADSCs; the SNARE interactions in vesicular transport pathway (hsa04130; 23.1 % of proteins). Proteins categorised into this pathway were syntaxin-7 (Uniprot: O15400), vesicle-associated membrane protein 2 (VAMP2; Uniprot: P63027) and vesicle-associated membrane protein 3 (VAMP3; Uniprot: Q15836). While this pathway did not appear to be enriched in the analysis of proteins found only in adipocytes, SNARE proteins have been implicated in vesicle trafficking in adipocytes (Olson, Knight *et al.*, 1997).

While there is some similarity in role and location of proteins found only in adipocytes and only in dADSCs (enrichment of mitochondrial proteins), overall the processes and locations of each list of identified proteins were quite different, with significant enrichment of nuclear proteins and particularly nucleic processes in dADSCs. The similarities between adipocytes and dADSCs will be discussed in the following section.

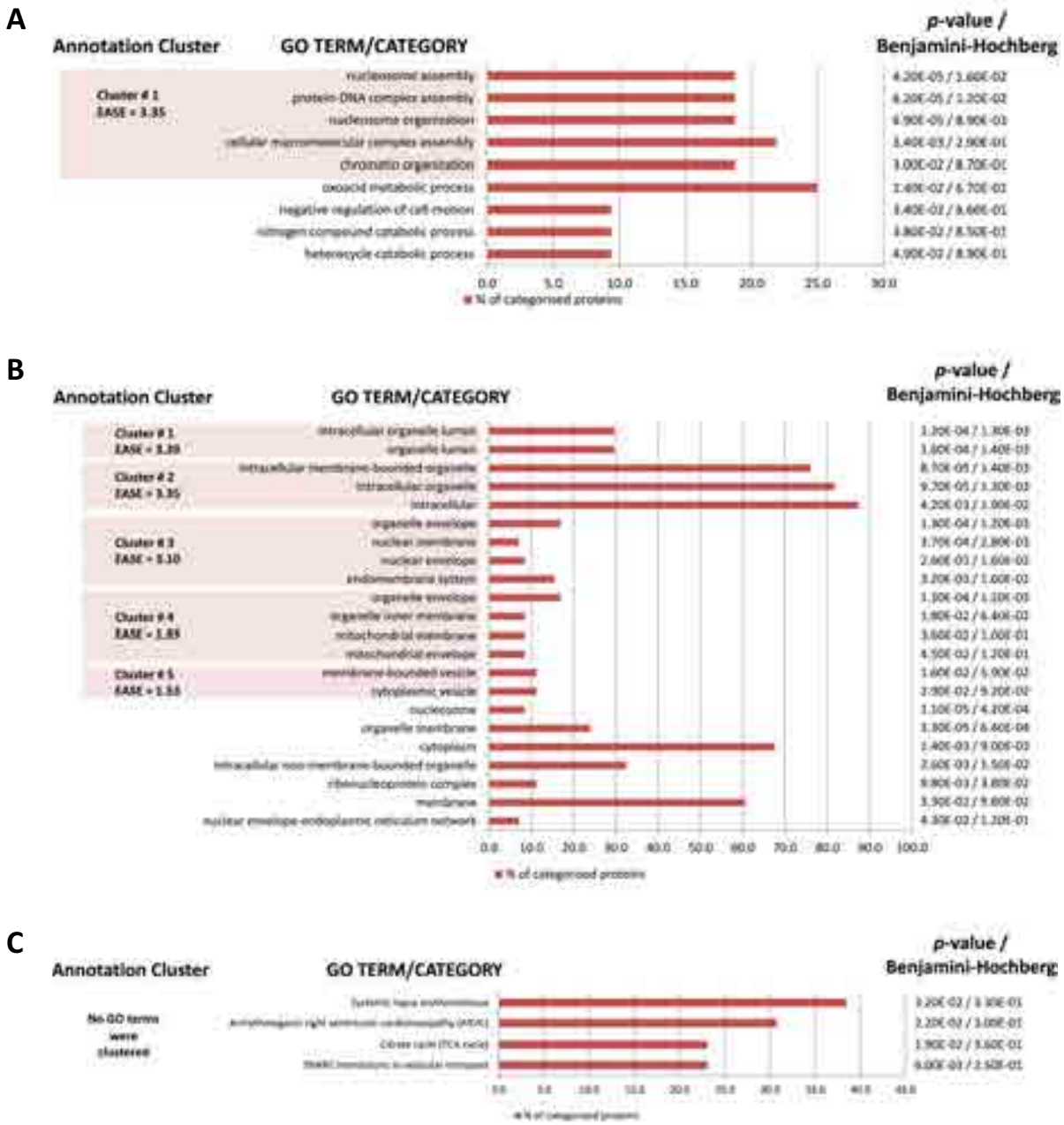


Figure 4.5 Functional categorisation of proteins detected only in dADSCs. Proteins were categorised according to biological process (**A**), cellular component (**B**) and KEGG pathway analysis (**C**). Similar GO terms are clustered together and listed in order of their enrichment (EASE score) with respect to the total human proteome. GO terms that were not included in any annotation clusters are listed in order of enrichment (increasing *P*-value).

4.3.3.3 Proteins expressed in cell membrane of both in vivo *and* in vitro *models of adipogenesis*

Forty four proteins were common to both native mature adipocytes and differentiated ADSCs (listed in Table S4.6, accompanying DVD). Since they were not found in the undifferentiated stem cell ADSCs, this indicates that they are expressed upon adipogenesis in both models. Twenty one of these proteins were successfully categorised by DAVID into nine biological processes (Figure 4.6a), the most enriched of which included various mitochondrial-related processes (each process ranging from

19-28 % of categorised proteins). Adipogenic differentiation related processes were also enriched with cellular lipid metabolic process up-regulated, and regulation of cell morphogenesis and cell-matrix adhesion which were grouped together.

Categorisation according to cellular component also revealed mitochondrial-related proteins to be the most significantly enriched (Figure 4.6b) which is in agreement with the major biological processes identified in both mature adipocytes and dADSCs. Together with the enrichment of the mitochondrial KEGG pathways of oxidative phosphorylation, fatty acid metabolism, and valine, leucine and isoleucine degradation (Figure 4.6c), the results indicated the enrichment of mitochondrial proteins and processes upon adipogenesis both *in vivo* and *in vitro*. Also similar to the analysis of proteins found only in adipocytes, categories relating to cell shape and organisation of the ECM (ECM-receptor interaction: 22.2 % of categorised proteins, and focal adhesion: 33.3 % of categorised proteins), were also enriched in both differentiated cell types, both of which are characteristic of adipogenesis.

Again, proteins typical of degenerative neural Alzheimer's, Parkinson's and Huntington's diseases were also enriched in these cell types. An adipogenic pathway of relevance, the PPAR signalling pathway, was also enriched however without statistical significance (P = 5.50E-02). Overall, three main features of adipogenesis were observed in the 44 proteins shared between mature adipocytes and differentiated ADSCs, i.e. processes involving the ECM, mitochondrial processes, and metabolic lipid processes.

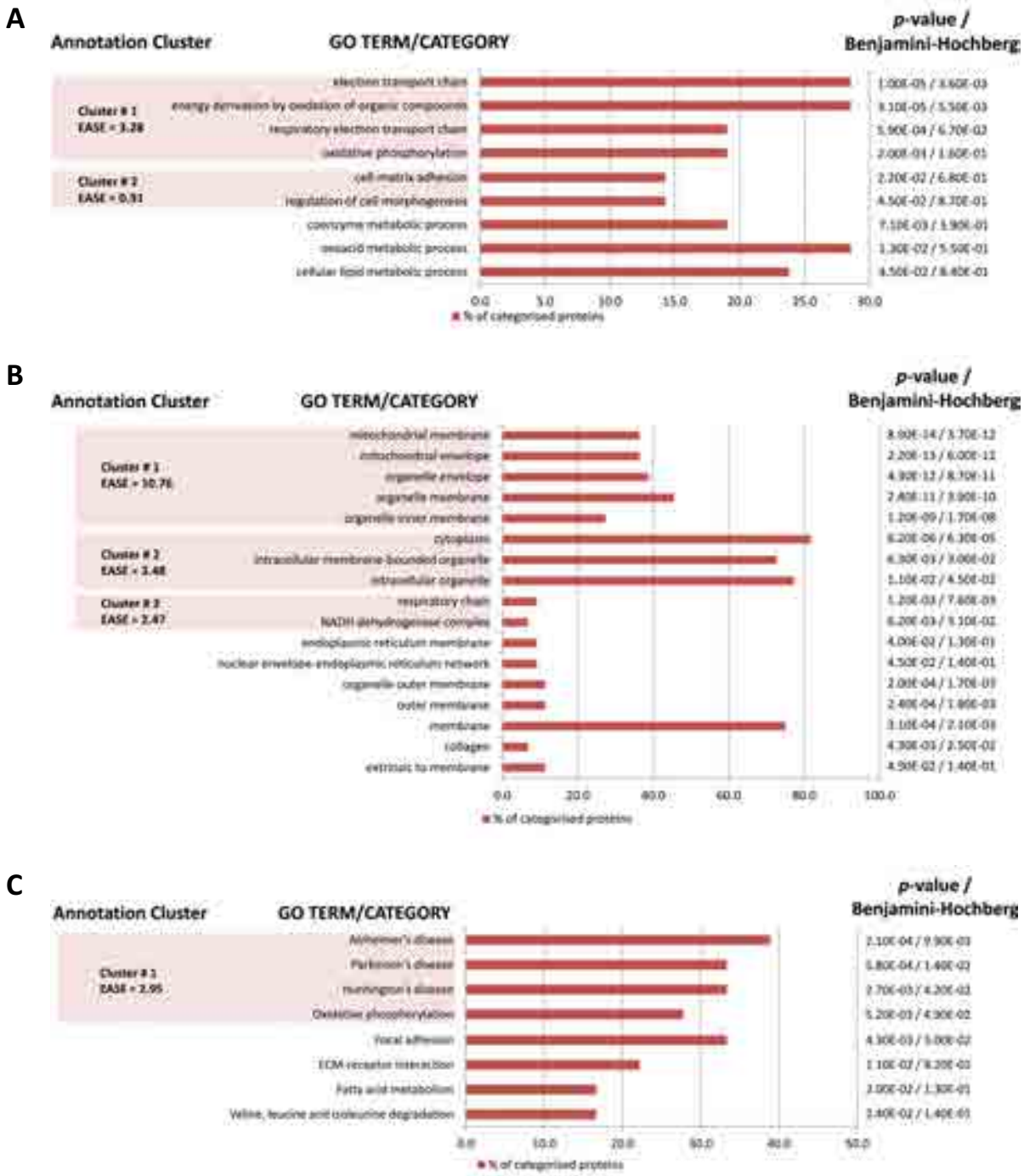


Figure 4.6 Functional categorisation of proteins identified in both dADSCs and mature adipocytes. Proteins were categorised according to biological process (**A**), cellular component (**B**) and KEGG pathway analysis (**C**). Similar GO terms are clustered together and listed in order of their enrichment (EASE score) with respect to the total human proteome. GO terms that were not included in any annotation clusters are listed in order of enrichment (increasing *P*-value).

4.3.3.4 Proteins identified reproducibly in ADSCs alone

Thirty four proteins were identified only in the stem cell ADSCs, which were the smallest number of 'unique' proteins (listed in Table S4.6, accompanying DVD). Of these, only 13 were categorised into

eight biological processes with five related processes clustered together (Figure 4.7a). The relatively small number of categorised proteins highlights a disadvantage of using GO annotation on small lists of proteins. Processes involving nucleotide metabolism dominated the list of categories found to be significantly enriched in this group of proteins, as was seen in the proteins identified only in dADSCs. Complementing this was the enrichment of ribo-nucleoprotein complex proteins (Figure 4.7b), a category that was also significantly enriched in proteins identified in dADSCs alone.

Regulation of cell morphology and anatomical structure morphogenesis were also enriched in the undifferentiated ADSCs (Figure 4.7a), most likely due to the cells growing and dividing during culture. Extracellular categories were also present (Figure 4.7b), namely collagen, basement membrane and extracellular matrix proteins, presumably to mediate adhesion to the cell culture vessel. In the KEGG pathway analysis, enrichment of extracellular matrix-receptor interaction (hsa04512) is consistent with the latter interpretation. The proteins grouped into this category were collagen type IV, alpha-1 chain (Uniprot: P02462), collagen type IV, alpha-2 chain (Uniprot: P08572), collagen type IV, alpha-5 chain (Uniprot: P29400) and collagen type V, alpha-1 chain (Uniprot: P20908).

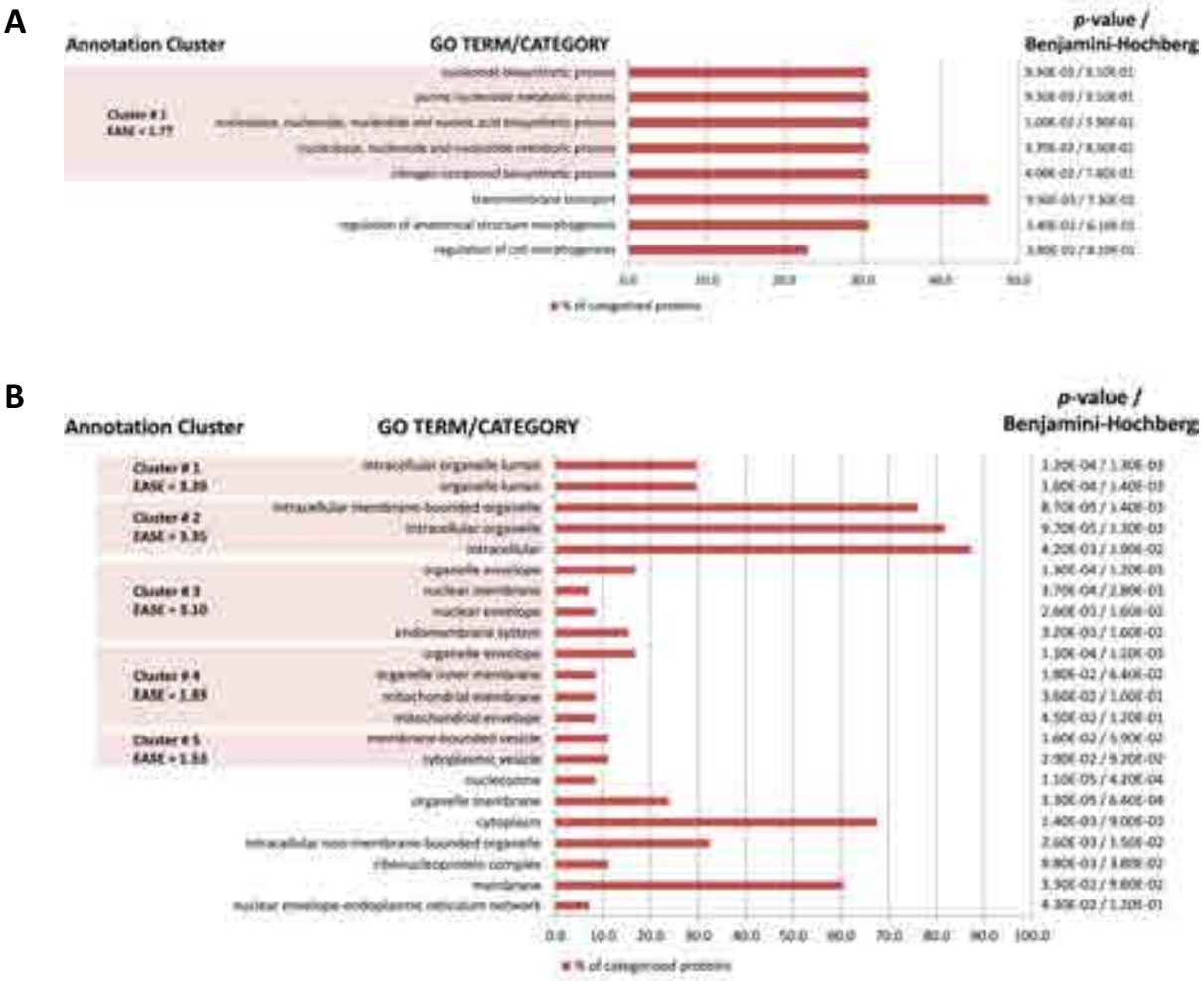


Figure 4.7 Functional categorisation of proteins identified only in ADSCs. Proteins were categorised according to biological process (**A**), cellular component (**B**). Similar GO terms are clustered together and listed in order of their enrichment (EASE score) with respect to the total human proteome. GO terms that were not included in any annotation clusters are listed in order of enrichment (increasing *P*-value).

4.3.4 Proteins present in all three cell types

Eighty two proteins shared between the mature adipocytes, ADSCs and the corresponding differentiated ADSCs were further investigated to identify differentially expressed proteins and, specifically, proteins that exhibited similar expression in adipogenic *in vivo* and *in vitro* product cells that were different to that of their expression in ADSCs. The 25 proteins found to be significantly changing ($P < 0.05$) were visualised on a heatmap and clustered using an ANOVA correlation-based distance (Figure 4.8). To identify PM proteins amongst this list, proteins were submitted to DAVID for cell component classification at GO level 3. The majority (72%) of these 25 common proteins were identified as membrane proteins, with 60 % categorised as PM proteins (shown in blue font in Figure 4.8) indicating significant changes occur on the cell surface during adipogenesis.

A number of cytoskeletal proteins were also identified, namely different types of actins that were down-regulated in dADSCs and adipocytes, and tubulins up-regulated in adipocytes, highlighting their role in the dynamic changes in cell shape occurring during differentiation. Two proteins in particular displayed higher abundance in both adipogenic products compared to the progenitor ADSCs indicating that their expression during *in vitro* adipogenesis of ADSCs mirrored that of *in vivo* adipogenesis into adipocytes. These proteins were perilipin-4 (Uniprot: Q96Q06), an adipogenic marker, and a mitochondrial protein trifunctional enzyme subunit alpha (Uniprot: P40939) (Figure 4.8). The expression pattern of mitochondrial protein trifunctional enzyme subunit alpha is perhaps representative of the significant increase in mitochondrial proteins in general seen in both *in vitro* and *in vivo* adipogenic product cells as shown in Sections 4.3.3.1-4.3.3.3. A third protein, myosin-9 (Uniprot: O00159), displayed a lower abundance in dADSCs and adipocytes when compared to ADSCs. Myosin-9 is classed as an unconventional myosin and has recently been demonstrated to be involved in insulin sensitive glucose uptake (Munnich, Taft *et al.*, 2014). Otherwise, the majority of proteins showed more similarity in abundance between dADSCs and ADSCs than adipocytes, indicating that the protein abundance changing in *in vitro* adipogenesis does not closely reflect the *in vivo* generated mature adipocyte, as clearly shown by broad inspection of the heatmap (Figure 4.8).

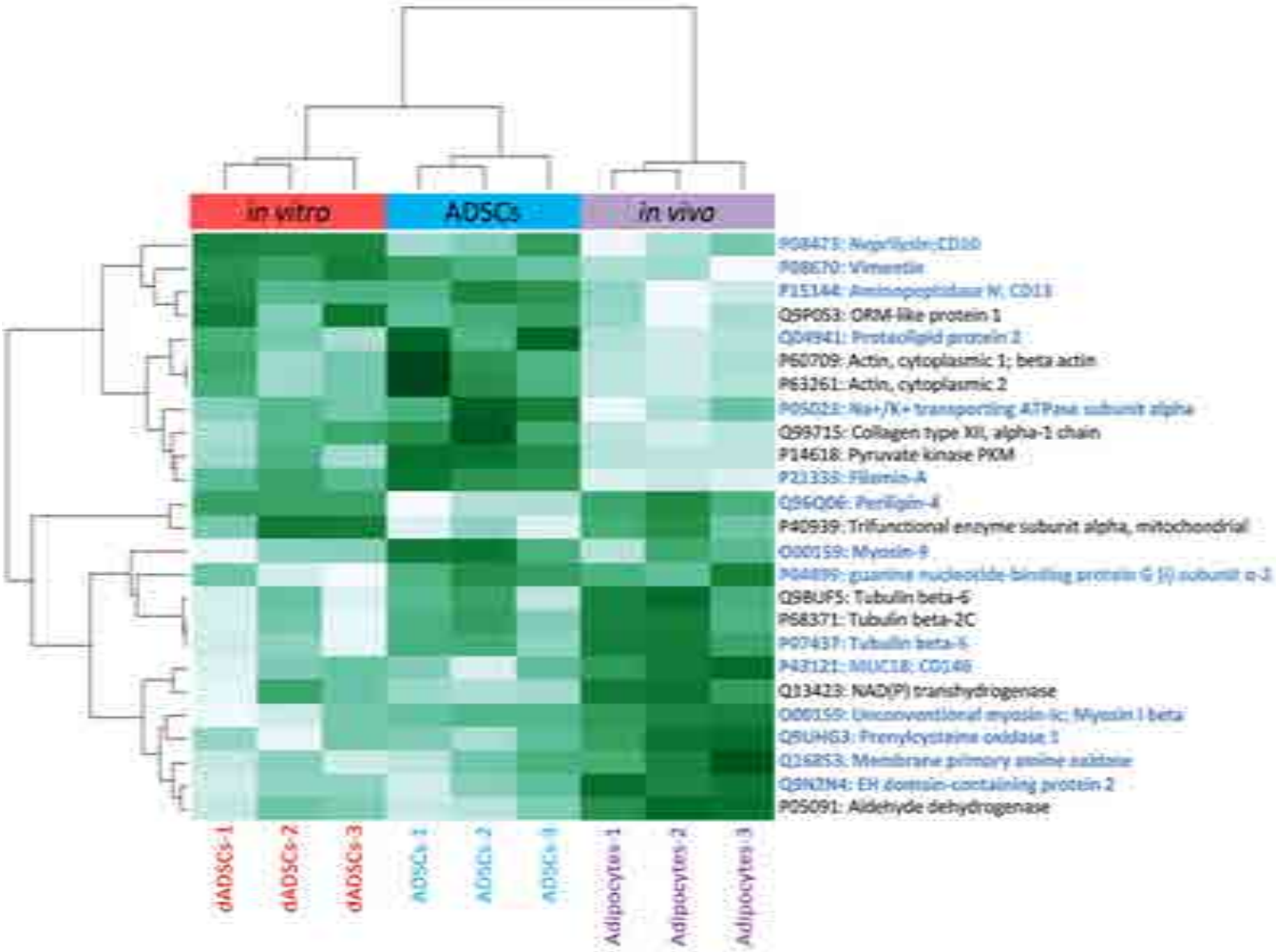


Figure 4.8 Proteins shared between ADSCs, dADSCs and adipocytes exhibiting differential expression ($P < 0.05$). Proteins are clustered by logNSAF values using a correlation based distance, and those identified as PM proteins by DAVID are highlighted with blue text. Green and white tones indicate up- and down-regulation respectively.

4.3.5 Pair-wise comparisons of ADSCs with their differentiated adipogenic product cells

For a broad overview of the changes in protein abundance occurring between ADSCs and their *in vitro* and *in vivo* adipogenic products, correlation plots of logNSAF ratios for each model of adipogenesis were generated (Figure 4.9). Proteins with differential expression (Student's *t*-test $P < 0.05$ and a fold change greater than 1.5) are marked in dark blue circles and up- and down-regulated proteins are shown above and below the diagonal respectively. The number of proteins differentially expressed upon adipocytes is higher than that observed in dADSCs. In addition to this, larger deviations in all protein ratios are also seen in adipocytes with many proteins further from the diagonal than in dADSCs. This indicates that ADSCs undergo a greater number and degree of change in protein abundance during *in vivo* adipogenesis than the *in vitro* induced differentiation.

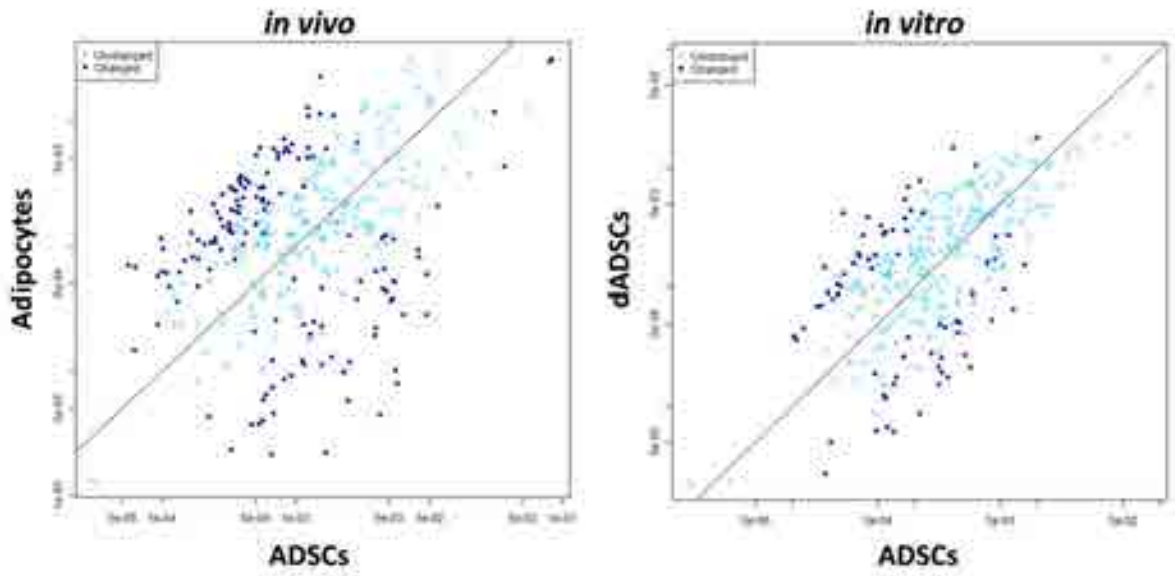


Figure 4.9 Correlation plots of logNSAF ratios from pair-wise comparison of ADSCs and adipocytes, and ADSCS and dADSCs. Differentially expressed proteins ($P < 0.05$) are marked with a dark blue circle while unchanged proteins are marked with light blue.

To identify differentially expressed proteins during *in vivo* and *in vitro* differentiation, dADSCs and adipocytes were separately compared to their progenitor ADSCs. A Venn diagram comparison of the two models and summary of results is shown in Figure 4.10 which is derived from Figure 4.3a. 121 proteins were found to be common between dADSCs and ADSCs, while 103 were common between adipocytes and ADSCs. Of the 121 common proteins between dADSCs and ADSCs, 15 (12.4 %) were found to be down-regulated in dADSCs, five proteins (4.1 %) were up-regulated, whilst 101 (83.5 %) were unchanged. In adipocytes, the greatest percentage of proteins were also those that were unchanged (72.8 %), while 17 (16.5 %) proteins were found to be down-regulated in adipocytes and nine proteins (8.7 %) up-regulated. A list of unchanged proteins in dADSCs and adipocytes can be

found in Supplementary Tables S4.2 and S4.3 respectively while the following sections will discuss the up and down-regulated proteins in each model of differentiation.

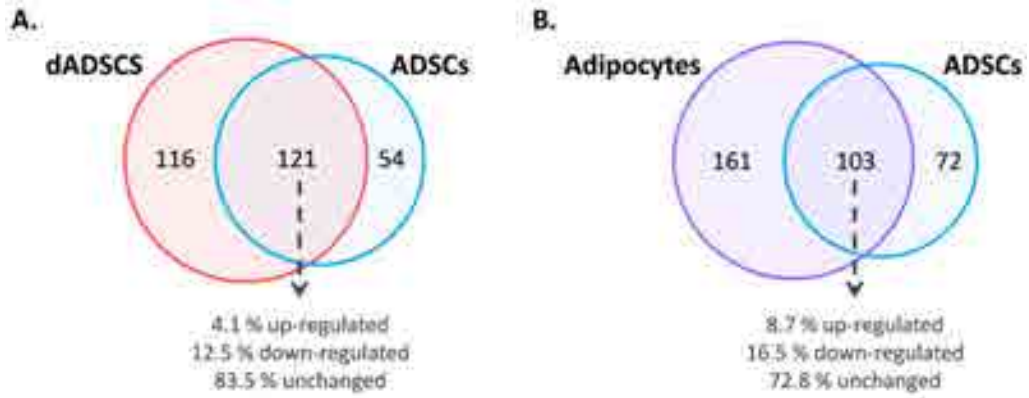


Figure 4.10 Comparison of relative abundance of proteins found in dADSCs and ADSCs (**A**), and in adipocytes and ADSCs (**B**). Percentages of up- and down-regulated proteins refer to changes in dADSCs and adipocytes in comparison to ADSCs

4.3.5.1 Proteins down-regulated in adipogenesis: in vitro *vs.* in vivo

For functional annotation, the 15 proteins found to be down-regulated in dADSCs after *in vitro* differentiation from ADSCs were categorised according to biological process and cell component using DAVID. Eight biological processes down-regulated during *in vitro* adipogenesis are shown in Figure 4.11a. The most significantly enriched category that was down-regulated in the *in vitro* differentiation was cell-substrate junction assembly (27.3 % of categorised proteins) which contained the proteins fibronectin (Uniprot: P02751), integrin alpha-5/CD49e (Uniprot: P08648) and talin-1 (Uniprot: Q9Y490). Several categories involving cytoskeletal rearrangement were also down-regulated, which may be due to a halt in cell division and proliferation upon induction of adipogenesis.

The down-regulated proteins in *in vitro* adipogenesis were categorised into seven significantly enriched cellular component categories displayed in Figure 4.11b. Importantly, a significantly down-regulated group of proteins were found to be PM located (P = 3.10E-04), again indicating changes occurring on the cell surface in response to differentiation. While proteins located in the cytoplasm comprised the largest percentage, the down-regulation of components such as extracellular matrix and collagen are further indicative of the cell's rearrangement during differentiation and reflect the loss of adhesion of the ADSCs to the culture plate upon differentiation. This observation is supported by the KEGG pathway analysis (Figure 4.11c), where only seven proteins were successfully categorised into two significantly enriched pathways; all seven were grouped under focal adhesion (hsa04510; P = 1.40E-06), and four of these also grouped under ECM-receptor interaction (hsa04512; P = 6.50E-04).

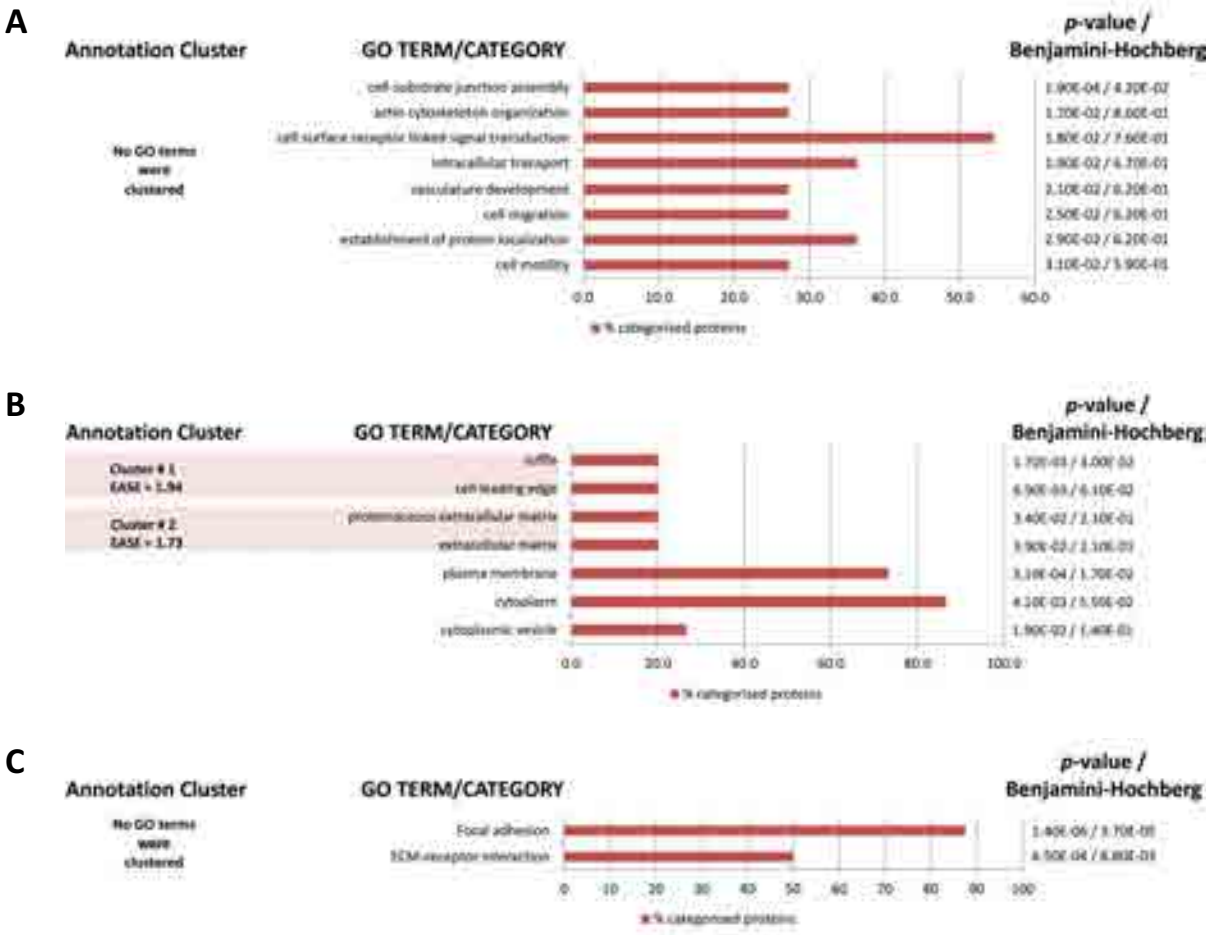


Figure 4.11 Functional categorisation of proteins down-regulated during *in vitro* adipogenic differentiation compared to precursor ADSCs. Proteins were categorised according to biological process (**A**), cellular component (**B**) and KEGG pathways (**C**). Similar GO terms are clustered together and listed in order of their enrichment (EASE score) with respect to the total human proteome. GO terms that were not included in any annotation clusters are listed in order of enrichment (increasing *P*-value).

For the 17 down-regulated proteins in the native, mature adipocytes that were generated *in vivo* from ADSCs, no biological processes were found to be enriched using GO annotation. All proteins were however categorised into intracellular components by cellular component analysis (Figure 4.12) and, in contrast to proteins down-regulated by *in vitro* differentiation, no changes were observed in this study for ECM-related processes or PM proteins, further indicating a difference in the two models of adipogenesis.

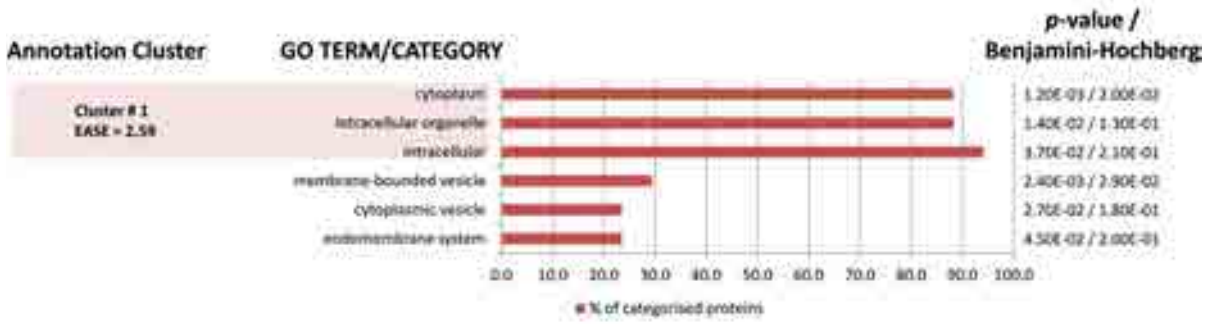


Figure 4.12 Functional categorisation of proteins down-regulated during *in vivo* adipogenic differentiation compared to precursor ADSCs. Proteins were categorised according to cellular component and similar GO terms are clustered together and listed in order of their enrichment (EASE score) with respect to the total human proteome. GO terms that were not included in any annotation clusters are listed in order of enrichment (increasing *P*-value).

When submitted for KEGG pathway analysis, only the focal adhesion pathway which involves cell-matrix adhesion proteins was found to be down-regulated (hsa04510; *P* = 4.70E-02) in native adipocytes. This correlates with the down-regulation of this pathway also seen in dADSCs, although the proteins categorised into this pathway in each adipogenesis model differed. This suggests that while the two models of adipogenesis may not mirror each other directly at the individual protein level, the focal adhesion pathway is changing in both models of differentiation.

To identify proteins exhibiting the greatest changes in abundance, proteins down-regulated during *in vitro* and *in vivo* adipogenesis were listed in order of greatest fold-change (see Table 4.2). The fold-changes in protein abundance seen *in vivo* were greater than *in vitro* differentiation, which was also seen in the logNSAF ratio plots (Figure 4.9). For example, the greatest difference in protein abundance seen in adipocytes compared with ADSCs was in collagen type XII, alpha-1 chain which had a 42-fold down-regulation, whilst in dADSCs the greatest change compared with the ADSCs was seen in Filamin-C with 8.4-fold down-regulation. The remodelling of the PM during *in vitro* adipogenesis is evident, with 11 out of the 15 down-regulated proteins categorised as PM located. Integrin alpha-5/CD49e is a fibronectin receptor which was down-regulated in the *in vitro* differentiation system along with fibronectin itself. Integrin alpha-5/CD49e was in fact not identified at all in the mature adipocytes, while interestingly, fibronectin was unchanged. Different heat shock protein HSP-90 subunits were down-regulated in both models, as were filamins (also called actin binding proteins) which is in line with the changes seen in actin proteins and cell remodelling proteins. Overall the results from this section shed light on the types of proteins and processes down-regulated in adipogenesis, many of which appear common to both *in vivo* and *in vitro* systems.

Table 4.2 Proteins significantly down-regulated during *in vitro* and *in vivo* adipogenic differentiation based on greatest fold change.

Uniprot Accession	Description	Fold Change	*P*-value
PROTEINS DOWN-REGULATED *IN VITRO*			
Q14315 *	Filamin-C	-8.4	2.56E-03
P02751 *	Fibronectin	-7.9	1.08E-02
Q9Y490 *	Talin-1	-6.4	3.52E-02
P14618	Pyruvate kinase PKM	-5.6	1.25E-02
P35579 *	Myosin-9	-5.1	3.34E-02
P08648 *	Integrin alpha-5; CD49e; Fibronectin receptor subunit alpha	-4.5	2.86E-02
P08123 *	Collagen type I, alpha-2 chain	-4.2	1.89E-02
Q9NZM1 *	Myoferlin; Fer-1-like protein 3	-3.8	1.34E-02
P02452 *	Collagen type I, alpha-1 chain	-3.4	3.65E-02
P07900	Heat shock protein HSP 90-alpha	-2.6	1.09E-02
P21333 *	Filamin-A	-2.4	4.43E-02
Q01082 *	Beta II spectrin	-2.4	2.66E-02
Q14697	Alpha-glucosidase 2	-2.3	3.29E-02
P04899 *	Guanine nucleotide-binding protein G(i) subunit alpha-2	-2.2	3.33E-02
P68104	Elongation factor 1-alpha 1	-1.7	1.24E-02
PROTEINS DOWN-REGULATED *IN VIVO*			
Q99715	Collagen type XII, alpha-1 chain	-42.0	4.86E-03
P21333	Filamin-A	-17.7	3.68E-04
P15144	Aminopeptidase N; CD13	-16.0	1.84E-02
P14618	Pyruvate kinase PKM	-13.9	1.04E-04
P08670	Vimentin	-8.3	4.24E-02
P05023 *	Sodium/potassium-transporting ATPase subunit alpha-1	-6.3	3.36E-02
Q04941	Proteolipid protein 2	-5.5	9.56E-03
P08238	Heat shock protein HSP 90-beta	-4.8	1.82E-03
O75396	Vesicle-trafficking protein SEC22b	-3.9	6.85E-03
P08962 *	CD63 antigen; Lysosomal-associated membrane protein 3	-3.8	3.79E-02
Q14974	Importin subunit beta-1	-2.7	1.51E-02
P63261	Actin, cytoplasmic 2	-2.6	1.13E-02
P60709	Actin, cytoplasmic 1	-2.6	1.58E-02
P04264 *	Keratin, type II cytoskeletal 1	-2.5	4.95E-02
Q14204	Cytoplasmic dynein 1 heavy chain 1	-2.4	2.29E-02
P14625	Endoplasmin; Heat shock protein 90 kDa beta member 1	-2.0	3.86E-02
O75844 *	CAAX prenyl protease 1 homolog	-1.8	2.85E-02

* PM protein as categorised by DAVID

4.3.5.2 Proteins up-regulated in adipogenesis: in vitro *vs.* in vivo

Of the five proteins up-regulated during *in vitro* adipogenesis, four proteins were successfully categorised by DAVID according to biological process, however only two proteins were categorised into processes with $P < 0.05$. These two proteins were mitochondrial proteins; very long-chain specific acyl-CoA dehydrogenase (Uniprot: P49748) and trifunctional enzyme subunit alpha (Uniprot: P40939). Both were grouped as part of three different biological processes shown in Figure 4.13a, two of which are lipid related processes.

All five proteins up-regulated in *in vitro* adipogenesis could however be categorised according to cellular component, with ten categories in total, nine of which form three clusters (Figure 4.13b). Overall the analysis revealed that four of the five proteins are mitochondrial, and this category was also the most up-regulated (Figure 4.13b). When submitted for KEGG pathway analysis, the two mitochondrial proteins categorised above in biological process were categorised into the fatty acid metabolism pathway (hsa00071; with P = 1.60E-02). Taken altogether, the results indicate up-regulation of lipid metabolic pathways and mitochondrial protein expression during *in vitro* adipogenesis; an observation which is validated by the types of proteins common to dADSCs and adipocytes (Figure 4.6), and the proteins identified in adipocytes alone (Figure 4.4).

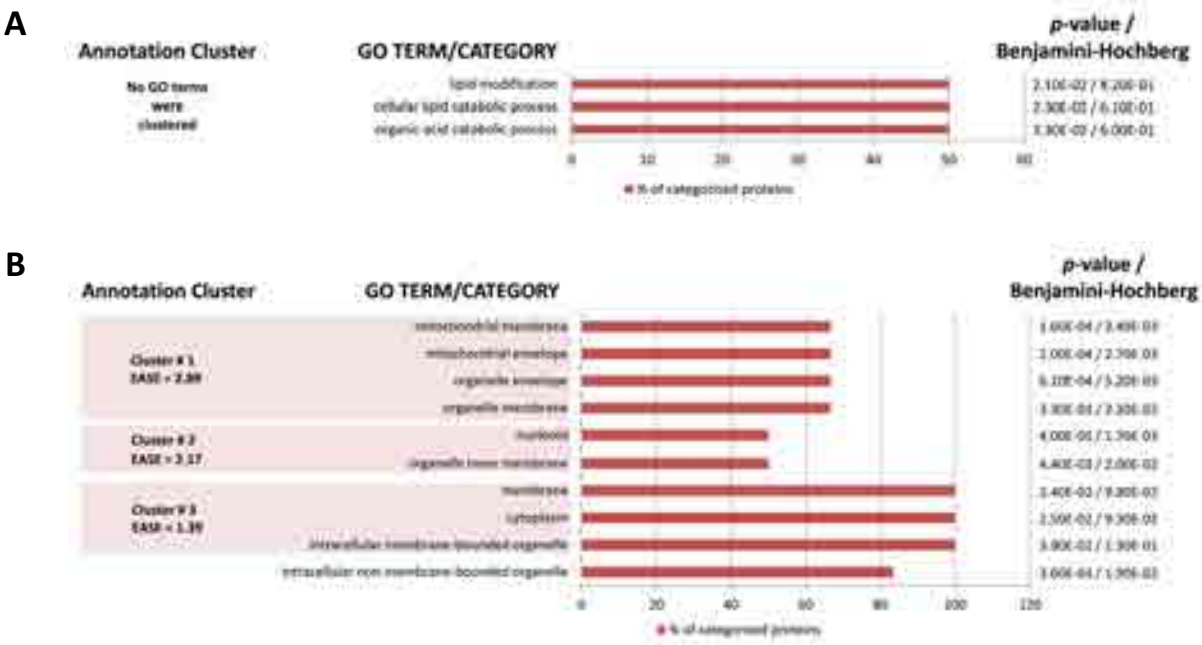


Figure 4.13 Functional categorisation of proteins up-regulated during *in vitro* adipogenic differentiation compared to precursor ADSCs. Proteins were categorised according to biological process (**A**) cellular component (**B**) and similar GO terms are clustered together and listed in order of their enrichment (EASE score) with respect to the total human proteome. GO terms that were not included in any annotation clusters are listed in order of enrichment (increasing *P*-value).

On the other hand, for the nine proteins up-regulated *in vivo* relative to the ADSCs, submission to DAVID for categorisation according to biological process resulted in only two proteins being categorised, but without significant enrichment of either category ($P > 0.05$). Grouping according to cellular component only indicated up-regulation of cytoplasmic proteins and of an 'intracellular organelle', while KEGG pathway analysis only grouped two proteins in a variety of amino acid metabolic pathways. The limited GO and KEGG categories associated with the up-regulated proteins in adipocytes compared to ADSCs may be because new proteins are produced in *in vivo* created adipocytes instead of a change being seen in differentially expressed proteins.

Whilst the two analyses above suffered from small protein lists that are limited for GO annotation, investigating the proteins individually provided some additional biological information in the context of adipogenesis. The five and nine proteins up-regulated in *in vitro* and *in vivo* adipogenesis respectively are listed in Table 4.3. Perilipin-4, a structural protein that coats adipocyte lipid droplets (Jefcoate, Wang *et al.*, 2008), was up-regulated to a similar degree in both models with 6.7-fold up-regulation in dADSCs, and 7.0-fold in adipocytes, confirming the abundance pattern observed in the heatmap in Figure 4.8 and linking this protein directly to adipogenesis. As seen in the heatmap, mitochondrial protein trifunctional enzyme subunit alpha was also up-regulated in both systems. Again, mitochondrial proteins were up-regulated in both models and each of these proteins have been previously reported to have roles in mitochondrial and lipid related processes (Fruhbeck, 2008; Reddy and Rao, 2006). Interestingly however, despite mitochondrial proteins being commonly up-regulated in both systems, trifunctional enzyme subunit alpha was the only common mitochondrial protein shared between the two lists. In each list, three proteins were classified as originating from the plasma membrane by DAVID. These findings will be further discussed in Section 4.4.

Table 4.3 Proteins significantly up-regulated during *in vitro* and *in vivo* adipogenic differentiation based on greatest fold change. Proteins up-regulated in both models are marked in bold.

Uniprot Accession	Description	Fold Change	*P*-value
PROTEINS UP-REGULATED *IN VITRO*			
Q96Q06 *	**Perilipin-4**	6.7	1.52E-02
P49748	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	4.7	1.72E-02
P40939	**Trifunctional enzyme subunit alpha, mitochondrial**	3.6	3.91E-02
P21796 *	Voltage-dependent anion-selective channel protein 1	3.4	3.55E-02
P27105 *	Erythrocyte band 7 integral membrane protein; Stomatin	1.8	3.65E-02
PROTEINS UP-REGULATED *IN VIVO*			
Q9NZN4 *	EH domain-containing protein 2	7.7	1.04E-02
P05091	Aldehyde dehydrogenase, mitochondrial	7.4	3.27E-03
Q96Q06 *	**Perilipin-4**	7.0	2.05E-02
P49327	Fatty acid synthase	4.8	1.00E-02
P43121 *	Melanoma cell adhesion molecule; CD146	4.2	2.24E-02
Q13423	NAD(P) transhydrogenase, mitochondrial	3.6	1.33E-03
Q9UHG3	Prenylcysteine oxidase 1	3.6	1.45E-02
P40939	**Trifunctional enzyme subunit alpha, mitochondrial**	2.6	3.11E-02
O00159	Unconventional myosin-Ic	2.4	6.60E-03

* PM protein as categorised by DAVID

4.3.6 Expression of ECM related proteins

Throughout the analysis presented in this chapter, ECM-related proteins and processes were enriched in different lists of the following: proteins found in adipocytes alone; proteins shared between adipocytes and dADSCs; proteins found in ADSCs alone and; proteins down-regulated during *in vitro* differentiation of ADSCs. Specifically, the ECM-related categories contained different collagen proteins (Table 4.4), demonstrating the involvement of different collagen types in adipogenesis. In this study, collagen type IV alpha chains were detected in ADSCs only, contrary to previous reports of increased collagen type IV expression in *in vitro* adipogenically differentiated products (Aratani and Kitagawa, 1988; Sillat, Saat *et al.*, 2012). Both chains of collagen type I were found in all cell types, and were down-regulated in dADSCs, in agreement with the report that a decrease of collagen type I in the ECM promoted adipogenic differentiation of mesenchymal stem cells (Rodriguez, Montecinos *et al.*, 2000)

Table 4.4 Distribution of collagen proteins* between ADSCs, dADSCs and adipocytes

ADSCs	dADSCs	Adipocytes
COLLAGENS		
== Present in all three cell types == **P02452** Collagen type I, alpha-1 chain **P08123** Collagen type I, alpha-2 chain **Q99715** Collagen type XII, alpha-1 chain		
	== Present in both dADSCs and adipocytes == **P02458** Collagen type II, alpha-1 chain **P02461** Collagen type III, alpha-1 chain **P12109** Collagen type VI, alpha-1 chain	
== Present in ADSCs alone == **P02462** Collagen type IV, alpha-1 chain **P29400** Collagen type IV, alpha-5 chain **P08572** Collagen type IV, alpha-2 chain **P20908** Collagen type V, alpha-1 chain **P05997** Collagen type V, alpha-2 chain **P12111** Collagen type VI, alpha-3 chain		== Present in adipocytes alone == **P12110** Collagen type VI, alpha-2 chain
== Present in ADSCs and adipocytes == **P05997** Collagen type V, alpha-2 chain **P12111** Collagen type VI, alpha-3 chain		

*Where multiple names exist for a protein, names relevant to the current study were selected and included here. Protein names are prefixed by their Uniprot accession number in bold

4.3.7 Comparison of mRNA transcript and protein abundance in the three cell types

The mRNA gene transcript expression of four accepted adipogenic marker proteins PPARγ2, FABP4, Perilipin-1 and Perilipin-4 and a stem cell marker CD73 were used (Chapter 2, Section 2.3.6) as a method of assessing the process of *in vitro* differentiation. The mRNA expression for these proteins in the three cell types (ADSCs, dADSCs and mature adipocytes) was compared with the actual protein expression levels in the present study (Figure 4.14). The only protein gene product not detected was PPARγ2, probably because it is not a membrane protein and would not have been enriched in the sample preparation and is therefore not included in the present comparison. When using the protein expression of the three adipogenic protein markers and one stem cell marker to evaluate differentiation, perilipin-1 and particularly perilipin-4 gave the best indication at the protein level, that the *in vitro* differentiation of ADSCs reflected that of *in vivo* adipogenesis (Figure 4.14b). Correlation of this at the mRNA level though the limited detection of perilipin-1 and -4 mRNA expression in ADSCs (Figure 4.14a-b) confirmed that they are both useful markers of adipogenesis. Despite the gene expression of FABP4 displaying similar expression in dADSCs and adipocytes and both being significantly different from ADSCs, at the gene product protein level the FABP4 protein was not detected in the dADSCs. This protein marker gave a strong indication that the late stages of *in vitro* differentiation had not occurred in the dADSCs. This early stage of differentiation of the dADSCs is further supported by levels of expression of the stem cell marker CD73, for which the

protein was highly expressed in the ADSCs, with no significant decrease upon differentiation in the dADSCs. Yet, at the mRNA level, CD73 was significantly decreased upon both *in vivo* and *in vitro* differentiation.

These results demonstrate that the transcriptome is clearly not always indicative of the proteome. Evaluating *in vitro* differentiation overall, the results indicate that dADSCs have entered the adipogenic lineage but do not show widespread protein abundance patterns that are similar to native adipocytes.

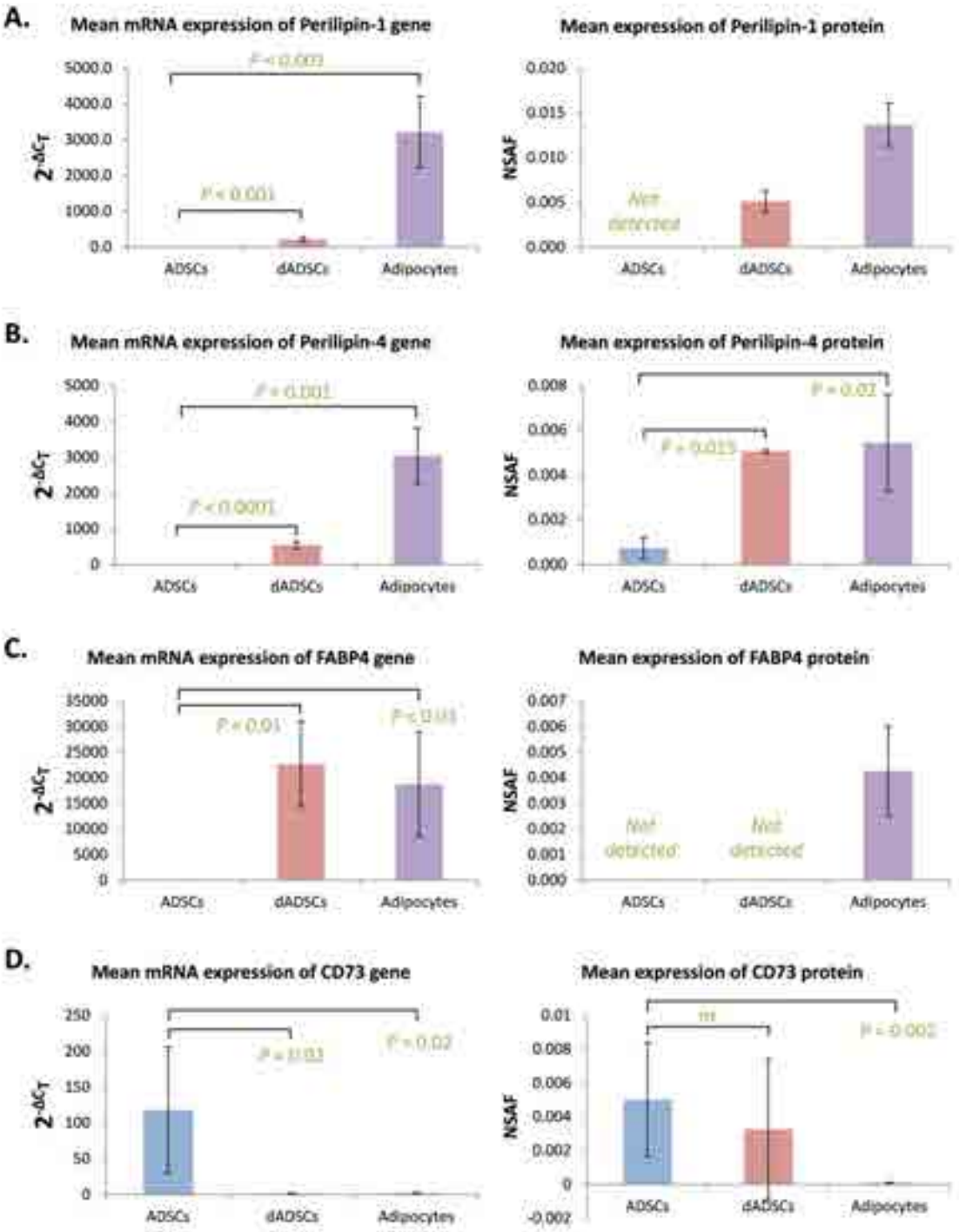


Figure 4.14 Comparison of mRNA transcript and protein gene-product abundance levels of the adipogenic markers Perilipin-1, Perilipin-4, and FABP4, and also a mesenchymal stem cell marker CD73. Proteins which were not detected reproducibly in a sample are marked as such. 'ns' = not significant.

4.4 Discussion

In this chapter, the membrane proteins of ADSCs from three liposuction patients and their *in vivo* and *in vitro* differentiated adipogenic protein products were investigated using a label-free quantitative shotgun proteomics approach. A total of 408 proteins were reproducibly identified across the three cell types, of which 67.1 % were membrane proteins which is in good accordance with the percentage membrane protein enrichment seen in previous studies (Dormeyer, van Hoof *et al.*, 2008b; Hao, Li *et al.*, 2013). Obtaining pure samples of membrane proteins is a challenge, with contamination of proteins originating from other parts of the cell a common problem (Cordwell and Thingholm, 2009; Gundry, Boheler *et al.*, 2008). In the present study, ultracentrifugation and a detergent enrichment (Lee, Chick *et al.*, 2011) was used to obtain maximum recovery and purity of hydrophobic membrane proteins. The use of SDS-PAGE separation of membrane proteins prior to trypsin digestion, avoided the membrane protein solubility problems of in-solution digest, and also helped with sample clean-up and removal of detergent prior to mass spectrometry. The percentage of specific PM proteins identified (36.4 %) was higher than the 22-24 % reported recently by Hao *et al*. (2013) who did not perform a detergent enrichment step after ultracentrifugation, and the 25 % by Dormeyer *et al*. (2008b) who used a de-glycosylation step to maximise membrane protein identification. Interestingly, proteins identified to originate from the PM showed more commonality in the three cell types in our study than the total proteins identified.

As expected, the proteins associated with fatty acid and lipid related processes were abundant in adipocytes, and some were common to both adipocytes and differentiated ADSCs. However, the enrichment of and/or up-regulation of mitochondrial proteins and processes in adipogenesis dominated the functional categorisation results. These proteins associated with mitochondrial processes were up-regulated in both adipocytes and dADSCs compared to ADSCs, and were also the proteins shared between adipocytes and dADSCs (i.e. differentially up-regulated in both models of adipogenesis). Adipose tissue is increasingly being viewed not only as an energy store, but as an organ with important roles in metabolism, satiety and immunity (Ferris and Crowther, 2011). As mitochondria play an essential role in cell energy metabolism, their function is heavily relied upon in adipocytes (Kusminski and Scherer, 2012) which explains their dominance in this study. In proteins identified in adipocytes alone, approximately 20 % were classified as mitochondrial in origin. Carbohydrate metabolism related processes such as the TCA cycle, also involved in energy production in mitochondria, were also reflected by proteins detected only in adipocytes. The TCA cycle feeds directly to the electron transport chain in mitochondria and the enrichment of related and dependent processes was encouraging. Mitochondrial remodelling upon *in vitro* adipogenesis has also been previously described in the literature and a previous study has identified 22 % of the

mature adipocyte proteome as mitochondrial (Xie, Yi *et al.*, 2010). Zhang and co-workers reported that inhibition of the mitochondrial electron transport chain suppressed adipogenic differentiation of human mesenchymal cells (Zhang, Marsboom *et al.*, 2013). The authors also reported an increase in oxygen consumption upon adipogenic differentiation, explaining the enrichment of oxoacid metabolic processes and proteins expressed in both dADSCs and adipocytes, and as the most significantly enriched category of proteins identified only in adipocytes in this study. In terms of comparing the two models of adipogenesis, despite different protein subsets identified, mitochondrial proteins and processes were characteristic of both differentiation models.

In the pair-wise comparison of ADSCs and their adipogenic progeny, the majority of proteins were in fact unchanged and therefore the resulting small numbers of differentially expressed proteins were not ideal for classifying by GO annotation according to biological process or KEGG pathway analysis. The two proteins up-regulated in both dADSCs and adipocytes were mitochondrial trifunctional enzyme subunit alpha and perilipin-4. Trifunctional enzyme subunit alpha has also been reported to be up-regulated in adipogenically differentiated mouse fibroblasts compared to undifferentiated fibroblasts (Ye, Zhang *et al.*, 2011) and is involved in catalysing long-chain fatty acid oxidation in the beta-oxidation pathway (Reddy and Rao, 2006). Up-regulation of this protein in both models of adipogenesis is perhaps representative of the overall increase in mitochondrial proteins and processes seen in adipogenesis. Perilipin-4 is a described adipogenic marker because of its involvement in lipid vesicle structure in adipocytes (Fruhbeck, 2008; Jefcoate, Wang *et al.*, 2008) therefore its up-regulation in adipocytes is expected, and the great increase of Perilipin-4 protein in dADSCs in the present study suggests its use as a marker to evaluate differentiation of stem cells into adipocytes. The abundance levels of the Perilipin-4 mRNA also increased upon differentiation of ADSCs although it did not reach the level of that in mature adipocytes.

Proteins down-regulated in differentiated ADSCs were enriched in biological processes that ultimately influence cell morphology and adhesion, reflecting the changes in these cells as they progress down the adipogenic lineage. Both fibronectin and the fibronectin receptor integrin alpha-5 were down-regulated during *in vitro* differentiation. A decrease in overall integrin abundance is required for adipogenesis as it decreases binding to the ECM and lowers focal adhesion, thereby allowing the cells to round and differentiate (Park, Han *et al.*, 2009). Fibronectin, a highly adhesive ECM component (Szabo, Feng *et al.*, 2009) that mediates cell attachment in culture, has an inhibiting effect on adipogenic differentiation of human ADSCs when used as a growth surface (Park, Han *et al.*, 2009). The decrease of integrin alpha-5 and fibronectin seen in this study are thus encouraging as potential markers for assessing *in vitro* differentiation, as is the fact that over 70% of proteins down-

regulated in dADSCs are PM proteins. Adding to this, the down-regulation of focal adhesion proteins and the focal adhesion pathway was seen in both mature adipocytes and differentiated ADSCs.

Down-regulation of different collagens was also seen in both *in vitro* and *in vivo* adipogenesis, and changes in collagen gene expression have also been described in mouse *in vitro* adipogenic differentiation (Gregoire, Smas *et al.*, 1998). Previous work using gene expression array analysis has reported the high expression of ECM related genes and the secretion of collagens by ADSCs (Alexeev, Arita *et al.*, 2014) indicating their ability to influence the ECM and their surrounding environment. In fact, ECM and focal adhesion were processes commonly enriched in this study of differentiation: as proteins found in adipocytes alone, as proteins shared in both models of adipogenesis, and also as different collagens. In differentiated ADSCs, the two alpha and beta isoforms of collagen type I was amongst the down-regulated proteins, whilst in adipocytes the alpha chain of the proteoglycan collagen type XII had the highest fold-change compared to ADSCs. A decrease in type I collagen in the ECM has been reported to favour the adipogenic differentiation of human bone marrow-derived MSCs (Rodriguez, Montecinos *et al.*, 2000). Both type I collagen chains and the alpha-1 chain of collagen type XII have also displayed high gene expression in human ADSCs (Alexeev, Arita *et al.*, 2014), however that study did not compare expression levels of the differentiated progeny. Type IV collagen alpha chains 1, 2 and 5 were detected only in ADSCs. A recent analysis of collagen type IV alpha 1 and 2 chain gene transcripts in bone marrow-derived MSCs increased in expression compared to *in vitro* adipogenic differentiation (Sillat, Saat *et al.*, 2012), and the same finding has been reported in mouse 3T3-L1 cells (Aratani and Kitagawa, 1988). The discrepancy between the up-regulation of collagen types in these reports and the down-regulation of collagen in this study may be due to the use of different stem cell progenitors (bone marrow or adipose derived) as the baseline 'control' comparison. Surprisingly, the characterisation of the proteome changes in adipogenic differentiation of human ADSCs by DeLany *et al*. (2005) did not report any collagen proteins.

Actin cytoskeletal proteins were down-regulated as a result of differentiation of ADSCs compared to ADSCs which was most likely due to the growth and proliferation of ADSCs on the surface of the flask in cell culture. Higher degrees of actin polymerisation have been reported in mesenchymal stem cells than in adipocytes, indicating that a decrease in actin density must need to occur for adipogenesis to proceed (Mathieu and Loboa, 2012). This requirement complements the need for few focal adhesions in adipogenesis (Mathieu and Loboa, 2012) as the space between cells decreases and the number of cell-cell contacts increases as the adipocytes swell with lipid.

Also down-regulated in both differentiated ADSCs and adipocytes was heat shock protein 90 (Hsp 90). The alpha isoforms were down-regulated in dADSCs, while the beta forms were down-regulated

in adipocytes. Hsp90 protein has been implicated in playing a crucial role in modulating adipogenesis through its pairing with PPARγ, and has been identified as a potential drug target in the control of obesity (Desarzens, Liao *et al.*, 2014). The inhibition of Hsp90 in murine cells impairs adipogenesis by impairing PPAR downstream events (Cuaranta-Monroy and Nagy, 2013), however the down-regulation of Hsp90 in murine cells has also been reported to be required for adipogenesis to proceed (Nguyen, Csermely *et al.*, 2013) which is in line with our results. Meanwhile Delany *et al*. (2005) reported no change in Hsp90 levels specifically after induction of adipogenesis in human ADSCs, instead highlighting the whole heat shock protein family as warranting further investigation in the study of obesity (DeLany, Floyd *et al.*, 2005). Heat shock proteins were not reported in the gene expression profiling of human bone marrow-derived MSCs undergoing adipogenesis (Hung, Chang *et al.*, 2004), nor in the proteomic analysis of human adipocytes and adipose-derived stromal vascular cells (Kheterpal, Ku *et al.*, 2011). The role of heat shock proteins in adipogenesis is therefore still inconclusive.

The functional annotation of proteins detected only in dADSCs revealed five biological processes involved in chromatin and nucleosome remodelling, which was further correlated by the nucleosome being the most significantly enriched cellular component category. These processes indicate another epigenetic level of adipogenesis regulation and modulation. In adipose-derived stromal vascular cells enriched with preadipocytes, non-methylation of the PPARγ2 promoter has been reported as an indication of adipogenic lineage commitment, while promoters for myogenic and endothelial differentiation were methylated (Noer, Sorensen *et al.*, 2006). Epigenetic related processes such as nucleotide metabolism were also enriched, but to a lesser degree in ADSCs, but not in mature adipocytes, indicating epigenetics may play a key role in control in the *in vitro* adipogenesis. This has been reported in other studies of *in vitro* adipogenic differentiation (Li, Xiao *et al.*, 2010; Musri and Parrizas, 2012).

A SNARE protein pathway was abundant amongst proteins detected only in dADSCs, and also not seen in mature adipocytes. The three proteins categorised in this pathway were syntaxin-7 (Uniprot: O15400), VAMP2 (Uniprot: P63027) and VAMP3 (Uniprot: Q15836). SNARE proteins are involved in glucose transport in adipocytes, which is mediated by the insulin-sensitive GLUT4 glucose transporter. The insulin-containing *in vitro* adipogenic differentiation cocktail therefore creates an environment that stimulates GLUT4 trafficking, which requires the presence of specific vesicle-associated membrane proteins (VAMPs) such as VAMP2 and VAMP3 (Olson, Knight *et al.*, 1997; Zhao, Yang *et al.*, 2009). Syntaxin-7 also plays a role in insulin-sensitive GLUT4 trafficking in adipocytes, however its action is not crucial to the process (Jewell, Oh *et al.*, 2010). Therefore, the presence of these proteins may reflect the action of the inducers added to induce the adipogenic differentiation.

As seen in Chapter 3 with the glycosyltransferase gene expression, gene transcript expression does not always reflect abundance at the protein level and has been previously reported (Greenbaum, Colangelo *et al.*, 2003; Gygi, Rochon *et al.*, 1999). The comparison of stem cell and adipocyte marker transcripts with their corresponding protein products showed examples of good (Perilipin-1 and -4) and poor correlation (FABP4 and CD73). The abundance of the FABP4 protein differed most from its transcript expression and was not detected at all in differentiated ADSCs. The FABP4 gene has been reported as a marker of late adipogenesis (Galateanu, Dinescu *et al.*, 2012) and while some of the transcript expression data indicated the dADSCs were in the later stages of adipogenesis due to high expression of FABP4, the FABP4 protein abundance taken alone suggests differentiation has not occurred or is still in the early stages. This example marks the importance of confirming correlation of gene and gene product abundance especially when investigating marker genes, as gene expression has increasingly been recognised as a poor indicator for protein abundance (Schwanhausser, Busse *et al.*, 2011).

The label-free quantitative shotgun proteomics approach applied in this study can produce large datasets which can be both an advantage and disadvantage since large datasets are laborious to analyse and require specialised data-mining software (Neilson, Mariani *et al.*, 2011). Stringent data filtering criteria were imposed in this study for pair-wise comparisons, with only proteins identified in all three biological replicates for a given cell type included in the comparison. Data-mining using relatively simple tools such as GO annotation analysis using DAVID, a freely available program, unearthed meaningful biological information pertinent to investigating adipogenesis and DAVID Functional Annotation Clustering tool allowed enriched categories to be quickly identified. Analysis of data at different GO levels however can offer varying degrees of information and a balance between high coverage of protein IDs (at low GO levels) or more biologically valuable information (at higher GO levels) must be carefully considered. While GO annotation invariably sacrifices some degree of protein ID coverage, as seen in this study, it can be a standardised and unbiased method for investigating functional trends in proteomics datasets (Neilson, Mariani *et al.*, 2011).

The change in cell shape from fibroblastic to spherical is described as the first hallmark of adipogenesis process and concurrent with this are changes in the components of the cytoskeleton and ECM (Gregoire, 2001). Such changes were seen in the differentiated stem cells used in this study and, along with the enrichment of mitochondrial proteins and processes, were indications of the *in vitro* adipogenic products reflecting those occurring in native adipocytes. However, at the individual protein level there were proteins for which abundance in dADSCs mirrored their progenitor ADSCs more closely instead, and others cases such as Perilipin-4 where dADSCs had similar abundance of the protein seen in native mature adipocytes. A large majority of proteins shared between ADSCs

and the two adipogenic products were unchanged with relatively few differentially expressed. Each cell type however had a subset of proteins which were identified in that cell only. Analysis of these 'unique' proteins indicated a significant enrichment of epigenetic processes in the differentiated ADSCs which were not seen to the same degree in ADSCs. Overall, while the membrane proteome of dADSCs share characteristics of both native adipocytes and ADSCS, dADSCs may also in fact have their own proteins and processes specialised for *in vitro* adipogenesis.

4.5 Concluding Remarks

The characterisation and comparison of the membrane proteomes of ADSCs, differentiated ADSCs and native mature adipocytes in this study revealed several processes that underpin adipogenesis *in vivo* and *in vitro*. Interestingly, *in vitro* differentiated ADSCs shared characteristics with both native adipocytes, such as lipid catabolism and mitochondrial activity, and with stem cells, such as abundance of nucleosome proteins indicating that differentiation of the stem cells may have been initiated but had not progressed to a mature adipocyte. This correlates with both the microscopy work conducted in Chapter 2 that showed dADSCs had entered the adipogenic lineage but had not progressed to the morphology of a mature adipocyte, and the glyco-profiling work carried out in Chapter 3 that identified significant up-regulation of bisecting GlcNAc structures in mature adipocytes but to a lesser degree in dADSCs compared to ADSCs. Overall, while the results suggest that *in vitro* differentiated ADSCs express the characteristics of native mature adipocytes only to a limited extent, the proteomics analysis, perhaps more importantly, unearthed potential epigenetic processes characteristic of *in vitro* adipogenesis. The biological pathways that underpin *in vitro* differentiation appeared not only to mimic aspects of *in vivo* differentiation, but appear to have a greater expression of the earlier epigenetic processes, and question the success of the *in vitro* differentiation environment and product.

Chapter 5:

Adipose-derived stem cell differentiation: targeting the cell membrane protein glycosylation

5.1 Introduction

In the data presented in Chapter 3, bisecting GlcNAc structures were found to be the most abundant structural type in the cell membrane protein *N*-glycome of native adipocytes compared to the adipose-derived stem cells (ADSCs), and were also significantly increased in *in vitro* differentiated stem cells (dADSCs) compared to ADSCs. While the function of bisecting GlcNAc structures in adipocytes and stem cells is currently unknown, there are growing reports detailing modification of protein function upon their addition to the *N*-linked glycan chitobiose core (Takahashi, Kuroki *et al.*, 2009). Some examples of this include the transfection of the gene that codes for the enzyme that adds bisecting GlcNAc to the core *N*-linked glycan (MGAT3) into HeLaS3 cells, where endocytosis of epidermal growth factor receptor (EGFR) protein was increased after introduction of bisecting GlcNAc structures to the protein glycosylation (Sato, Takahashi *et al.*, 2001). Addition of bisecting GlcNAc structures onto the *N*-linked glycans of integrin-α5β1 in melanoma cells inhibited binding to its cell-matrix ligand fibronectin, resulting in decreased integrin-mediated cell migration (Isaji, Gu *et al.*, 2004). These examples of the modification of plasma membrane protein glycans with bisecting GlcNAc highlight the way in which a specific sub-type of protein *N*-glycosylation can influence and modify proteins positioned at the interface between a cell and its environment. The addition of a bisecting GlcNAc to *N*-linked glycans can also influence further glycosylation as it prevents the addition of GlcNAc residues to form tri- or tetra-antennary structures (Taniguchi and Korekane, 2011).

The changes in the membrane glycome and proteome between the stem cells and their adipogenic products observed in Chapters 3 and 4 respectively, led, in this chapter, to an integrated glycomics and proteomics approach to further investigate specific differences between the cell types. This integrated approach is not often undertaken in proteomics analyses as *N*-linked glycans, if released at all, are typically discarded. Specifically, the aim was to target and determine whether the bisecting GlcNAc glycosylation of membrane proteins observed in the differentiation of adipose cells is protein-specific or characteristic of all adipocyte membrane glycoproteins. Protein-specific glycosylation implies a specific function of a particular glycan type that directly controls protein function, such as, for example, the addition of mannose-6-phosphate to lysosomal hydrolases which directs the enzyme to the lysosome (Hanisch and Breloy, 2009). As bisecting GlcNAc structures are seen to be so prevalent on the native adipocyte membrane proteins (Chapter 3), the identification of a single glycoprotein or closely related panel of glycoproteins carrying these structures may shed light on their functional role in adipocytes.

In this chapter, de-glycosylated membrane proteins from native adipocytes, ADSCs and *in vitro* differentiated ADSCs were separated using SDS-PAGE gels and the separated protein bands were analysed by glycomic profiling and shotgun proteomics techniques, providing the rare ability to determine a correlation between expressed proteins and their glycosylation profile. Technology improvements in future will enable both the identification of the proteins and the glycosylation at a particular site to be determined. Tryptic digestion of the protein bands was performed with prior *in-gel* de-glycosylation using PNGase F and compared to tryptic digests of non-PNGase F treated samples in the corresponding samples carried out in Chapter 4. Released *N*-linked glycans were analysed using porous graphitised liquid chromatography (PGC-LC) coupled with negative ion electrospray tandem mass spectrometry (ESI-MS/MS) as described in Chapter 3. Peptides were analysed using reverse phase-liquid chromatography (RP-LC) coupled to ESI-MS/MS.

5.2 Materials and Methods

5.2.1 Materials

Fetal bovine serum was purchased from Bovogen (Victoria, Australia). Lithium dodecyl sulphate (LDS) sample buffer, 3-(N-morpholino)propansulfonic acid (MOPS) buffer, and NuPAGE Novex 4-12 % (w/v) polyacrylamide Bis-Tris precast gel was purchased from Invitrogen (Life Technologies, Australia). Bradford Reagent, dithiothreitol (DTT), formic acid, ammonium bicarbonate (NH_4HCO_3), acetonitrile (ACN) and iodoacetamide (IAA) were from Sigma Aldrich (St Louis, MO, USA). µC $_8$ ZipTips were from Eppendorf (Hamburg, Germany), and 10 % (w/v) polyacrylamide TGX gel and Coomassie Brilliant Blue G-250 dye were from Biorad (NSW, Australia). HPLC-grade water was obtained using a Milli-Q water purification system (Millipore, MA, USA). Magic C18 resin was from Michrom Bioresources (Auburn, CA) and sequencing grade modified trypsin was from Promega (WI, USA). Anti-CD36 and anti-rabbit antibodies were obtained from RabMAb (Abcam, Australia).

5.2.2 Adipocyte and ADSC isolation from lipo-aspirate

The lipo-aspirate sample used in the work described in this chapter was from the first of the biological replicates used for analysis in Chapters 3 and 4. The lipo-aspirate was processed for collection of adipocytes and ADSCs, and the culture and differentiation of ADSCs in 75 cm^2 flasks as described in Chapter 2, Section 2.2.2. Cell membrane proteins were extracted from native adipocytes, ADSCs and dADSCs as outlined in Chapter 3, Section 3.2.3.

5.2.3 Trypsin digestion of SDS-PAGE separated membrane proteins with and without prior in-gel *PNGase F de-N-glycosylation*

Membrane proteins were quantified using Bradford assay (Bradford, 1976). Approximately 20 µg of each protein sample was reduced with 2.5 µl 100 mM DTT at 70 °C for 10 min in LDS sample buffer diluted with water to give a total volume of 20 µl. 1 µl 500mM IAA was added to each sample and proteins alkylated at room temperature for 30 min. Samples were then loaded onto a NuPAGE Novex 4-12 % (w/v) Bis-Tris precast gel and subjected to electrophoresis for 1 hr at 200 V. The gel was fixed with 10 % (v/v) methanol, 1 % (v/v) acetic acid before overnight staining with Coomassie G-250. Proteins were visualised using an Odyssey Infrared Imaging System (LiCor).

Each gel lane was excised from the gel and cut into approximately ten fractions. Each fraction was chopped into smaller pieces and transferred into an eppendorf tube. Gel pieces were washed twice with 50 % (v/v) ACN and then dehydrated with 100 % ACN. Rehydration with 100 mM NH_4HCO_3 for 5 min before dehydration again with 100 % ACN was repeated until all pieces were de-stained. For *N*-linked glycan release, gel pieces were rehydrated with 5 µl PNGase F (2.5 units) for 10 min at 37 ℃, before topping up each sample with approximately 20 µl 100 mM NH_4HCO_3 to ensure gel pieces were covered for overnight incubation at 37 ℃. *N*-linked glycans were extracted with 3 x 50 µl water followed by sonication each time and the supernatants pooled before drying under vacuum. Released *N*-linked glycans were reduced, desalted and further purified using carbon SPE cleanup as described in Chapter 3, Section 3.2.5, prior to PGC-LC ESI-MS/MS analysis.

Following *N*-linked glycan release, eight gel pieces from regions of the gel containing differential protein expression as determined visually from Coomassie staining (at approximately 60-80 kDa, 50-55 kDa and 55-60 kDa gel regions) were subsequently dehydrated with 100 % (v/v) ACN until they turned white. For protein digestion, 1 µl trypsin (50 mg/µl in 1 mM HCl) diluted in 2 µl 25 mM NH_4HCO_3 was applied to the dehydrated gel pieces for 10 min at 37 ℃. Approximately 20 µl 100 mM NH_4HCO_3 was then added to each sample to ensure gel pieces were covered for overnight incubation at 37 ℃. The same was carried out for non-PNGase F treated gel pieces. Peptides were extracted from gel pieces and purified as described in Chapter 4, Section 4.2.3. Purified peptide extracts were then dried using a vacuum centrifuge and resuspended in 2 % (v/v) formic acid for RP-nanoLC ESI-MS/MS analysis.

5.2.4 PGC-LC ESI-MS/MS analysis of N-*linked glycans and RP-nanoLC ESI-MS/MS analysis of de-glycosylated peptides*

Analysis of the *N*-linked glycans released from each protein band was performed as described in Chapter 3, Section 3.2.7, with MS/MS fragmentation spectra collected for the three most abundant

ions. Base peak chromatograms (BPC) of *m/z* 500-2200 were generated and the MS averaged over 30-70 min for each gel fraction.

Peptides from de-glycosylated protein bands were resuspended in 20 µl 2 % (v/v) formic acid for analysis by RP-nanoLC ESI-MS/MS using a LTQ linear ion trap mass spectrometer (Thermo, San Jose, CA). Peptides were injected onto a peptide Captrap pre-column (Michrom Bioresources; Bruker, Australia) for desalting with buffer A (0.1 % (v/v) formic acid, 2.5 % (v/v) acetonitrile) at a flow rate of 5 µl/min using a Tempo nanoLC system (Applied Biosystems; Life Technologies, Australia). Peptides were separated on a reversed-phase column made in-house by packing a fused silica capillary with an integrated electrospray tip with 100 Å, 5 µm Magic C18 resin to approximately 7 cm (with 70 µm inner diameter). At a flow rate of 650 nl/min, peptides were eluted from the column using 10-28 % buffer B (95 % (v/v) acetonitrile, 0.1 % (v/v) formic acid) for 40 min, followed by five min of 28-100 % buffer B, washing with 100 % buffer B for a further 5 min before reducing to 5% buffer B over 15 min. A 1.8 kV electrospray voltage was applied via a liquid junction upstream of the analytical column. Spectra were scanned in positive mode over the range *m/z* 400-1500 and Xcalibur software (Version 2.06, Thermo) was used for automated peak recognition, dynamic exclusion and MS/MS of the top six most intense precursor ions at 35 % normalisation collision energy. Peptides from non-PNGase F treated samples were analysed as described in Chapter 4, Section 4.2.4.

5.2.5 Peptide analysis

The LTQ-XL '.raw' output files were converted into mzXML format and searched against the curated human Uniprot/Swissprot database (20, 264 proteins as of 14 March 2014). This was done using the Global Proteome Machine (GPM) software (version 2.1.1) using the X!Tandem algorithm (Craig and Beavis, 2004) for each of the eight gel band fractions individually. Peptide identification was determined using a 2 Da parent ion and 0.4 Da fragment ion tolerance. Carbamidomethyl was considered as a complete modification, while oxidation of methionine and threonine, carbamylation of N-termini and lysine, and de-amidation of asparagine and glutamine were considered as partial modifications. Reverse database searching was also conducted for estimating false discovery rates (FDRs).

5.2.6 Protein identification

The lists of proteins identified in the membrane fraction of ADSCs, dADSCs and adipocytes were exported from GPM and filtered manually. Contaminants such as bovine serum albumin, porcine trypsin and Lys-C were removed and a cut-off applied after the first reversed hit (first false discovery identification) which resulted in protein FDRs of between 0.58 - 1.78 % across the different samples. Proteins were selected as *N*-linked glycoproteins if annotated in the UniProt-KB database

(http://www.uniprot.org/) with at least one experimentally determined *N*-linked glycosylation site and retained if identified with a minimum of two or more unique peptides. Identified glycoproteins were submitted to the Database for Annotation, Visualisation and Integrated Discover (DAVID) for functional categorisation (Huang, Sherman *et al.*, 2009a; Huang, Sherman *et al.*, 2009b) according to Cellular Component at a cut-off of GO level 3 and Biological Process at a cut-off of GO level 4.

5.2.7 Anti-CD36 Western blotting

Human torso lipo-aspirates used to derive ADSCs for membrane protein extraction, gel separation and Western blotting analysis were from the same three biological replicates used for the qRT-PCR work carried out in Chapter 2, although were different samples to that of the three individuals used for glycomic and proteomic analysis (Chapters 3 and 4 respectively). Proteins were quantified using Bradford assay and approximately 10 µg membrane proteins was reduced using 100 mM DTT in LDS sample buffer at 70 °C for 10 min before loading onto a 10 % (v/v) polyacrylamide TGX gel. Proteins were separated over 60 min at 200 V before semi-dry transfer to a nitrocellulose membrane using the Turbo blot system (Biorad). The membrane was blocked in 3 % (w/v) skim milk in 0.1% (v/v) Tween-20 in phosphate buffered saline (PBST) for one hour at room temperature before washing twice with PBST. The primary antibody (anti-CD36, rabbit anti-human) was applied at 4 °C overnight at a concentration of 1:1000 in 1 % (w/v) skim milk in PBST. The membrane was subsequently washed three times with PBST before incubation with the secondary antibody (anti-rabbit) at a concentration of 1:10 000 in 3% (w/v) milk in PBST for one hour at room temperature. Finally, the membrane was washed three times with PBST before imaging using the Odyssey Infrared Imaging System (LiCor).

5.2.8 CD36 mRNA transcript expression by qRT-PCR

For measurement of CD36 mRNA transcript expression, qRT-PCR was performed as described in Chapter 2, Sections 2.2.9-2.2.11 and was carried out in biological triplicate using the same individuals as used for Western blotting of the protein products. The primer pair used for this protein is shown in Table 5.1 and TATA-box binding protein was used to normalise CD36 mRNA expression as in Chapter 2.

Table 5.1 qRT-PCR CD36 primer pair sequence

Gene Name	Symbol	Forward (5'-3')	Reverse (5'-3')	Primer Target/s	Product Size (bp)
CD36	CD36	GCAACAAACCACACACTGGG	AGTCCTACACTGCAGTCCTCA	Transcript variants 1 & 3	130
TATA-box binding protein	TBP	GCACCACTCCACTGTATCCC	GCTGCGGTACAATCCCAGAA	Transcript variant 1	116

5.3 Results

Using a targeted approach, the cell membrane glycosylation of native adipocytes, ADSCs and differentiated ADSCs was further explored. This work complements the results of the glycomics analysis of the membrane protein glycosylation of these cell types presented in Chapter 3 in which bisecting GlcNAc structures were found to be a major *N*-linked glycan type, particularly in native adipocytes, that were significantly increased in the membrane *N*-glycome of differentiated adipose-derived stem cells.

5.3.1 SDS-PAGE separation of adipocyte, ADSC and dADSC membrane proteins

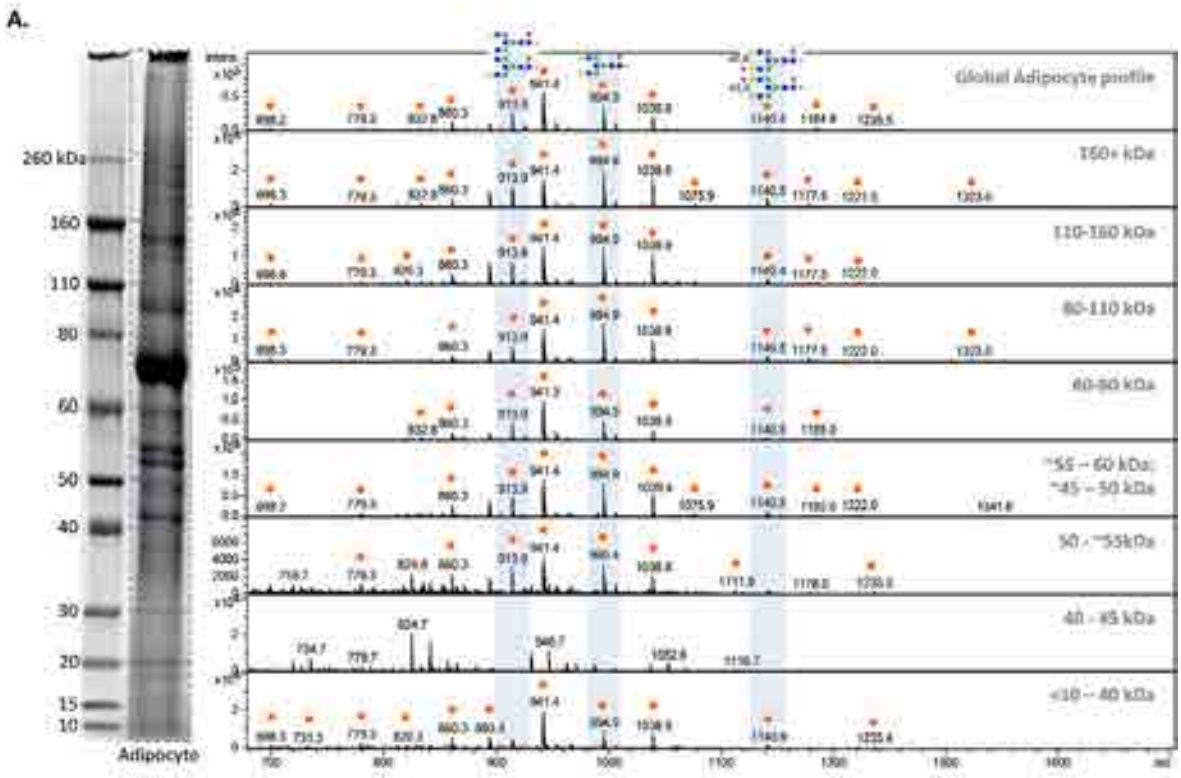
To determine whether the global *N*-glycomic profiles observed in Chapter 3 (Figure 3.11) were representative of the glycosylation of the majority of membrane glycoproteins of each cell type or were protein specific, membrane proteins from adipocytes, ADSCs and dADSCs were separated using SDS-PAGE. The gel lanes were cut into fractions and glycomic analysis carried out on the individual fractions. For some regions of the gel, proteomic analysis was also carried out so that glycan structures could be attributed to the identified glycoproteins within a defined molecular weight range.

5.3.1.1 Analysis of N*-linked glycans released from gel fractions following SDS-PAGE of membrane proteins from native adipocytes, ADSCs and dADSCs*

The *N*-linked glycan profiles from each gel fraction of the separated membrane proteins of native adipocytes, ADSCs and dADSCs were found to be reflective of the corresponding global *N*-linked glycan profile. Ion signals corresponding to glycan structures (Figure 5.1a-c) are marked with an orange circle and correspond to the *m/z* values listed in Table S3.1 (Supplementary data).

The native adipocytes contained the three main ion signals corresponding to bisecting GlcNAc structures that were prominent in the global profile (*m/z* $[913.9]^{2-}$; $[994.9]^{2-}$; $[1140.9]^{2-}$). These signals were similarly prominent across all of the molecular weight sub-divisions of the separated membrane

proteins and not seen in the average MS of ADSCs and dADSCs (highlighted panels; Figure 5.1). This indicates that bisecting GlcNAc structures are present on the majority of glycoproteins expressed on adipocytes membranes and that this glycosylation type is not particularly protein-specific in adipocytes although its presence may be differently distributed across the proteins in the molecular weight bands. Separated ADSC membrane protein glycomic profiles (Figure 5.1b) contained ion signals corresponding to glycan structures also reflective of the global *N*-linked glycomic profile, while more variation was seen in the glycosylation of the dADSCs (Figure 5.1c) between the gel fractions and global profile. In both the dADSCs and ADSCs however there were more protein fractions where glycans were not detected compared to the same molecular weight sub-fractions of the native adipocyte proteins, particularly in the lower molecular weight ranges (<50 kDa) suggesting less overall protein glycosylation. In dADSCs, the ion signal with *m/z* $[974.4]^{2-}$, as described in the global *N*-glycome profile (Chapter 3, Section 3.3.4.3), corresponding to Gal-Gal epitope-containing structures, was detected only in one mass range of 55-60 kDa, indicating that its expression may be specific to one or a small panel of proteins.



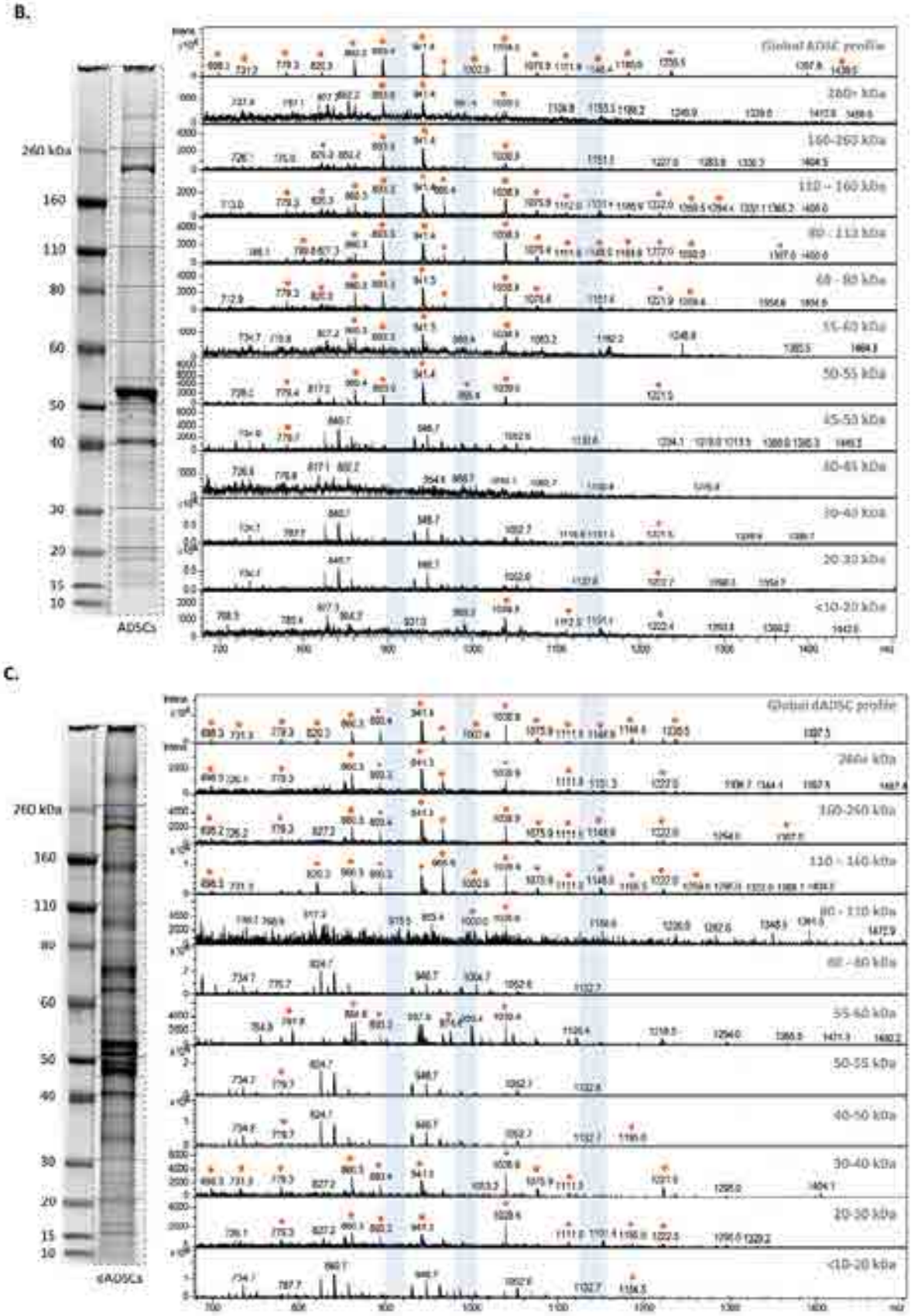


Figure 5.1 Average MS profiles of *N*-linked glycans released from membrane proteins separated using SDS-PAGE compared with the global *N*-glycomic profiles in native adipocytes (**A**), ADSCs (**B**) and differentiated ADSCs (**C**). Signals corresponding to glycan structures are marked with an orange circle and structures are given in Supplementary Table S3.1. The corresponding protein molecular weight range is labelled for each MS profile.

5.3.1.2 Identification of abundant glycoproteins in native adipocytes, ADSCs and dADSCs

To more exactly identify the differential protein expression between the three cell types, glycoproteins were specifically targeted for identification in areas of the gel displaying obvious different abundances of protein. Visual differences in abundance of protein expression were seen between the molecular masses of 60-80 kDa and 50-60 kDa (Figures 5.1-5.3). Although equal loading of protein amount was indicated by the Bradford assay, there appeared to be less total protein in the ADSC sample, but despite this, the major protein bands are clearly different.

For each identification, the log(e) score denotes assignment confidence, with a lower score indicating a higher confidence or lower chance of a random identification. Functional categorisation of glycoproteins was carried out using DAVID for Cellular Component annotation at a cut-off of GO level 3. Some proteins were not classified and it is important to note that some proteins (e.g. CD36, vimentin, prenylcysteine oxidase and adipocyte plasma-membrane associated protein) can be found at multiple locations suggesting degradation or isoforms of the protein. The full list of proteins identified in all gel bands can be found in Table S5.1 (see accompanying DVD).

The top ten glycoproteins identified for each gel fraction of interest were identified and are shown in Tables 5.2-5.9. The trypsin digest for protein identification was carried out on the same sample before and after de-glycosylation with PNGase F to increase the coverage and consequent identification of the glycoproteins.

The differences in the protein profiles of the membrane proteins of the three cell types were particularly striking on the gel at approximately 70 kDa (Figure 5.2). At this mass region, there was a major protein band in the gel lane separating the adipocyte proteins, with two smaller bands in the dADSC protein lane at this molecular range, and little of this molecular weight at all amongst the ADSC proteins. The 'top' glycoprotein (identified with the highest confidence or lowest log(e) score) in this region in the adipocyte sample (Band 1; Figure 5.2), and in fact the most abundant protein identified, was CD36 with 56 unique peptides and a total of 580 peptides. CD36 is a marker of the adipogenic lineage (Bunnell, Estes *et al.*, 2008) and was also present in the corresponding region in dADSCs (Band 2; Figure 5.2) however at much lower abundance with 15 unique and 69 total peptides. Although CD36 has a predicted amino acid sequence molecular weight of approximately 53 kDa, its ten *N*-glycosylation sites can result in a predicted weight of up to 90 kDa (Febbraio and Silverstein, 2007), indicating that the CD36 detected here is highly glycosylated as also indicated by the typical thickness of a heterogeneous glycoprotein band. Further variants with molecular weights lower than 53 kDa can also be produced from alternate splicing events, and were detected in adipocytes at both the 50-60 kDa and ~40 kDa molecular mass regions (Table 5.4 and 5.7). In

comparison, although CD36 was present, the top glycoprotein identification in the corresponding gel region in dADSCs (Band 2; Figure 5.2) was CD73, which is a mesenchymal stem cell marker. The identification of a stem cell marker here indicates that the dADSCs share characteristics of both mature adipocytes and their progenitor stem cells, as also concluded in previous chapters of this thesis.

As seen in Tables 5.2 - 5.9, it is clear that significantly improved protein identifications are made if proteins are de-glycosylated before trypsin digestion. For example, Galectin-3-binding protein (in Band 2) was detected with only 1 peptide when the sample was not de-glycosylated whereas 6 unique peptides were seen after de-glycosylation. Overall, the major difference observed in this molecular mass region was the differential abundance of CD36 with 9 times the total peptides and 3.7 times the unique peptides found in the native adipocyte membranes compared with differentiated ADSCs. This difference between the two adipocyte cell types was not apparent using the shotgun proteomics approach (Chapter 4).

Differences in the membrane protein profiles between the three cell types were also observed in the 50-60 kDa (Figure 5.3) mass range. Glycoproteins identified in both dADSCs and adipocyte cell membranes were prenylcysteine oxidase 1, serpin H1 and calreticulin, whilst adipocyte plasma membrane associated protein was found in all three cell types. Within each band, there were glycoproteins that were not detected either in the corresponding gel mass range (for example, haemoglobin subunit beta; Table 5.4, and sushi domain-containing protein 2; Table 5.5) or at all in the samples that were not de-glycosylated (for example, calreticulin; Table 5.5, and protein-lysine 6-oxidase; Table 5.6).

The 40 kDa mass region showed a major band in all three cell types, with relatively high abundance in the ADSCs (Figure 5.1 & 5.4). CD36 was again identified in the adipocyte Band 6 sample (Figure 5.4) but is most likely a cleavage product as the non-glycosylated mass of CD36 is 53 kDa. Again, several glycoproteins in Band 7 specifically were detected in the de-glycosylated samples and not detected in the glycosylated samples, namely HLA class I histocompatibility antigen B-53 alpha chain, synaptophysin-like protein 1, inositol monophosphatase 3, cation-dependent mannose-6-phosphate receptor, pentraxin-related protein PTX3 and inhibitor of nuclear factor kappa-B kinase-interacting protein.

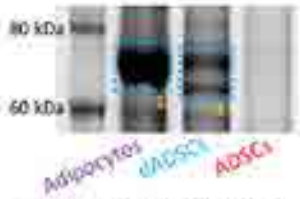


Figure 5.2 Individual gel bands chosen for targeted proteomics analysis and glycoprotein identification in the 60-80 kDa mass range, as listed in Tables 5.2 (adipocytes) and 5.3 (dADSCs). Insufficient protein was available at this mass range in the ADSCs

Table 5.2 Top ten g ycoprote n dent f cat ons n ad pocytes (Band 1, F gure 5.2), sted n order of number of un que pept des, before and after de-g ycosy at on. A together, a tota of 126 prote ns were dent f ed n th s band w th a 0.79% FDR

Uniprot accession	Name	GO Cellular Component	Theoretical Molecular Weight (kDa)	Results from de-glycosylated samples			Results from glycosylated samples		
				Log(e)	No. of unique peptides	No. total peptides	Log(e)	No. of unique peptides	No. total peptides
P16671	CD36	M	53 0	-689 8	56	580	-206 8	6	68
Q14108	Lysosome memb a e p o e 2/CD36 a ge - ke 2	M	54 3	- 05 2		26	- 6	2	2
P23229	eg a p a-6/CD49f	M	26 6	-86	9	4	-26 6	3	3
Q16853	Memb a e p ma y am e ox dase	M	84 6	-60 9	7	2	-229 3	20	80
P14314	G ucos dase 2 subu be a	uc	59 4	-54 6	6		-50	6	7
Q07954	o ow-de s y pop o e ecep o - e a ed p o e	M	504 2	-5 2	6	2	- 4		
Q16363	Lam subu a p a-4	ECM	202 4	-4 6	5	6	-53 5	5	7
Q01082	Spec be a c a , b a	M	274 4	-40	5	6	-92 7	7	8
P05556	eg be a- /CD29	M	88 4	-52 8	4	9	- 4 8	4	6
P43121	MUC 8/CD 46	M	7 6	-36 7	4	7	- 78	4	22

M = p asma memb a e; uc = u -ca ego sed; ECM = ex ace u a ma x

Table 5.3 Top ten g ycoprote n dent f cat ons n dADSCs (Band 2, F gure 5.2), sted n order of number of un que pept des, before and after de-g ycosy at on. A together, a tota of 142 prote ns were dent f ed n th s band w th a 0.70% FDR

Uniprot accession	Name	GO Cellular Component	Theoretical Molecular Weight (kDa)	Results from de-glycosylated samples			Results from glycosylated samples		
				Log(e)	No. of unique peptides	No. total peptides	Log(e)	No. of unique peptides	No. total peptides
P21589	5'- uc eo dase/CD73	M	63 3	-244 9	22	65	-9 0	9	
P04843	Do c y -d p osp oo gosacc a de-p o e g ycosy a sfe ase subu	ER	68 5	- 65 8	7	39	- 4	2	2
P16671	CD36	M	53	- 66 2	5	69	-37 8	3	3
P08195	4 2 ce -su face a ge eavy c a /CD98	M	68	- 9 5	3	23	-20 0		
Q14108	Lysosome memb a e p o e 2/CD36 a ge - ke 2	M	54 3	- 33 4	2	36	-62 7	7	0
P23219	os ag a d G/H sy ase	M	68 6	- 09 3	9	52	-2 7	2	2
Q08380	Ga ec -3-b d g p o e	ECM	65 3	-47 8	6	9	-2 4		
P16070	CD44 a ge	M	8 5	-52 9	5		-5	6	6
P18084	eg be a-5	M	88	-48 2	5	0	-6		
P16070	CD44	M	8 5	-52 9	5		-5	6	6

M = p asma memb a e; ER = e dop asm c e cu um; ECM = ex ace u a ma x

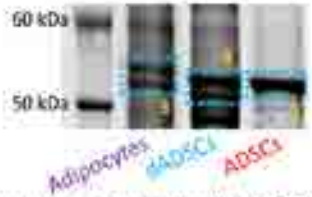


Figure 5.3 Individual gel bands chosen for targeted proteomics analysis and glycoprotein identification in the 50-60 kDa mass range, as listed in Tables 5.4 (adipocytes), Table 5.5 (dADSCs) and Table 5.6(ADSCs)

Table 5.4 Top ten g ycoprote n dent f cat ons n ad pocytes (Band 3, F gure 5.3), sted n order of number of un que pept des, before and after de-g ycosy at on. A tota of 173 prote ns were dent f ed n th s band w th a 0.58% FDR.

Uniprot accession	Name	GO Cellular Component	Theoretical Molecular Weight (kDa)	Results from de-glycosylated samples			Results from glycosylated samples		
				Log(e)	No. of unique peptides	No. total peptides	Log(e)	No. of unique peptides	No. total peptides
P16671	CD36	M	53	- 76	7	59	-75 5	6	3
Q9UHG3	e y cys e e ox dase	M	56 6	- 93	6	8	- 59	4	9
P51888	o a g	ECM	43 8	- 42	3	74	- 3 8	2	2
P50454	Se p H	ER	46 4	- 22 3		22	- 6	2	2
P12111	Co age ype V a p a-3 c a	ECM	343 5	-97		6	-40 3	5	6
P27797	Ca e cu	ECM	48	- 4 3	0	28	-35 9	4	5
Q16853	Memb a e p ma y ain e ox dase	M	84 6	-78 7	8	2	-85 6	8	9
Q9HDC9	Ad pocy e p asma memb a e-assoc a ed p o e	M	46 5	-7 7	8	6	- 24 2	2	4
P68871	Haemog ob subu be a	cy oso	6 0	-60 8	7	3	De ec ed ge mo ecu a we g eg o ≤ 20 kDa - 8 6	2	2
P02452	Co age ype a p a- c a	ECM	38 8	-52 2	6	9	De ec ed ge mo ecu a we g eg o 0-260 kDa -2 7	3	3

M = p asma memb a e; ER = e dop asm c e cu um; ECM = ex ace u a ma x

Table 5.5 G ycoprote n dent f cat ons n dADSCs (Band 4, F gure 5.3), sted n order of number of un que pept des, before and after de-g ycosy at on. A tota of 114 prote ns were dent f ed n th s band w th a 0.88% FDR

Uniprot accession	Name	GO Cellular Component	Theoretical Molecular Weight (kDa)	Results from de-glycosylated samples			Results from glycosylated samples		
				Log(e)	No. of unique peptides	No. total peptides	Log(e)	No. of unique peptides	No. total peptides
Q9UHG3	e y cys e e ox dase	M	56 6	- 72	5	58	-2 7		
P50454	Se p H	ER	46 4	- 44 8	3	26	-2 8	3	4
P36269	Gamma-g u amy a sfe ase 5	M	62 2	-94	9	29	-8 4	2	2
P35613	Bas g /CD 47	M	42 2	-70 8	7	5	-2 2		
P27797	Ca e cu	ER	48	-48 9	6	5	o de ec ed		
Q08431	Lac ad e	M	43	-43 8	5	8	o de ec ed		
Q9HDC9	Ad pocy e p asma memb a e-assoc a ed p o e	M	46 5	-28 3	4	8	De ec ed ge mo ecu a we g eg o 60- 0kDa - 2 9	2	2
Q9UGT4	Sus doma -co a g p o e 2	M	90	-36 4	3	4	De ec ed ge mo ecu a we g eg o 60- 0kDa - 2 0		2
P15144	Am opep dase N/CD 3	M	02 4	- 9 5	3	3	-27 7	3	3
Q12797	Aspa y/aspa ag y be a- yd oxy ase	ER	85 8	-20 8	2	4	De ec ed ge mo ecu a we g eg o ≤ 20 -2 8		

M = p asma memb a e; ER = e dop asm c e cu um

Table 5.6 Glycoprotein identifications in ADSCs (Band 5, Figure 5.3), listed in order of number of unique peptides, before and after de-glycosylation. A total of 126 proteins were identified in this band with a 1.78% FDR

Uniprot accession	Name	GO Cellular Component	Theoretical Molecular Weight (kDa)	Results from de-glycosylated samples			Results from glycosylated samples		
				Log(e)	No. of unique peptides	No. total peptides	Log(e)	No. of unique peptides	No. total peptides
P28300	o e - ys e 6-ox dase	M	46 9	- 2 7	2	4		o de ec ed	
Q9HDC9	Ad pocy e p asma memb a e-assoc a ed p o e	M	46 5	- 3	2	3	- 0	2	2
P35613	Basig /CD 47	M	42 2	-9 4	2	2	-66 4	3	7

M = p asma memb a e; ER = e dop asm c e cu um

Figure 5.4 Individual gel bands chosen for targeted proteomics analysis and glycoprotein identification in the 40 kDa mass range, as listed in Tables 5.7 (adipocytes), Table 5.8 (dADSCs) and Table 5.9 (ADSCs)

Table 5.7 Top ten glycoprotein identifications in adipocytes (Band 6, Figure 5.4), listed in order of number of unique peptides, before and after de-glycosylation. A total of 135 proteins were identified in this band with a 0.74% FDR

Uniprot accession	Name	GO Cellular Component	Theoretical Molecular Weight (kDa)	Results from de-glycosylated samples			Results from glycosylated samples		
				Log(e)	No. of unique peptides	No. total peptides	Log(e)	No. of unique peptides	No. total peptides
Q12797	Aspa y/aspa ag y be a- yd oxy ase	ER	85 8	- 24	3	3	-55 5	5	8
P14543	N doge -	ECM	36 3	- 23 6	3	37	-60 4	6	6
P16671	CD36	M	53	- 24		70	-25	3	7
O94905	E -2	ER	37 8	-84 2	9	2	-68 7	6	7
P07585	Deco ; Bo e p o eog yca	ECM	39 7	-78 4	8	42		o de ec ed	
P04439	HLA c ass s ocompa b y a ge , A-3 a p a c a	GA	40 8	-76 5	8	6	-3 7	4	5
Q9HDC9	Ad pocy e p asma memb a e-assoc a ed p o e	M	46 5	-70 5	8	20	- 0	2	2
P51888	o a g	ECM	43 8	-70 2	8	4	- 3 8	2	2
Q9UHG3	e y cys e e ox dase	M	56 6	-77 7	7	24	- 6		
P35613	Bas g /CD 47	M	42 2	-54 7	6	9	-2 2		

M = p asma memb a e; ER = e dop asm c e cu um; GA = Go g appa a us; ECM = ex ace u a ma x

Table 5.8 Top ten glycoprotein identifications in dADSCs (Band 7, Figure 5.4), listed in order of number of unique peptides, before and after de-glycosylation. A total of 107 proteins were identified in this band with a 0.93% FDR

Uniprot accession	Name	GO Cellular Component	Theoretical molecular weight (kDa)	Results from de-glycosylated samples			Results from glycosylated samples		
				Log(e)	No. of unique peptides	No. total peptides	Log(e)	No. of unique peptides	No. total peptides
Q9HDC9	Ad pocy e p asma memb a e-assoc a ed p o e	M	46 5	-200	8	70	- 2 9	2	2
P50454	Se p H	ER	46 4	- 34 3		2	-2 8	3	4
A4D0R2	Sy ap op ys - ke p o e	M	26 4	-54 9	5	7		o de ec ed	
Q9NX62	os a mo op osp a ase 3	GA	38 7	-53 6	5	5		o de ec ed	
P20645	Ma ose-6-p osp a e ecep o	M	3	-53	5	6		o de ec ed	
P26022	e ax - e a ed p o e TX3	ECM	4 9	-35 9	5	7		o de ec ed	
P30491	HLA c ass s ocompa b y a ge , B-53 a p a c a	M	40 5	-9 5	4	5		o de ec ed	
P35613	Bas g /CD 47	M	29 2	-39 2	4	8	-2 2		
Q70UQ0	b o of uc ea fac o kappa-B k ase- e ac g p o e	ER	39 3	-24 8	3	4		o de ec ed	
Q15293	Re cu oca b -	ER	38 9	-20 7	3	3		o de ec ed	

M = p asma memb a e; ER = e dop asm c e cu um; GA = Go g appa a us; ECM = ex a-ce u a ma x

Table 5.9 Six glycoprotein identifications were made in ADSCs (Band 8, Figure 5.4), listed in order of number of unique peptides, before and after de-glycosylation. A total of 55 proteins were found in this band with a 1.82% FDR

Uniprot accession	Name	GO Cellular Component	Theoretical molecular weight (kDa)	Results from de-glycosylated samples			Results from glycosylated samples		
				Log(e)	No. of unique peptides	No. total peptides	Log(e)	No. of unique peptides	No. total peptides
P20645	Ma ose-6-p osp a e ecep o	M	3	-3 9	4	9	-5 3		2
P04439	HLA c ass s ocompa b y a ge , A-3 a p a c a	M	40 8	-26 5	4	5	-3 7	4	5
P08962	CD63	M	25 6	-30 9	3	2	- 3 8	2	4
Q94905	E -2	ER	37 8	-23 4	3	5		o de ec ed	
P16671	CD36	M	53	-2	2	3	-25	3	7
Q10588	AD - bosy cyc ase/cyc c AD - bose yd o ase 2	M	35 7	- 5 2	2	6	- 5 3	2	2

M = p asma memb a e; ER = e dop asm c e cu um

5.3.2 Effect of de-glycosylation on determination of CD36 abundance in adipocytes

Of all the membrane glycoproteins identified (Section 5.3.1), CD36 displayed the most striking differential expression across the three cell types, both visually on the gel and in the number of unique and total peptides identified between adipocytes and dADSCs. The observation of the abundance of the major CD36 glycoprotein in adipocytes was significantly improved by de-glycosylation and gel separation before trypsin digest. In the adipocytes, the total peptides matching CD36 were increased after de-glycosylation by 8.5 fold with 3.5 times the number of unique peptides presented in Chapter 4. Similarly in the dADSCs, although at a lower abundance, CD36 was seen at a 6 and 2.5 fold increase in coverage respectively after de-glycosylation. CD36 did not emerge as a statistically significant protein displaying differential expression in the statistical analyses conducted on glycosylated samples analysed by shotgun proteomics on the whole membrane protein fraction (Chapter 4).

For closer inspection of CD36 sequence coverage by MS, the amino acid sequence coverage (matching of fragment spectra to *in silico* fragments in GPM) for glycosylated CD36 (data from Chapter 4) and de-glycosylated CD36 (results from Band 1, Figure 5.2) were compared (Figure 5.5). Sixty eight total peptides were identified in the glycosylated counterpart whilst 580 were identified when the glycoprotein was de-glycosylated prior to trypsin digestion and analysis, resulting in a 41.74 % increase in sequence coverage. Matching peptide sequences are shown in red (Figure 5.5) and were higher in number in the de-glycosylated sample, covering much more of the sequence. The majority of the protein sequence in the glycosylated sample is highlighted in green and contains eight several *N*-linked glycosylation sites (N-X-S/T; where X can equal any amino acid except proline). This strongly indicates that the presence of *N*-glycosylation on the protein hinders the ability of trypsin to cleave the protein and/or prevents the matching of detected glycopeptide MS/MS spectra with theoretical *in silico* peptide sequence masses, thereby lowering the number of identified peptides matching the protein in question.

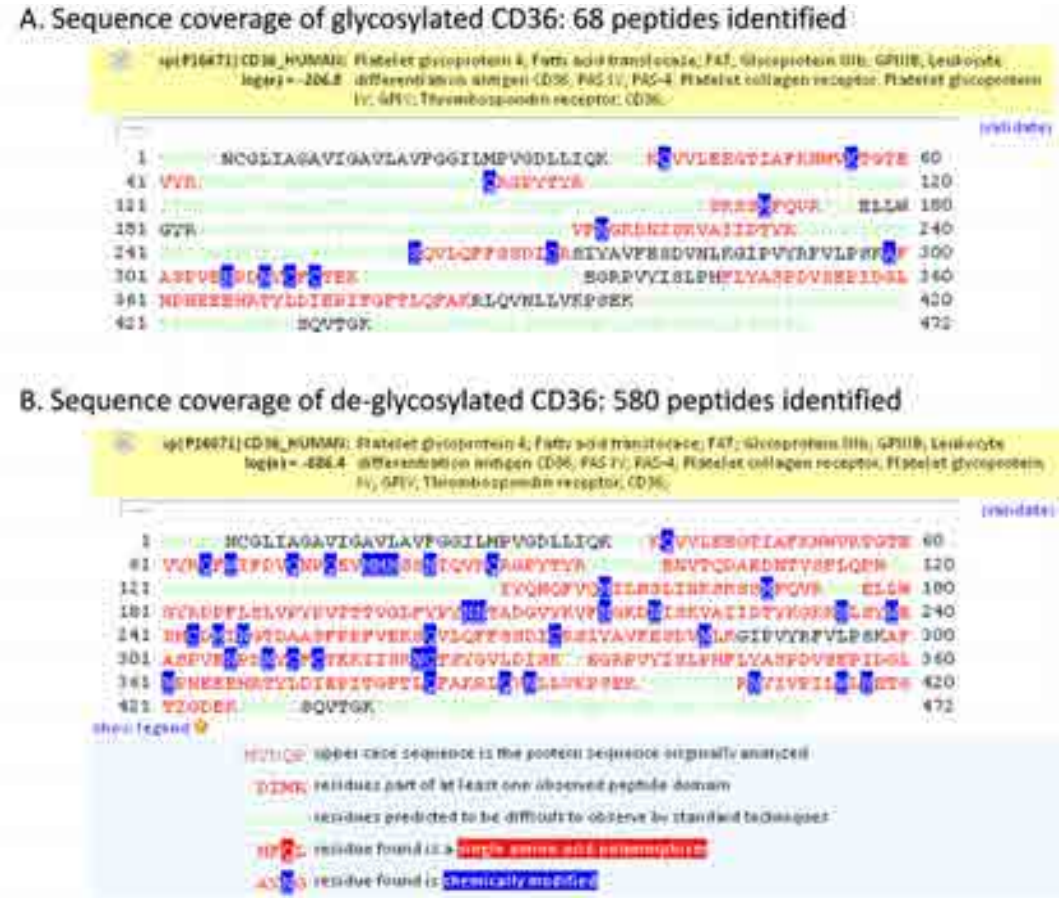


Figure 5.5 CD36 protein sequence coverage in GPM using the X!Tandem algorithm for peptide MS/MS spectra matching with *in silico* fragments with *N*-linked glycans present (27.54 % sequence coverage)(**A**) and removed (de-glycosylated, 69.28 % sequence coverage) (**B**).

5.3.3 Confirming CD36 differential expression using complementary Western blotting and qRT-PCR

To confirm CD36 as the most abundant protein in native adipocytes, CD36 expression was further measured using two targeted approaches of Western blotting and qRT-PCR.

Western blotting with anti-CD36 antibody of the gel separated, de-glycosylated proteins and qRT-PCR measurement of the CD36 gene transcripts of the three cell types was conducted in biological triplicate. The results of the Western blot showed the binding of the anti-CD36 antibody only in the 70 80 kDa molecular weight range, with binding strongest in adipocytes followed by dADSCs, and almost absent in ADSCs. This supports the mass spectrometric identification. No binding was seen in the 40 kDa molecular range, despite identification of CD36 by MS/MS at this mass in adipocytes and dADSCs (Table 5.7 and 5.9) indicating the antibody may not recognise this isoform. The region of CD36 recognised by the antibody is between amino acids 350-450 which contains one *N*-glycosylation site motif at Asn$^{4\ 7}$. The detection of CD36 by MS/MS at the 40 kDa range is most likely a cleavage product therefore, containing an area of the protein not recognised by the antibody.

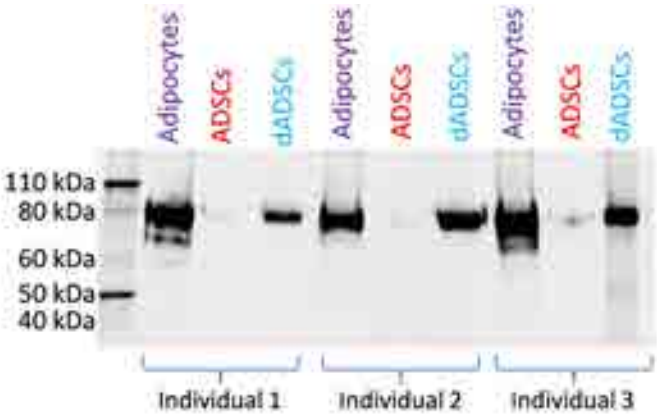


Figure 5.6 Western blot probing for CD36 in adipocyte, ADSC and dADSC membrane proteins in three biological replicates.

The expression pattern of CD36 observed by Western blotting was also seen in analysis at the mRNA transcript level. qRT-PCR measurement of CD36 transcripts showed significantly higher expression in both mature adipocytes (102.9-fold increase) and dADSCs (7.1-fold increase) compared to ADSCs (Figure 5.7a), with expression highest in adipocytes as was observed by Western blot at the protein level. This is in contrast to the shotgun proteomic analysis of these samples, which did not reveal any significant differences in the abundance of the CD36 protein between ADSCs and both adipogenic products (Figure 5.7b; data from Chapter 4), despite the fact that CD36 is known as an adipogenic marker. Overall, the results of both techniques carried out at the different gene-product levels (mRNA and antibody detected protein expression), strongly indicated that CD36 protein abundance was masked in the proteomic analysis when de-glycosylation of the proteins was not carried out prior to MS analysis. This highlights the crucial need for de-glycosylation prior to proteomics analyses regardless of whether the experiment has a glycomics based aim, to allow the proper detection and identification of glycoproteins and their abundance. The results of this section have also demonstrated the importance of verifying abundance of proteins using orthogonal techniques.

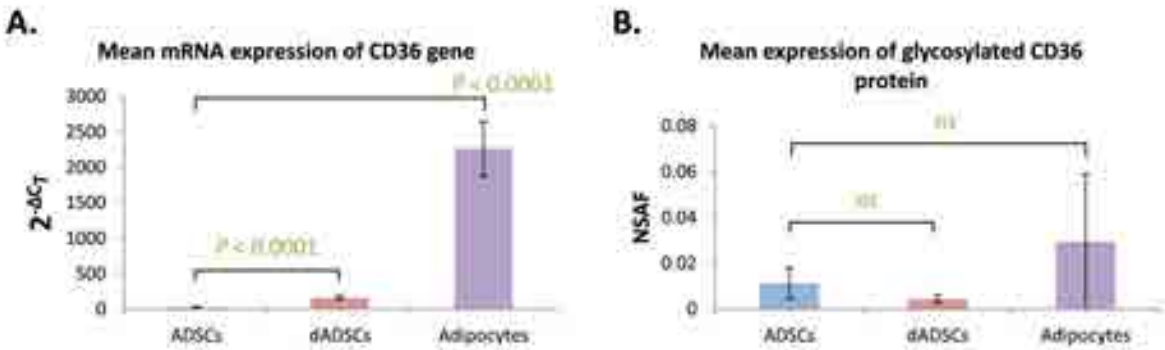


Figure 5.7 Comparison of mRNA transcript (A) and shotgun proteomic expression levels (B) of the adipogenic marker CD36. *'ns'* = not significant. For transcript analysis, expression levels were normalised to the TBP housekeeping gene and are shown as $2^{-\Delta CT}$ mean ± standard deviation. *P*-values from student's *t*-test comparison of ADSCs and dADSCs, and ADSCs and adipocytes are shown.

5.4 Discussion

In this chapter, a targeted approach was undertaken to investigate the membrane glycosylation of adipocytes, ADSCs and differentiated ADSCs from the same individual to complement the results of the glycomic and proteomics analyses presented in Chapters 3 and 4 respectively. In particular, the observation that bisecting GlcNAc structures are the dominant structural type in the adipocyte membrane *N*-glycome (Chapter 3) was further investigated.

The analysis of *N*-linked glycans released from adipocyte, ADSC and dADSC membrane proteins across the SDS-PAGE gel molecular weight range showed similarity to the global *N*-glycome profiles for each sample. For adipocytes, this indicated that bisecting GlcNAc structures were present on the majority of adipocyte glycoproteins across the molecular weight range and was not protein specific. This is not surprising as the proteins are processed through the same cellular glycosylation machinery, but certainly does not preclude the variability of site glycosylation on individual glycoproteins. It has been shown that glycoproteins have highly heterogeneous site glycosylation with the accessibility of individual sites on a protein able to determine the types of *N*-glycan structures with are attached (Thaysen-Andersen and Packer, 2014). While no information about the specific function of bisecting GlcNAc glycans in adipocytes was obtained in this thesis, their widespread expression on the majority of native adipocyte glycoproteins, their appearance on the membrane proteins of differentiated ADSCs, and corresponding significantly lower abundance in ADSCs, could indicate an important involvement in the unique physiology or function carried out by adipocytes in the body. To investigate the role of bisecting structures in adipocytes, one possibility would be to silence or knockout the MGAT3 gene in ADSCs following by testing of adipogenic differentiation capability using Oil Red O to investigate the impact of a loss or decrease of bisecting GlcNAc structures in adipogenesis.

The proteomic analysis of de-glycosylated membrane proteins separated by SDS-PAGE revealed different sets of glycoproteins compared to the global shotgun proteomic analysis of the corresponding gel-separated glycosylated membrane proteins. Glycoproteins that were newly identified in the de-glycosylated samples and not in the glycosylated counterparts were identified to be from the plasma membrane, endoplasmic reticulum, Golgi apparatus and extra-cellular matrix according to GO annotation. GO annotation according to biological processes indicated these membrane proteins are involved in various cellular functions such as protein and glycoprotein metabolism (Erlin-2, Uniprot: O94905 and Decorin, Uniprot: P07585) and endocytosis (mannose-6-phosphate receptor, Uniprot: P20645, lactadherin, Uniprot Q08431, and pentraxin-related protein PTX3, Uniprot P26022) but for the interest of this study none of the processes were adipogenic. However, this de-glycosylation and gel separation approach unearthed the adipogenic marker CD36

as the major glycoprotein in adipocytes which was found to appear in dADSCs upon adipogenic differentiation. This observation was confirmed using orthogonal techniques of Western blotting and mRNA expression of CD36.

The greatest difference in protein abundance between adipocytes, ADSCs and dADSCs was evident at the mass range of 60-80 kDa, however CD36 did not emerge as an abundant protein or as having significantly increased or decreased expression in the shotgun proteomics study conducted in Chapter 4. The removal of *N*-linked glycans prior to mass spectrometry analysis resulted in an increase from 27.54 % to 69.28 % coverage of the CD36 sequence. De-glycosylation of proteins prior to proteomic analysis has previously been described to improve identification of *N*-glycosylated proteins in shotgun experiments (Bailey and Schulz, 2013; Comunale, Mattu *et al.*, 2004). Similarly to this study, new glycoproteins were identified after *N*-linked glycan removal that were not observed in the corresponding glycosylated protein samples prior to de-glycosylation. Given that CD36 contains up to ten *N*-linked glycosylated sites (Febbraio and Silverstein, 2007), removal of *N*-linked glycans was particularly crucial and resulted a 41.74 % increase in sequence coverage and a higher number of peptides identified. Nine of these *N*-linked glycosylation sites present in regions of the protein that were not identified in the analysis of glycosylated samples, most likely due either to incomplete cleavage around the glycosylation site or to the discrepancy between any detected glycopeptide masses and the *in silico* peptide mass matching. *N*-linked glycans have been reported to obscure the action of proteases such as trypsin and AspN (Bailey and Schulz, 2013; Dormeyer, van Hoof *et al.*, 2008a). Lower MS ionisation of glycopeptides compared to un-modified peptides (Leymarie, Griffin *et al.*, 2013) and glycopeptide site heterogeneity resulting in a lower copy number of each peptide (Thaysen-Andersen and Packer, 2014) may also account for the decreased coverage of the glycosylation CD36.

Many glycoproteins, such as integrin alpha-6 and prolow-density lipoprotein receptor-related protein, were not identified at predicted mass on the SDS-PAGE gel matching to their theoretical molecular weight, and some glycoproteins were seen throughout the gel lane. This could be due to the presence of isoforms with varying mass, cleavage products, or possibly protease activity within the sample prior to gel loading. While this is not necessarily a problem in a shotgun SDS-PAGE LC-MS experiment where the results from all gel fractions are be combined to produce a merged list of results, the targeting of particular gel fractions as done in this study must be approached with caution. One way in which this problem could be addressed is through the use of two-dimensional gel electrophoresis (2DE) of intact membrane proteins which can offer protein resolution prior to MS identification (Coorssen, 2013). For CD36, the use of Western blotting and measurement of transcript levels served a dual purpose to confirm expression at the protein level between 60-80 kDa as well as

the differences observed between the samples. CD36 abundance was confirmed as being highest in native adipocytes by both orthogonal techniques, was induced by differentiation and was significantly higher than the ADSCs at the transcript level. Overall the data confirm the crucial need for sample de-glycosylation for accurate identification of glycoprotein abundance by classic shotgun proteomics methods and the need for orthogonal methods to validate any differential protein expression.

While bisecting GlcNAc structures were identified as being present on the majority of adipocyte glycoproteins, to more fully understand their functional relevance, a site-specific global glycoproteomic analysis of glycoproteins carrying bisecting GlcNAc structures should be conducted in future work, in addition to immunoprecipitation and characterisation of targeted glycoproteins of interest. CD36 is an essential component of the fatty acid uptake process and the action of PPARγ (a key mediator in adipogenesis) on CD36 occurs early in adipogenesis, making it one of the key markers of the process (Jefcoate, Wang *et al.*, 2008). In regards to its function, silencing of the CD36 gene in mouse pre-adipocytes has been demonstrated to impair lipid accumulation upon adipogenic differentiation *in vitro* and in mature adipocytes *in vivo* (Christiaens, Van Hul *et al.*, 2012). Impaired fatty acid accumulation in CD36 null mice compared to wild type has also been reported in heart, skeletal muscle and adipose tissue (Coburn, Knapp *et al.*, 2000). As an adipogenic marker therefore, CD36 may be a suitable abundant glycoprotein candidate to determine whether, and or where it is glycosylated with bisecting GlcNAc structures.

5.5 Concluding Remarks

In conclusion, this chapter unearthed new findings to complement the results of previous chapters and highlighted important analytical considerations when investigating the (glyco)proteome. Most importantly, the results strongly promote the de-glycosylation of samples prior to proteomics analysis as an important and crucial consideration for uncovering true *N*-glycoprotein expression and relative expression of all proteins. The strength of integrating various methodologies was also exhibited in this study, to confirm the differential expression of CD36 uncovered by this approach. This glycoprotein was seen to be the most abundant protein in adipocytes and was seen to be induced upon differentiation of adipose-derived stem cells. Overall, given that CD36 and bisecting GlcNAc structures were up-regulated in both *in vivo* and *in vitro* adipogenic cell products, both warrant further investigation at a glycoproteomics based and functional level to determine any potential dual functional role they may share in adipogenesis and the unique biology of adipocytes.

Chapter 6:

General Discussion

6.1 Thesis overview

Since the sequencing of the human genome in the 1990s, numerous studies have been undertaken to investigate gene expression at both the transcriptomic and proteomic levels of analysis. In the current 'post-genomics' era, there is increasing awareness of the importance and role played by the post-translational modifications of proteins. In the context of protein glycosylation for this thesis, a 'reverse genomics' approach was employed, whereby a glycomic analysis was undertaken first to determine the products of the glycosylation genes, followed by a shotgun proteomics study to characterise the products of the protein synthetic genes, with each analysis complemented by investigation of the mRNA expression levels in specific genes of interest.

The modulation of cell surface glycoproteins has been implicated in the differentiation process, and glycosylation changes have been characterised in various stem cell differentiation models (Hamouda, Ullah *et al.*, 2013; Heiskanen, Hirvonen *et al.*, 2009; Hemmoranta, Satomaa *et al.*, 2007; Satomaa, Heiskanen *et al.*, 2009). In the last decade, adipose-derived stem cells (ADSCs) have emerged as a key mesenchymal stem cell (MSC) alternative in stem cell research and therapeutics, compared to the more traditional bone marrow-derived MSCs or MSCs from umbilical cord blood. Stem cells are defined by their unique ability to differentiate into different specialised cells types, however little is known about ADSC differentiation in humans, with the majority of work carried out using a mouse model (Ali, Hochfeld *et al.*, 2013). One of the classical mesenchymal lineages is the adipogenic lineage which ADSCs naturally commit to *in vivo* to become mature adipocytes. Therefore, the aim of this thesis was to use ADSCs and their product native adipocytes to create a model of *in vitro* and *in vivo* adipogenesis with which to investigate stem cell differentiation, specifically to evaluate changes in cell membrane glycobiology that are associated with adipocyte cell differentiation. The characterisation of cell surface glycans in stem cells in the undifferentiated (ADSCs), differentiating (dADSCs) and terminally differentiated (adipocytes) states, using cells from the same donor, offers the opportunity to illuminate the role of glycans as cell markers and signalling effectors. To our knowledge, the use of such a model has not been described before in the literature for any type of stem cell differentiation.

In the opening study of this thesis, *in vitro* adipogenic differentiation of the ADSCs was optimised and evaluated using expression levels of established gene-markers and staining of the generated lipid in the adipose cells. This initial work confirmed that the dADSCs had entered the adipogenic lineage as determined by the standard tests described in the literature, but that the majority of differentiated stem cells did not display the morphology of a mature adipocyte (Chapter 2). The membrane glycome and proteome of ADSCs, *in vitro* differentiated ADSCs and native adipocytes were then

characterised and compared (Chapter 3 and 4). These investigations indicated that while *in vitro* differentiated ADSCs share some characteristics with *in vivo* produced native adipocytes, there were marked differences in the relative abundance of specific glycan sub-classes and proteins between the cell types. Following this, a more targeted glycome and glycoproteome investigation was pursued (Chapter 5) which exposed the importance of glycoprotein de-*N*-glycosylation in shotgun proteomics experiments for optimal identification of the major glycoproteins in adipogenesis. The major findings of this thesis are summarised in the following sub-sections.

6.1.1 Bisecting GlcNAc structures are characteristic of adipocyte N-*linked glycosylation*

The detailed structures of the *N*- and *O*-linked glycans on the membranes of ADSCs and their *in vivo* and *in vitro* adipogenically differentiated progeny were compared to determine the involvement of any specific structures that were markers of adipogenesis. This comprehensive analysis was enabled by the high resolving power of the carbon chromatography (PGC-LC) coupled together with negative mode ESI-MS/MS fragmentation that allowed the separation and characterisation of the many isomeric and isobaric released *N*-linked glycan structures. Although glycosylation changes in stem cell differentiation has been investigated before and described in the literature (including adipogenic differentiation of bone-marrow derived stem cells by Hamouda *et al*., (2013), osteogenic differentiation of bone-marrow derived stem cells by Heiskanen *et al*., (2009) and differentiation of embryonic stem cells into embryoid bodies by Terashima *et al*. (2014)), to our knowledge the protein glycosylation of ADSCs or mature adipocytes has not been investigated previously.

The adipocyte *N*-linked glycome was predominantly comprised of bisecting GlcNAc-type complex structures that were almost absent in the ADSCs. This structural type was seen to significantly increase during *in vitro* adipocyte differentiation (4.8-fold increase) although to a much lesser degree than in mature adipocytes (80.3-fold increase). The high expression of these structures in native adipocytes and their emergence upon differentiation offers a potential new way of evaluating degree of *in vitro* adipogenic differentiation. Specifically, there were three biosynthetically related bisecting GlcNAc structures (structure #36 with composition (Hex)2 (HexNAc)3 + (Man)3(GlcNAc)2, #50 with composition (Hex)2 (HexNAc)3 (Fuc)1 + (Man)3(GlcNAc)2, and #84 with composition (Hex)2 (HexNAc)3 (Fuc)1 (NeuAc)1 + (Man)3(GlcNAc)2) which were significantly increased in native adipocytes and dADSCs compared to ADSCs, and may represent a potential panel of structures to be monitored. Importantly, the increased expression profile of bisecting GlcNAc structures in the differentiated progeny of ADSCs was also reflected in the increased mRNA expression of the MGAT3 gene which is crucial to the synthesis of this structural type because it encodes the glycosyltransferase capable of transferring GlcNAc to the beta-linked mannose of the *N*-linked trimannosyl core. Expression levels of MGAT3 were significantly higher in both differentiated ADSCs

(149-fold increase) and mature adipocytes (279-fold increase) compared to ADSCs. The higher fold changes seen at the transcript level compared to the glycan products may suggest a delay in time from enzyme expression to structural synthesis.

6.1.2 The dADSC membrane proteome shares proteins and processes with native adipocytes

The glycosylation process modulates the function and role of a given protein and therefore a global glycomic analysis alone investigated only one important aspect of the processes underpinning differentiation. After identifying interesting changes in the *N*-glycosylation of ADSCs compared to their *in vitro* and *in vivo* adipogenic progeny, the work focused on the comparison of the membrane proteins in the same cells. Using a shotgun proteomics approach, the membrane proteome of ADSCs, dADSCs and mature adipocytes was analysed and revealed that, qualitatively, there was close similarity between the membrane proteomes of mature adipocytes and differentiated ADSCs particularly when identified proteins that are not part of the plasma membrane (PM) were removed from the analysis. Thus the glycome seems to be a better differentiator of the stage of adipocyte differentiation than the proteome.

The proteins identified and/or up-regulated in both adipocytes or dADSCs (and not in the progenitor ADSCs) were revealed by GO annotation to be involved in adipogenic and lipid-related processes, demonstrating shared adipogenic characteristics between the *in vivo* and *in vitro* differentiation models. For some proteins, such as adipogenic markers Perilipin-1 and -4, up-regulation was further validated by a corresponding increase in mRNA transcript expression. This correlation between protein and mRNA expression was not the case for all marker proteins, for example the increased protein expression of the adipogenic marker FABP4, a marker of early adipogenesis (Jefcoate, Wang *et al.*, 2008) and stem cell marker CD73 was not reflected in the respective mRNA expression. Based on their expression at the mRNA level alone, these two markers in fact indicated that differentiation was still only in the early stages, which agrees with the observation of both the immature adipocyte cell morphology and the low increase in abundance of the bisecting GlcNAc on the *N*-linked glycans of the dADSCs. While this finding did not detract from the overall conclusion that *in vitro* differentiation mimics *in vivo* differentiation to some degree, it also highlighted the importance of using orthogonal techniques to study this process.

6.1.3 Extracellular matrix-related proteins exhibit differential expression during adipogenesis

During stem cell differentiation, the re-arrangement of the extracellular matrix (ECM) is a requirement for re-shaping and re-modelling a cell's morphology, and the expression of different collagen types has been reported to direct cells into specific lineages (Alexeev, Arita *et al.*, 2014; Rodriguez, Montecinos *et al.*, 2000; Sillat, Saat *et al.*, 2012). In the proteomic analysis of membrane

proteins, several different collagen types were identified as exhibiting differential expression across the three cell types, for example while both alpha-1 and -2 chains of collagen type I were found in all cell types, they were in fact down-regulated in dADSCs. These collagen proteins were associated with ECM-related processes according to GO annotations and their detection in this study strongly indicates a role for collagens in the *in vivo* and *in vitro* adipogenic differentiation of ADSCs, supporting previous reports in the literature (Aratani and Kitagawa, 1988; Rodriguez, Montecinos *et al.*, 2000; Sillat, Saat *et al.*, 2012).

6.1.4 Epigenetic processes potentially underpin in vitro *adipogenesis*

In the proteomic analysis, proteins were identified in dADSCs that were not detected in either progenitor ADSCs or mature adipocytes. This is in contrast to the results of the glycomics analysis where no 'unique' structures were identified in either dADSCs or ADSCs. These proteins were grouped into numerous GO category processes that are ultimately found to be involved in epigenetic control of gene expression. At the very least this strongly indicates a role of epigenetic control of *in vitro* differentiation. Epigenetic regulation of adipogenesis has previously been reported (Li, Xiao *et al.*, 2010; Musri and Parrizas, 2012; Noer, Sorensen *et al.*, 2006) and therefore our results contribute to this growing area. It has been proposed that further knowledge of epigenetic mechanisms governing adipogenesis may lead to potential therapy targets in the prevention and treatment of obesity *in vivo* (Li, Xiao *et al.*, 2010) and of metabolic disorders in which adipocytes play key roles.

6.1.5 Bisecting GlcNAc N-*linked glycans are carried by most adipocyte proteins*

A more targeted approach was employed to further investigate the expression of bisecting GlcNAc *N*-linked structures, particularly on adipocyte membrane glycoproteins, by releasing and analysing *N*-linked glycans from membrane proteins separated by SDS-PAGE. The glycan profiles of the mass separated proteins revealed that the expression of bisecting GlcNAc structures is not protein-specific, rather they are present on the majority of adipocyte glycoproteins regardless of molecular weight. Precise glycosylation site occupancy of bisecting GlcNAc on specific proteins will involve further work using the recently developed glycoproteomic analytical techniques applied to complex mixtures of proteins, which is currently starting to appear in the literature (Parker, Thaysen-Andersen *et al.*, 2013; Thaysen-Andersen, Venkatakrishnan *et al.*, 2015).

6.2 Analytical considerations

6.2.1 Adipose samples and processing

Several of the methodologies and protocols utilised during the course of this thesis have inherent technical difficulties, and had to be modified for application to the unique physiology of fat cells. The

fragility of adipocytes makes their collection from lipo-aspirates challenging and their high lipid content required specialised downstream sample processing to isolate the membrane fractions without fat contamination. Firstly, care was taken to maximise collection of adipocytes to collect areas of the floating adipocyte supernatant with the least amount of lipid from burst cells. Downstream from this, preparation of the cells for analysis first required removal of the lipid contents which required extra high speed centrifugation steps, a step not required in the lysis of ADSCs and dADSCs. Where possible, and depending on sample availability, lipo-aspirates were also collected from the same location in the body for each biological replicates for consistency. The three lipo-aspirate samples used to derive cells for the glycomic and proteomic analyses were all from the abdomen of the donors, whilst the three samples used for qRT-PCR and Western blotting originated from the abdomen, hip, and flank areas. The statistical analysis undertaken on adipocytes from the latter samples displayed low standard deviation values, indicating low variation between the samples despite originating from different areas of the body.

6.2.2 Membrane protein isolation

PM proteins play crucial roles in cell-cell signalling and a cell's interaction with its environment (Cordwell and Thingholm, 2009). However, the enrichment of membrane proteins is not without challenges as is the downstream analysis of these proteins by using mass spectrometry. The membrane isolation method used in this thesis produced a protein population containing 67.1 % of membrane proteins which was in good accord with previous methods (Dormeyer, van Hoof *et al.*, 2008b; Hao, Li *et al.*, 2013). A recent report described an optimised PM isolation method from cultured cells that did not rely on the use of an ultracentrifuge, instead performing the membrane isolation and collection directly from the cell culture vessel (Mun, Lee *et al.*, 2013). This procedure resulted in a reduced percentage of high mannose glycan structures on the enriched proteins. High mannose glycans typically originate largely from intracellular endoplasmic reticulum (ER) and Golgi membrane proteins, and the method allowed the researchers to better determine the true degree of high mannose structures actually on the PM. Such a procedure that is optimised specifically for glycomics analysis of PM protein glycosylation would be worth consideration in future membrane glycomics investigations of *in vitro* cultured cells. Additional alternatives include the use of a discontinuous sucrose gradient to fractionate cellular organelles and selectively avoid inclusion of ER and Golgi membrane proteins in membrane proteomics studies (Lin, Chik *et al.*, 2015), or the targeted enrichment of PM proteins through biotin labelling of surface protein amino acid residues following by magnetic streptavidin bead affinity enrichment (Zhao, Zhang *et al.*, 2004).

6.2.3 Shotgun proteomics underestimates the abundance of highly glycosylated proteins

Protein identification of the gel-separated de-*N*-glycosylated proteins revealed differential expression of glycoproteins by native adipocytes, ADSCs and dADSCs that was not seen in the shotgun proteomics experiments on the total membrane proteins of these cells. Specifically, the results indicated that an adipogenic marker, the glycoprotein CD36, that displayed the greatest degree of differential expression between the cell types using this molecular mass-based sample preparation, is by far the most abundant glycoprotein in adipocytes as well as being up-regulated upon differentiation of the ADSCs. Critically, as a highly glycosylated protein, CD36 was not detected as a high abundance protein in the corresponding non-de-glycosylated membrane protein samples analysed by shotgun proteomics after gel separation. Other glycoproteins including calreticulin, lactadherin, erlin-2, mannose-6-phosphate receptor and decorin were also not detected in the whole sample proteomic analysis without de-glycosylation. This finding emphasises the need for a de-glycosylation step in proteomics experiments in order to achieve accurate glycoprotein expression data. This observation has also been reported recently by Bailey and Schulz (2013) and has previously been employed sparingly in previous proteomics studies (Dormeyer, van Hoof *et al.*, 2008a; Dormeyer, van Hoof *et al.*, 2008b; Jessani, Humphrey *et al.*, 2004).

6.2.4 1D SDS-PAGE separation affects protein resolution and targeted proteomics approaches

Many glycoproteins were identified in the 1D SDS-PAGE separation carried out in Chapter 5 in gel areas which differed from their predicted, theoretical mass, or were identified at multiple molecular weights. This may be due to proteolytic events or the presence of protein isoforms, and ultimately may have impacted the subsequent targeted analysis of gel areas displaying differential expression. To combat this in future experiments, the use of a high resolving 2D SDS-PAGE as part of a top down proteomic (Coorssen, 2013) approach may circumvent this problem.

6.3 Future directions and concluding remarks

To our knowledge, the work presented in this thesis is the first to investigate changes in the glycome and proteome of cell membranes during *in vivo* and *in vitro* adipose-derived stem cell adipogenesis. Several studies have been undertaken to investigate cellular changes during differentiation, however none have done so using both an *in vitro* and *in vivo* differentiation model and therein lies the strength of the model used here.

In addition to the important technical and analytical considerations encountered throughout this thesis, various biological insights were uncovered in the different analyses undertaken. The systems glycobiology approach used here revealed different degrees of change at the different 'omics levels,

despite sharing a glycosylation-based commonality. An example of this was the pronounced way in which *N*-linked glycosylation of adipocytes and dADSCs were both directed to a higher percentage of bisecting GlcNAc-type structures, whilst changes in the membrane proteome of the adipogenic products indicated common yet also divergent changes. Attempts to complement the glycomic and proteomic findings through measurement of the relevant gene transcript expression indicated that gene expression and protein product expression are not always correlated. As hypothesised, the differences in the *in vivo* and *in vitro* differentiation environments produced adipogenic products that differed in their protein glycosylation as unearthed by glycomics, proteomics and transcriptomics analyses and suggest that the *in vitro* differentiation process does not produce the full native adipocyte phenotype.

A major recommended area of future work would be an investigation into the role of the major protein glycan component of bisecting GlcNAc structures in adipocytes and in adipogenesis, particularly as a potential target for the *in vivo* treatment of obesity. The correlation of MGAT3 expression with bisecting GlcNAc expression in this study provides supporting data towards the involvement of membrane protein glycosylation in stem cell adipogenesis. This finding offers a potential avenue for determining the functional relevance of these structures; for example, through determining whether the silencing of MGAT3 has any effect on ADSC differentiation. For a further understanding of the function of bisecting GlcNAc structures on *N*-linked glycoproteins, it would also be of interest to confirm and locate the bisecting GlcNAc structures on the adipogenic marker CD36. CD36 has been shown to play an important role in lipid accumulation in adipogenic differentiation (Christiaens, Van Hul *et al.*, 2012) and was the most abundant glycoprotein identified in adipocytes in this thesis, making it a promising candidate for a targeted glycoproteomic site-specific analysis. Although glycoproteomic approaches are often overlooked due to limitations of available technology for such analysis, efforts to further facilitate system-wide glycoproteomics analyses are growing (Thaysen-Andersen and Packer, 2014). Global glycomics analyses alone (like that undertaken here) provide structural detail on the glycan repertoire of a cell but offer no glycosylation site or protein information whilst proteomics analyses are limited by the lower detection and impaired identification of glycoproteins (as seen here). A systems glycobiology approach is therefore required for ascertaining the functional significance of glycosylation in adipogenesis. As prior knowledge of the glycome and proteome is important in the interpretation of glycopeptide-based glycoproteomics data, the work conducted in this thesis provides a suitable platform for a future global glycoproteomics analysis. In addition, considerations for future work could include:

- ***Co-culturing of ADSCs with native adipocytes***. In addition to the application of the differentiating cocktail. This may help to create an environment that more closely mimics the stem cell niche where ADSCs undergo adipogenic differentiation *in vivo*.
- ***Inclusion of donor's gender, age or BMI in the analysis***. Unfortunately, due to constraints outlined in the original ethics application, this information was not accessible for the lipo-aspirates used in this study. Such information may be useful when adipocytes or adipose-derived stem cells are used to study metabolic disorders or obesity.
- ***Application of new culturing techniques for mesenchymal stem cells including the use of xenogenic (non-human) serum-free media***. This possibility is worth investigation for future work, particularly when stem cells are differentiated with the intention for use in reconstructive surgery.
- ***Inclusion of a protein de-glycosylation step prior to trypsin digestion and MS analysis to assist in protein identification***. This step is a straightforward addition to a typical shotgun proteomics workflow and offers improved proteome coverage regardless of whether glycoproteins are the focus of the study.

In regards to the translational aim of differentiating ADSCs *in vitro* for use in regenerative or cosmetic medicine, the results presented in this thesis indicate that despite the fact that approximately 75 % of ADSCs appear to undergo differentiation *in vitro*, the glycome of these adipogenic products differs greatly from the native mature adipocyte, suggesting possible consequences for their use as a therapeutic. The different glycan surface profile of *in vitro* differentiated cells, if used as a therapeutic, may affect the cell-cell interactions, membrane transport and growth characteristics and therefore their efficacy and risk profile. Adipogenic differentiation of murine ADSCs over 7 days only, using the same adipogenic media components as in this study, has been shown to result in dADSCs that are able to vascularise and successfully re-incorporate when injected to the fat pads of mice (Mizuno, Itoi *et al.*, 2008). Based on the immature morphology and unchanged glycomic profiles of Day 7 dADSCs in this thesis, this finding suggests potentially that either glycosylation may not play a vital role in this system, that the authors achieved a more mature stage of differentiation in the murine ADSCs prior to injection, or that further glycosylation changes may occur after injection into the adipose tissue environment. Previous work has demonstrated different degrees of adipogenic differentiation depending on the species origin of the ADSCs and also the serum used in the culture medium, for example human ADSCs showed the greatest percentage of differentiating cells when using fetal bovine serum (FBS) while ADSCs from cows, dogs, horse showed higher rates of differentiation when using rabbit serum (Masters' thesis; H. Hinneburg 2010). The work of Mizuno *et al*. (2008) involved the use of FBS, which was also used in this thesis, however the difference in

species origin of ADSCs may explain why perhaps more differentiation might have been seen in the murine model than reported in this thesis. A glycomic analysis of differentiated ADSCs over a time course post-incorporation of the dADSCs into the murine fat pad may help shed light on any glycosylation changes. Distinguishing the re-incorporated cells from original cells could be achieved by first labelling the cells *in vitro* with labels such as super-paramagnetic iron oxide particles or carboxylated nanodiamonds, which have been demonstrated to be biocompatible with ADSCs (Blaber, Hill *et al.*, 2013).

Overall, the use of different levels of 'omics investigations, namely glycomic, transcriptomic and proteomic analyses in this study has formed a more holistic systems glycobiology approach to explore the role of glycosylation in adipogenesis. The work here provides a framework for the future investigation of adipose-derived stem cells and adipocytes, their related pathologies and potential therapeutic applications.

References

Abdul Rahman S, Bergstrom E, Watson CJ, Wilson KM, Ashford DA, Thomas JR, Ungar D, Thomas-Oates JE. 2014. Filter-aided N-glycan separation (FANGS): a convenient sample preparation method for mass spectrometric N-glycan profiling. *J Proteome Res*, 13:1167-1176.

Alexeev V, Arita M, Donahue A, Bonaldo P, Chu ML, Igoucheva O. 2014. Human adipose-derived stem cell transplantation as a potential therapy for collagen VI-related congenital muscular dystrophy. *Stem Cell Res Ther*, 5:21.

Ali AT, Hochfeld WE, Myburgh R, Pepper MS. 2013. Adipocyte and adipogenesis. *Eur J Cell Biol*, 92:229-236.

Alpert AJ. 1990. Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *Journal of Chromatography A*, 499:177-196.

Amos PJ, Kapur SK, Stapor PC, Shang H, Bekiranov S, Khurgel M, Rodeheaver GT, Peirce SM, Katz AJ. 2010. Human adipose-derived stromal cells accelerate diabetic wound healing: impact of cell formulation and delivery. *Tissue Eng Part A*, 16:1595-1606.

An HJ, Gip P, Kim J, Wu S, Park KW, McVaugh CT, Schaffer DV, Bertozzi CR, Lebrilla CB. 2012. Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells. *Mol Cell Proteomics*, 11:M111 010660.

Anugraham M, Jacob F, Nixdorf S, Everest-Dass AV, Heinzelmann-Schwarz V, Packer NH. 2014. Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: glycan structures reflect gene expression and DNA methylation status. *Mol Cell Proteomics*.

Aratani Y, Kitagawa Y. 1988. Enhanced synthesis and secretion of type IV collagen and entactin during adipose conversion of 3T3-L1 cells and production of unorthodox laminin complex. *J Biol Chem*, 263:16163-16169.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT*, et al.* 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, 25:25-29.

Baer PC, Geiger H. 2012. Adipose-derived mesenchymal stromal/stem cells: tissue localization, characterization, and heterogeneity. *Stem Cells Int*, 2012:812693.

Bailey UM, Schulz BL. 2013. Deglycosylation systematically improves N-glycoprotein identification in liquid chromatography-tandem mass spectrometry proteomics for analysis of cell wall stress responses in Saccharomyces cerevisiae lacking Alg3p. *J Chromatogr B Analyt Technol Biomed Life Sci*, 923-924:16-21.

Bantscheff M, Lemeer S, Savitski MM, Kuster B. 2012. Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Analytical and Bioanalytical Chemistry*, 404:939-965.

Barrera NP, Robinson CV. 2011. Advances in the Mass Spectrometry of Membrane Proteins: From Individual Proteins to Intact Complexes. *Annual Review of Biochemistry*, 80:247-271.

Bashkatov AN, Genina ÉA, Kochubey VI, Tuchin VV. 2005. Optical properties of the subcutaneous adipose tissue in the spectral range 400-2500 nm. *Optics and Spectroscopy*, 99:836-842.

Benjamini Y, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57:289-300.

Bern M, Brito AE, Pang PC, Rekhi A, Dell A, Haslam SM. 2013. Polylactosaminoglycan glycomics: enhancing the detection of high-molecular-weight N-glycans in matrix-assisted laser desorption ionization time-of-flight profiles by matched filtering. *Mol Cell Proteomics*, 12:996-1004.

Blaber SP, Hill CJ, Webster RA, Say JM, Brown LJ, Wang S-C, Vesey G, Herbert BR. 2013. Effect of Labeling with Iron Oxide Particles or Nanodiamonds on the Functionality of Adipose-Derived Mesenchymal Stem Cells. *PLoS One*, 8:e52997.

Blaber SP, Webster RA, Hill CJ, Breen EJ, Kuah D, Vesey G, Herbert BR. 2012. Analysis of in vitro secretion profiles from adipose-derived cell populations. *J Transl Med*, 10:172.

Bordier C. 1981. Phase separation of integral membrane proteins in Triton X-114 solution. *J Biol Chem*, 256:1604-1607.

Bowers RR, Lane MD. 2008. Wnt signaling and adipocyte lineage commitment. *Cell Cycle*, 7:1191-1196.

Bradford MM. 1976. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Analytical Biochemistry*.

Brzoska AJ, Hassan KA. 2014. Quantitative PCR for detection of mRNA and gDNA in environmental isolates. *Methods Mol Biol*, 1096:25-42.

Bunnell BA, Estes BT, Guilak F, Gimble JM. 2008. Differentiation of adipose stem cells. *Methods Mol Biol*, 456:155-171.

Bunnell BA, Flaat M, Gagliardi C, Patel B, Ripoll C. 2008. Adipose-derived stem cells: isolation, expansion and differentiation. *Methods*, 45:115-120.

Callewaert N, Geysens S, Molemans F, Contreras R. 2001. Ultrasensitive profiling and sequencing of N-linked oligosaccharides using standard DNA-sequencing equipment. *Glycobiology*, 11:275-281.

Campbell MP, Nguyen-Khuong T, Hayes CA, Flowers SA, Alagesan K, Kolarich D, Packer NH, Karlsson NG. 2014. Validation of the curation pipeline of UniCarb-DB: Building a global glycan reference MS/MS repository. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844:108-116.

Cao J, Shen C, Wang H, Shen H, Chen Y, Nie A, Yan G, Lu H, Liu Y, Yang P. 2009. Identification of N-glycosylation sites on secreted proteins of human hepatocellular carcinoma cells with a complementary proteomics approach. *J Proteome Res*, 8:662-672.

Capela A, Temple S. 2006. LeX is expressed by principle progenitor cells in the embryonic nervous system, is secreted into their environment and binds Wnt-1. *Dev Biol*, 291:300-313.

Carbon S, Ireland A, Mungall CJ, Shu S, Marshall B, Lewis S, the Ami GOH, the Web Presence Working G. 2009. AmiGO: online access to ontology and annotation data. *Bioinformatics*, 25:288-289.

Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. 2008. GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res*, 7:1650-1659.

Chamrad DC, Blueggel M, Koerting G, Glandorf J, Vagts J, Hufnagel P, Thiele H. 2007. P5-M Proteinscape - Software Platform for Managing Proteomics Data. *Journal of Biomolecular Techniques : JBT*, 18:2-3.

Chik JH, Zhou J, Moh ES, Christopherson R, Clarke SJ, Molloy MP, Packer NH. 2013. Comprehensive glycomics comparison between colon cancer cell cultures and tumours: Implications for biomarker studies. *J Proteomics*, 108C:146-162.

Christiaens V, Van Hul M, Lijnen HR, Scroyen I. 2012. CD36 promotes adipocyte differentiation and adipogenesis. *Biochim Biophys Acta*, 1820:949-956.

Christiansen MN, Chik J, Lee L, Anugraham M, Abrahams JL, Packer NH. 2014. Cell surface protein glycosylation in cancer. *Proteomics*, 14:525-546.

Christodoulides C, Lagathu C, Sethi JK, Vidal-Puig A. 2009. Adipogenesis and WNT signalling. *Trends in Endocrinology & Metabolism*, 20:16-24.

Chu CS, Niñonuevo MR, Clowers BH, Perkins PD, An HJ, Yin H, Killeen K, Miyamoto S, Grimm R, Lebrilla CB. 2009. Profile of Native N-linked Glycan Structures from Human Serum Using High Performance Liquid Chromatography on a Microfluidic Chip and Time-of-Flight Mass Spectrometry. *Proteomics*, 9:1939-1951.

Coburn CT, Knapp FF, Jr., Febbraio M, Beets AL, Silverstein RL, Abumrad NA. 2000. Defective uptake and utilization of long chain fatty acids in muscle and adipose tissues of CD36 knockout mice. *J Biol Chem*, 275:32523-32529.

Comunale MA, Mattu TS, Lowman MA, Evans AA, London WT, Semmes OJ, Ward M, Drake R, Romano PR, Steel LF*, et al.* 2004. Comparative proteomic analysis of de-N-glycosylated serum from hepatitis B carriers reveals polypeptides that correlate with disease status. *Proteomics*, 4:826-838.

Cooper CA, Gasteiger E, Packer NH. 2001. GlycoMod--a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics*, 1:340-349.

Coorssen JR. 2013. Top-down proteomics: 2D gels are an integral part of the process. In: Almeida A, Eckersall D, Bencurova E, Dolinska S, Mlynarcik P, Vincova M, Bhide M editors. Farm animal proteomics 2013: Proceedings of the 4th Management Committee Meeting and 3rd Meeting of Working Groups 1, 2 & 3 of COST Action FA1002 Košice, Slovakia 25-26 April 2013. Wageningen: Wageningen Academic Publishers. p. 31-33.

Cordwell SJ, Thingholm TE. 2009. Technologies for plasma membrane proteomics. *Proteomics*, 10:611-627.

Craig R, Beavis RC. 2004. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*, 20:1466-1467.

Craig R, Cortens JP, Beavis RC. 2004. Open source system for analyzing, validating, and storing protein identification data. *J Proteome Res*, 3:1234-1242.

Cuaranta-Monroy I, Nagy L. 2013. PPARgamma needs a helping hand to make fat. *Cell Death Differ*, 20:1599-1600.

De Gemmis P, Lapucci C, Bertelli M, Tognetto A, Fanin E, Vettor R, Pagano C, Pandolfo M, Fabbri A. 2006. A real-time PCR approach to evaluate adipogenic potential of amniotic fluid-derived human mesenchymal stem cells. *Stem Cells Dev*, 15:719-728.

DeLany JP, Floyd ZE, Zvonic S, Smith A, Gravois A, Reiners E, Wu X, Kilroy G, Lefevre M, Gimble JM. 2005. Proteomic analysis of primary cultures of human adipose-derived stem cells: modulation by Adipogenesis. *Mol Cell Proteomics*, 4:731-740.

Dennis G, Jr., Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. 2003. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*, 4:P3.

Desantos-Garcia JL, Khalil SI, Hussein A, Hu Y, Mechref Y. 2011. Enhanced sensitivity of LC-MS analysis of permethylated N-glycans through online purification. *Electrophoresis*, 32:3516-3525.

Desarzens S, Liao WH, Mammi C, Caprio M, Faresse N. 2014. Hsp90 blockers inhibit adipocyte differentiation and fat mass accumulation. *PLoS One*, 9:e94127.

Dominici M, Le Blanc K, Mueller I, Slaper-Cortenbach I, Marini F, Krause D, Deans R, Keating A, Prockop D, Horwitz E. 2006. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy*, 8:315-317.

Domon B, Costello C. 1988. A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal*, 5:397-409.

Dormeyer W, van Hoof D, Braam SR, Heck AJ, Mummery CL, Krijgsveld J. 2008a. Plasma membrane proteomics of human embryonic stem cells and human embryonal carcinoma cells. *J Proteome Res*, 7:2936-2951.

Dormeyer W, van Hoof D, Mummery CL, Krijgsveld J, Heck AJ. 2008b. A practical guide for the identification of membrane and plasma membrane proteins in human embryonic stem cells and human embryonal carcinoma cells. *Proteomics*, 8:4036-4053.

Dube DH, Prescher JA, Quang CN, Bertozzi CR. 2006. Probing Mucin-Type O-Linked Glycosylation in Living Animals. *Proc. Natl. Acad. Sci.*, 103:4819-4824.

Eng JK, McCormack AL, Yates JR. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom*, 5:976-989.

Estrella RP, Whitelock JM, Roubin RH, Packer NH, Karlsson NG. 2009. Small-scale enzymatic digestion of glycoproteins and proteoglycans for analysis of oligosaccharides by LC-MS and FACE gel electrophoresis. *Methods Mol Biol*, 534:171-192.

Everest-Dass AV, Abrahams JL, Kolarich D, Packer NH, Campbell MP. 2013. Structural Feature Ions for Distinguishing N- and O-Linked Glycan Isomers by LC-ESI-IT MS/MS. *J Am Soc Mass Spectrom*.

Everest-Dass AV, Jin D, Thaysen-Andersen M, Nevalainen H, Kolarich D, Packer NH. 2012. Comparative structural analysis of the glycosylation of salivary and buccal cell proteins: innate protection against infection by Candida albicans. *Glycobiology*, 22:1465-1479.

Febbraio M, Silverstein RL. 2007. CD36: implications in cardiovascular disease. *Int J Biochem Cell Biol*, 39:2012-2030.

Fernandes D, Spencer D. 2008. Glycoprotein Analysis. In: Fraser-Reid BO, Tatsuta K, Thiem J editors. Glycoscience: Springer Berlin Heidelberg. p. 1873-1884.

Ferris WF, Crowther NJ. 2011. Once fat was fat and that was that: our changing perspectives on adipose tissue. *Cardiovasc J Afr*, 22:147-154.

Fink T, Lund P, Pilgaard L, Rasmussen JG, Duroux M, Zachar V. 2008. Instability of standard PCR reference genes in adipose-derived stem cells during propagation, differentiation and hypoxic exposure. *BMC Mol Biol*, 9:98.

Fruhbeck G. 2008. Overview of adipose tissue and its role in obesity and metabolic disorders. *Methods Mol Biol*, 456:1-22.

Fujitani N, Furukawa J, Araki K, Fujioka T, Takegawa Y, Piao J, Nishioka T, Tamura T, Nikaido T, Ito M*, et al.* 2013. Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers. *Proc Natl Acad Sci U S A*, 110:2105-2110.

Galateanu B, Dinescu S, Cimpean A, Dinischiotu A, Costache M. 2012. Modulation of Adipogenic Conditions for Prospective Use of hADSCs in Adipose Tissue Engineering. *Int J Mol Sci*, 13:15881-15900.

Galili U. 2013. Anti-Gal: an abundant human natural antibody of multiple pathogeneses and clinical benefits. *Immunology*, 140:1-11.

Gammulla CG, Pascovici D, Atwell BJ, Haynes PA. 2011. Differential proteomic response of rice (Oryza sativa) leaves exposed to high- and low-temperature stress. *Proteomics*, 11:2839-2850.

Gimble JM, Bunnell BA, Guilak F. 2012. Human adipose-derived cells: an update on the transition to clinical translation. *Regenerative Medicine*, 7:225-235.

Gorzelniak K, Engeli S, Sharma AM. 1998. Standardizing the counting of adipocytes in cell culture. *Biotechniques*, 24:536-538.

Greenbaum D, Colangelo C, Williams K, Gerstein M. 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol*, 4:117.

Greenberg AS, Obin MS. 2006. Obesity and the role of adipose tissue in inflammation and metabolism. *Am J Clin Nutr*, 83:461S-465S.

Gregoire FM. 2001. Adipocyte differentiation: from fibroblast to endocrine cell. *Exp Biol Med (Maywood)*, 226:997-1002.

Gregoire FM, Smas CM, Sul HS. 1998. Understanding adipocyte differentiation. *Physiol Rev*, 78:783-809.

Gronthos S, Franklin DM, Leddy HA, Robey PG, Storms RW, Gimble JM. 2001. Surface protein characterization of human adipose tissue-derived stromal cells. *J Cell Physiol*, 189:54-63.

Guan H, Yang K. 2008. RNA isolation and real-time quantitative RT-PCR. *Methods Mol Biol*, 456:259-270.

Gundry RL, Boheler KR, Van Eyk JE, Wollscheid B. 2008. A novel role for proteomics in the discovery of cell-surface markers on stem cells: Scratching the surface. *Proteomics Clin Appl*, 2:892-903.

Gundry RL, White MY, Murray CI, Kane LA, Fu Q, Stanley BA, Van Eyk JE. 2009. Preparation of Proteins and Peptides for Mass Spectrometry Analysis in a Bottom-Up Proteomics Workflow. *Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.]*, CHAPTER:Unit10.25-Unit10.25.

Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotech*, 17:994-999.

Gygi SP, Rochon Y, Franza BR, Aebersold R. 1999. Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol*, 19:1720-1730.

Hamouda H, Ullah M, Berger M, Sittinger M, Tauber R, Ringe J, Blanchard V. 2013. N-glycosylation profile of undifferentiated and adipogenically differentiated human bone marrow mesenchymal stem cells: towards a next generation of stem cell markers. *Stem Cells Dev*, 22:3100-3113.

Hanisch F-G, Breloy I. 2009. Protein-specific glycosylation: signal patches and cis-controlling peptidic elements. *Biological Chemistry*, 390:619.

Hao J, Li W, Dan J, Ye X, Wang F, Zeng X, Wang L, Wang H, Cheng Y, Liu L*, et al.* 2013. Reprogramming- and pluripotency-associated membrane proteins in mouse stem cells revealed by label-free quantitative proteomics. *J Proteomics*, 86:70-84.

Hart GW, Copeland RJ. 2010. Glycomics Hits the Big Time. *Cell*, 143:672-676.

Harvey DJ. 2005. Fragmentation of negative ions from carbohydrates: part 3. Fragmentation of hybrid and complex N-linked glycans. *J Am Soc Mass Spectrom*, 16:647-659.

Harvey DJ, Crispin M, Scanlan C, Singer BB, Lucka L, Chang VT, Radcliffe CM, Thobhani S, Yuen C-T, Rudd PM. 2008. Differentiation between isomeric triantennary N-linked glycans by negative ion tandem mass spectrometry and confirmation of glycans containing galactose attached to the bisecting (β1-4-GlcNAc) residue in N-glycans from IgG. *Rapid Communications in Mass Spectrometry*, 22:1047-1052.

Harvey DJ, Royle L, Radcliffe CM, Rudd PM, Dwek RA. 2008. Structural and quantitative analysis of N-linked glycans by matrix-assisted laser desorption ionization and negative ion nanospray mass spectrometry. *Anal Biochem*, 376:44-60.

Hasehira K, Tateno H, Onuma Y, Ito Y, Asashima M, Hirabayashi J. 2012. Structural and quantitative evidence for dynamic glycome shift on production of induced pluripotent stem cells. *Mol Cell Proteomics*, 11:1913-1923.

Heiskanen A, Hirvonen T, Salo H, Impola U, Olonen A, Laitinen A, Tiitinen S, Natunen S, Aitio O, Miller-Podraza H*, et al.* 2009. Glycomics of bone marrow-derived mesenchymal stem cells can be used to evaluate their cellular differentiation stage. *Glycoconj J*, 26:367-384.

Hemmoranta H, Satomaa T, Blomqvist M, Heiskanen A, Aitio O, Saarinen J, Natunen J, Partanen J, Laine J, Jaatinen T. 2007. N-glycan structures and associated gene expression reflect the

characteristic N-glycosylation pattern of human hematopoietic stem and progenitor cells. *Exp Hematol*, 35:1279-1292.

Hennen E, Faissner A. 2012. LewisX: a neural stem cell specific glycan? *Int J Biochem Cell Biol*, 44:830-833.

Hoja-Lukowicz D, Link-Lenczowski P, Carpentieri A, Amoresano A, Pochec E, Artemenko KA, Bergquist J, Litynska A. 2013. L1CAM from human melanoma carries a novel type of N-glycan with Galbeta1-4Galbeta1- motif. Involvement of N-linked glycans in migratory and invasive behaviour of melanoma cells. *Glycoconj J*, 30:205-225.

Huang DW, Sherman BT, Lempicki RA. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*, 37:1-13.

Huang DW, Sherman BT, Lempicki RA. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*, 4:44-57.

Hung SC, Chang CF, Ma HL, Chen TH, Low-Tone Ho L. 2004. Gene expression profiles of early adipogenesis in human mesenchymal stem cells. *Gene*, 340:141-150.

Illouz Y-G, Sterodimas A, Aktan T, Duman S, Cihantimur B. 2011. Cellular and Molecular Aspects of Adipose Tissue. Adipose Stem Cells and Regenerative Medicine: Springer Berlin Heidelberg. p. 1-12.

Isaji T, Gu J, Nishiuchi R, Zhao Y, Takahasi M, Miyoshi E, Honke K, Sekiguchi K, Taniguchi N. 2004. Introduction of bisecting GlcNAc into integrin α5β1 reduces ligand binding and down-regulates cell adhesion and cell migration. *J Biol Chem*, 279:19747-19754.

James AW. 2013. Review of Signaling Pathways Governing MSC Osteogenic and Adipogenic Differentiation. *Scientifica (Cairo)*, 2013:684736.

Jefcoate CR, Wang S, Liu X. 2008. Methods that resolve different contributions of clonal expansion to adipogenesis in 3T3-L1 and C3H10T1/2 cells. *Methods Mol Biol*, 456:173-193.

Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M*, et al.* 2009. STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res*, 37:D412-416.

Jensen PH, Karlsson NG, Kolarich D, Packer NH. 2012. Structural analysis of N- and O-glycans released from glycoproteins. *Nat Protoc*, 7:1299-1310.

Jeong JA, Ko KM, Park HS, Lee J, Jang C, Jeon CJ, Koh GY, Kim H. 2007. Membrane proteomic analysis of human mesenchymal stromal cells during adipogenesis. *Proteomics*, 7:4181-4191.

Jessani N, Humphrey M, McDonald WH, Niessen S, Masuda K, Gangadharan B, Yates JR, 3rd, Mueller BM, Cravatt BF. 2004. Carcinoma and stromal enzyme activity profiles associated with breast tumor growth in vivo. *Proc Natl Acad Sci U S A*, 101:13756-13761.

Jewell JL, Oh E, Thurmond DC. 2010. Exocytosis mechanisms underlying insulin release and glucose uptake: conserved roles for Munc18c and syntaxin 4. *Am J Physiol Regul Integr Comp Physiol*, 298:R517-531.

Johnson RS, Davis MT, Taylor JA, Patterson SD. 2005. Informatics for protein identification by mass spectrometry. *Methods*, 35:223-236.

Joshi-Tope G, Gillespie M, Vastrik I, D'Eustachio P, Schmidt E, de Bono B, Jassal B, Gopinath GR, Wu GR, Matthews L*, et al.* 2005. Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res*, 33:D428-432.

Josic D, Clifton JG. 2007. Mammalian plasma membrane proteomics. *Proteomics*, 7:3010-3029.

Kailemia MJ, Ruhaak LR, Lebrilla CB, Amster IJ. 2013. Oligosaccharide Analysis by Mass Spectrometry: A Review of Recent Developments. *Analytical Chemistry*, 86:196-212.

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu T*, et al.* 2008. KEGG for linking genomes to life and the environment. *Nucleic Acids Res*, 36:D480-484.

Kasper BT, Koppolu S, Mahal LK. 2014. Insights into miRNA regulation of the human glycome. *Biochem Biophys Res Commun*, 445:774-779.

Katrlik J, Svitel J, Gemeiner P, Kozar T, Tkac J. 2010. Glycan and lectin microarrays for glycomics and medicinal applications. *Med Res Rev*, 30:394-418.

Katz AJ, Tholpady A, Tholpady SS, Shang H, Ogle RC. 2005. Cell Surface and Transcriptional Characterization of Human Adipose-Derived Adherent Stromal (hADAS) Cells. *Stem Cells*, 23:412-423.

Kelleher NL. 2004. Peer Reviewed: Top-Down Proteomics. *Analytical Chemistry*, 76:196 A-203 A.

Kenny D, A. Hayes C, Jin C, G. Karlsson N. 2011. Perspective and Review of Mass Spectrometric Based Sulfoglycomics of N-Linked and O-Linked Oligosaccharides. *Current Proteomics*, 8:278-296.

Kenny DT, Issa SMA, Karlsson NG. 2011. Sulfate migration in oligosaccharides induced by negative ion mode ion trap collision-induced dissociation. *Rapid Communications in Mass Spectrometry*, 25:2611-2618.

Kern S, Eichler H, Stoeve J, Kluter H, Bieback K. 2006. Comparative analysis of mesenchymal stem cells from bone marrow, umbilical cord blood, or adipose tissue. *Stem Cells*, 24:1294-1301.

Kheterpal I, Ku G, Coleman L, Yu G, Ptitsyn AA, Floyd ZE, Gimble JM. 2011. Proteome of human subcutaneous adipose tissue stromal vascular fraction cells versus mature adipocytes based on DIGE. *J Proteome Res*, 10:1519-1527.

Kim E, Kim W, Oh K-J, Han B, Lee S, Bae K-H. 2015. Recent Advances in Proteomic Studies of Adipose Tissues and Adipocytes. *International Journal of Molecular Sciences*, 16:4581-4599.

Kim M, Kim I, Lee SK, Bang SI, Lim SY. 2011. Clinical trial of autologous differentiated adipocytes from stem cells derived from human adipose tissue. *Dermatol Surg*, 37:750-759.

Klapoetke S, Zhang J, Becht S, Gu X, Ding X. 2010. The evaluation of a novel approach for the profiling and identification of N-linked glycan with a procainamide tag by HPLC with fluorescent and mass spectrometric detection. *Journal of Pharmaceutical and Biomedical Analysis*, 53:315-324.

Kolf CM, Cho E, Tuan RS. 2007. Mesenchymal stromal cells. Biology of adult mesenchymal stem cells: regulation of niche, self-renewal and differentiation. *Arthritis Res Ther*, 9:204.

Komoda H, Okura H, Lee CM, Sougawa N, Iwayama T, Hashikawa T, Saga A, Yamamoto-Kakuta A, Ichinose A, Murakami S*, et al.* 2010. Reduction of N-glycolylneuraminic acid xenoantigen on human adipose tissue-derived stromal cells/mesenchymal stem cells leads to safer and more useful cell sources for various stem cell therapies. *Tissue Eng Part A*, 16:1143-1155.

Kozak RP, Royle L, Gardner RA, Fernandes DL, Wuhrer M. 2012. Suppression of peeling during the release of O-glycans by hydrazinolysis. *Analytical Biochemistry*, 423:119-128.

Kumar A, Galaev I, Mattiasson B, Cabral JMS. 2007. Cell Partitioning in Aqueous Two-Phase Polymer Systems. Cell Separation: Springer Berlin Heidelberg. p. 151-171.

Kurimoto A, Kitazume S, Kizuka Y, Nakajima K, Oka R, Fujinawa R, Korekane H, Yamaguchi Y, Wada Y, Taniguchi N. 2014. The absence of core fucose up-regulates GnT-III and Wnt target genes: a possible mechanism for an adaptive response in terms of glycan function. *J Biol Chem*, 289:11704-11714.

Kusminski CM, Scherer PE. 2012. Mitochondrial dysfunction in white adipose tissue. *Trends Endocrinol Metab*, 23:435-443.

Lanctot PM, Gage FH, Varki AP. 2007. The glycans of stem cells. *Curr Opin Chem Biol*, 11:373-380.

Lee A, Chick JM, Kolarich D, Haynes PA, Robertson GR, Tsoli M, Jankova L, Clarke SJ, Packer NH, Baker MS. 2011. Liver membrane proteome glycosylation changes in mice bearing an extra-hepatic tumor. *Mol Cell Proteomics*, 10:M900538MCP900200.

Lee A, Kolarich D, Haynes PA, Jensen PH, Baker MS, Packer NH. 2009. Rat Liver Membrane Glycoproteome: Enrichment by Phase Partitioning and Glycoprotein Capture. *Journal of Proteome Research*, 8:770-781.

Lee HK, Lee BH, Park SA, Kim CW. 2006. The proteomic analysis of an adipocyte differentiated from human mesenchymal stem cells using two-dimensional gel electrophoresis. *Proteomics*, 6:1223-1229.

Leymarie N, Griffin PJ, Jonscher K, Kolarich D, Orlando R, McComb M, Zaia J, Aguilan J, Alley WR, Altmann F*, et al.* 2013. Interlaboratory Study on Differential Analysis of Protein Glycosylation by Mass Spectrometry: The ABRF Glycoprotein Research Multi-Institutional Study 2012. *Molecular & Cellular Proteomics : MCP*, 12:2935-2951.

Li HX, Xiao L, Wang C, Gao JL, Zhai YG. 2010. Review: Epigenetic regulation of adipocyte differentiation and adipogenesis. *J Zhejiang Univ Sci B*, 11:784-791.

Lin CH, Chik JH, Packer NH, Molloy MP. 2015. Multidimensional fractionation is a requirement for quantitation of Golgi-resident glycosylation enzymes from cultured human cells. *J Proteome Res*, 14:747-755.

Lin CS, Xin ZC, Deng CH, Ning H, Lin G, Lue TF. 2010. Defining adipose tissue-derived stem cells in tissue and in culture. *Histol Histopathol*, 25:807-815.

Lindroos B, Suuronen R, Miettinen S. 2011. The potential of adipose stem cells in regenerative medicine. *Stem Cell Rev*, 7:269-291.

Luque-Garcia JL, Neubert TA. 2009. On-membrane tryptic digestion of proteins for mass spectrometry analysis. *Methods Mol Biol*, 536:331-341.

Lutolf MP, Gilbert PM, Blau HM. 2009. Designing materials to direct stem-cell fate. *Nature*, 462:433-441.

Malik R, Dulla K, Nigg EA, Korner R. 2010. From proteome lists to biological impact--tools and strategies for the analysis of large MS data sets. *Proteomics*, 10:1270-1283.

Maniatis S, Zhou H, Reinhold V. 2010. Rapid De-O-glycosylation Concomitant with Peptide Labeling Using Microwave Radiation and an Alkyl Amine Base. *Analytical Chemistry*, 82:2421-2425.

Mariño K, Bones J, Kattla JJ, Rudd PM. 2010. A systematic approach to protein glycosylation analysis: a path through the maze. *Nat Chem Biol*, 6:713-723.

Markarian CF, Frey GZ, Silveira MD, Chem EM, Milani AR, Ely PB, Horn AP, Nardi NB, Camassola M. 2014. Isolation of adipose-derived stem cells: a comparison among different methods. *Biotechnol Lett*, 36:693-702.

Martel J, Wu CY, Young JD. 2010. Critical evaluation of gamma-irradiated serum used as feeder in the culture and demonstration of putative nanobacteria and calcifying nanoparticles. *PLoS One*, 5:e10343.

Mathieu PS, Loboa EG. 2012. Cytoskeletal and focal adhesion influences on mesenchymal stem cell shape, mechanical properties, and differentiation down osteogenic, adipogenic, and chondrogenic pathways. *Tissue Eng Part B Rev*, 18:436-444.

May P, Bock HH, Nimpf J, Herz J. 2003. Differential glycosylation regulates processing of lipoprotein receptors by gamma-secretase. *J Biol Chem*, 278:37386-37392.

Mechref Y. 2011. Analysis of glycans derived from glycoconjugates by capillary electrophoresis-mass spectrometry. *Electrophoresis*, 32:3467-3481.

Mechref Y. 2012. Use of CID/ETD mass spectrometry to analyze glycopeptides. *Curr Protoc Protein Sci*, Chapter 12:Unit 12 11 11-11.

Mehlem A, Hagberg CE, Muhl L, Eriksson U, Falkevall A. 2013. Imaging of neutral lipids by oil red O for analyzing the metabolic status in health and disease. *Nat Protoc*, 8:1149-1154.

Melo-Braga MN, Schulz M, Liu Q, Swistowski A, Palmisano G, Engholm-Keller K, Jakobsen L, Zeng X, Larsen MR. 2014. Comprehensive quantitative comparison of the membrane proteome, phosphoproteome, and sialiome of human embryonic and neural stem cells. *Mol Cell Proteomics*, 13:311-328.

Minteer D, Marra KG, Rubin JP. 2013. Adipose-derived mesenchymal stem cells: biology and potential applications. *Adv Biochem Eng Biotechnol*, 129:59-71.

Mitchell JB, McIntosh K, Zvonic S, Garrett S, Floyd ZE, Kloster A, Di Halvorsen Y, Storms RW, Goh B, Kilroy G*, et al.* 2006. Immunophenotype of human adipose-derived cells: temporal changes in stromal-associated and stem cell-associated markers. *Stem Cells*, 24:376-385.

Mizuno H, Itoi Y, Kawahara S, Ogawa R, Akaishi S, Hyakusoku H. 2008. In vivo adipose tissue regeneration by adipose-derived stromal cells isolated from GFP transgenic mice. *Cells Tissues Organs*, 187:177-185.

Moh ESX, Thaysen-Andersen M, Packer NH. 2015. Relative vs absolute quantitation in disease glycomics. *PROTEOMICS Clinical Applications*:n/a-n/a.

Moloney DJ, Panin VM, Johnston SH, Chen J, Shao L, Wilson R, Wang Y, Stanley P, Irvine KD, Haltiwanger RS*, et al.* 2000. Fringe is a glycosyltransferase that modifies Notch. *Nature*, 406:369-375.

Moloney DJ, Shair LH, Lu FM, Xia J, Locke R, Matta KL, Haltiwanger RS. 2000. Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *J Biol Chem*, 275:9604-9611.

Morelle W, Michalski JC. 2007. Analysis of protein glycosylation by mass spectrometry. *Nat Protoc*, 2:1585-1602.

Moremen KW, Tiemeyer M, Nairn AV. 2012. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol*, 13:448-462.

Moreno-Navarrete JM, Fernández-Real JM. 2012. Adipocyte Differentiation. In: Symonds ME editor. Adipose Tissue Biology: Springer New York. p. 17-38.

Mun JY, Lee KJ, Seo H, Sung MS, Cho YS, Lee SG, Kwon O, Oh DB. 2013. Efficient adhesion-based plasma membrane isolation for cell surface N-glycan analysis. *Anal Chem*, 85:7462-7470.

Munnich S, Taft MH, Manstein DJ. 2014. Crystal structure of human myosin 1c--the motor in GLUT4 exocytosis: implications for Ca2+ regulation and 14-3-3 binding. *J Mol Biol*, 426:2070-2081.

Muramatsu T, Muramatsu H. 2004. Carbohydrate antigens expressed on stem cells and early embryonic cells. *Glycoconj J*, 21:41-45.

Murray I, West C, Hardy W, James A, Park T, Nguyen A, Tawonsawatruk T, Lazzari L, Soo C, Peault B. 2014. Natural history of mesenchymal stem cells, from vessel walls to culture vessels. *Cellular and Molecular Life Sciences*, 71:1353-1374.

Musri MM, Parrizas M. 2012. Epigenetic regulation of adipogenesis. *Curr Opin Clin Nutr Metab Care*, 15:342-349.

Nairn AV, Aoki K, dela Rosa M, Porterfield M, Lim JM, Kulik M, Pierce JM, Wells L, Dalton S, Tiemeyer M*, et al.* 2012. Regulation of glycan structures in murine embryonic stem cells: combined transcript profiling of glycan-related genes and glycan structural analysis. *J Biol Chem*, 287:37835-37856.

Nairn AV, York WS, Harris K, Hall EM, Pierce JM, Moremen KW. 2008. Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *J Biol Chem*, 283:17298-17313.

Nakano M, Saldanha R, Gobel A, Kavallaris M, Packer NH. 2011. Identification of glycan structure alterations on cell membrane proteins in desoxyepothilone B resistant leukemia cells. *Mol Cell Proteomics*, 10:M111 009001.

Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, Assadourian G, Lee A, van Sluyter SC, Haynes PA. 2011. Less label, more free: Approaches in label-free quantitative mass spectrometry. *Proteomics*, 11:535-553.

Neilson KA, George IS, Emery SJ, Muralidharan S, Mirzaei M, Haynes PA. 2014. Analysis of rice proteins using SDS-PAGE shotgun proteomics. *Methods Mol Biol*, 1072:289-302.

Neilson KA, Keighley T, Pascovici D, Cooke B, Haynes PA. 2013. Label-free quantitative shotgun proteomics using normalized spectral abundance factors. *Methods Mol Biol*, 1002:205-222.

Neilson KA, Mariani M, Haynes PA. 2011. Quantitative proteomic analysis of cold-responsive proteins in rice. *Proteomics*, 11:1696-1706.

Nguyen-Khuong T, Everest-Dass AV, Kautto L, Zhao Z, Willcox MDP, Packer NH. 2015. Glycomic characterization of basal tears and changes with diabetes and diabetic retinopathy. *Glycobiology*, 25:269-283.

Nguyen MT, Csermely P, Soti C. 2013. Hsp90 chaperones PPARgamma and regulates differentiation and survival of 3T3-L1 adipocytes. *Cell Death Differ*, 20:1654-1663.

Nielsen R, Mandrup S. 2014. Genome-wide profiling of transcription factor binding and epigenetic marks in adipocytes by ChIP-seq. *Methods Enzymol*, 537:261-279.

Nims RW, Gauvin G, Plavsic M. 2011. Gamma irradiation of animal sera for inactivation of viruses and mollicutes--a review. *Biologicals*, 39:370-377.

Nishimura S, Manabe I, Nagasaki M, Seo K, Yamashita H, Hosoya Y, Ohsugi M, Tobe K, Kadowaki T, Nagai R*, et al.* 2008. In vivo imaging in mice reveals local cell dynamics and inflammation in obese adipose tissue. *J Clin Invest*, 118:710-721.

Noer A, Sorensen AL, Boquest AC, Collas P. 2006. Stable CpG hypomethylation of adipogenic promoters in freshly isolated, cultured, and differentiated mesenchymal stem cells from adipose tissue. *Mol Biol Cell*, 17:3543-3556.

Olson AL, Knight JB, Pessin JE. 1997. Syntaxin 4, VAMP2, and/or VAMP3/cellubrevin are functional target membrane and vesicle SNAP receptors for insulin-stimulated GLUT4 translocation in adipocytes. *Mol Cell Biol*, 17:2425-2435.

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics*, 1:376-386.

Pabst M, Altmann F. 2008. Influence of Electrosorption, Solvent, Temperature, and Ion Polarity on the Performance of LC-ESI-MS Using Graphitic Carbon for Acidic Oligosaccharides. *Analytical Chemistry*, 80:7534-7542.

Pabst M, Bondili JS, Stadlmann J, Mach L, Altmann F. 2007. Mass + Retention Time = Structure: A Strategy for the Analysis of N-Glycans by Carbon LC-ESI-MS and Its Application to Fibrin N-Glycans. *Analytical Chemistry*, 79:5051-5057.

Palmisano G, Melo-Braga MN, Engholm-Keller K, Parker BL, Larsen MR. 2012. Chemical deamidation: a common pitfall in large-scale N-linked glycoproteomic mass spectrometry-based analyses. *J Proteome Res*, 11:1949-1957.

Pan S, Chen R, Aebersold R, Brentnall TA. 2011. Mass Spectrometry Based Glycoproteomics - From a Proteomics Perspective. *Molecular & Cellular Proteomics : MCP*, 10:R110.003251.

Park IS, Han M, Rhie JW, Kim SH, Jung Y, Kim IH. 2009. The correlation between human adipose-derived stem cells differentiation and cell adhesion mechanism. *Biomaterials*, 30:6835-6843.

Parker BL, Thaysen-Andersen M, Solis N, Scott NE, Larsen MR, Graham ME, Packer NH, Cordwell SJ. 2013. Site-Specific Glycan-Peptide Analysis for Determination of N-Glycoproteome Heterogeneity. *Journal of Proteome Research*, 12:5791-5800.

Peng J, Elias JE, Thoreen CC, Licklider LJ, Gygi SP. 2003. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res*, 2:43-50.

Pettersson P, Cigolini M, Sjostrom L, Smith U, Bjorntorp P. 1984. Cells in human adipose tissue developing into adipocytes. *Acta Med Scand*, 215:447-451.

Pevzner PA, Mulyukov Z, Dancik V, Tang CL. 2001. Efficiency of database search for identification of mutated and modified proteins via mass spectrometry. *Genome Res*, 11:290-299.

Poulos SP, Dodson MV, Hausman GJ. 2010. Cell line models for differentiation: preadipocytes and adipocytes. *Exp Biol Med (Maywood)*, 235:1185-1193.

Prieto C, De Las Rivas J. 2006. APID: Agile Protein Interaction DataAnalyzer. *Nucleic Acids Research*, 34:W298-W302.

Prior MJ, Larance M, Lawrence RT, Soul J, Humphrey S, Burchfield J, Kistler C, Davey JR, La-Borde PJ, Buckley M*, et al.* 2011. Quantitative Proteomic Analysis of the Adipocyte Plasma Membrane. *Journal of Proteome Research*, 10:4970-4982.

Ragni E, Vigano M, Rebulla P, Giordano R, Lazzari L. 2012. What is beyond a qRT-PCR study on mesenchymal stem cell differentiation properties: how to choose the most reliable housekeeping genes. *J Cell Mol Med*, 17:168-180.

Rakus JF, Mahal LK. 2011. New technologies for glycomic analysis: toward a systematic understanding of the glycome. *Annu Rev Anal Chem (Palo Alto Calif)*, 4:367-392.

Reddy JK, Rao MS. 2006. Lipid metabolism and liver inflammation. II. Fatty liver disease and fatty acid oxidation. *Am J Physiol Gastrointest Liver Physiol*, 290:G852-858.

Ringden O, Okas M, Uhlin M, Uzunel M, Remberger M, Mattsson J. 2008. Unrelated cord blood and mismatched unrelated volunteer donor transplants, two alternatives in patients who lack an HLA-identical donor. *Bone Marrow Transplant*, 42:643-648.

Robinson LN, Artpradit C, Raman R, Shriver ZH, Ruchirawat M, Sasisekharan R. 2012. Harnessing glycomics technologies: integrating structure with function for glycan characterization. *Electrophoresis*, 33:797-814.

Rocha V, Labopin M, Sanz G, Arcese W, Schwerdtfeger R, Bosi A, Jacobsen N, Ruutu T, de Lima M, Finke Jr*, et al.* 2004. Transplants of Umbilical-Cord Blood or Bone Marrow from Unrelated Donors in Adults with Acute Leukemia. *New England Journal of Medicine*, 351:2276-2285.

Rodriguez JP, Montecinos L, Rios S, Reyes P, Martinez J. 2000. Mesenchymal stem cells from osteoporotic patients produce a type I collagen-deficient extracellular matrix favoring adipogenic differentiation. *J Cell Biochem*, 79:557-565.

Rosen ED, MacDougald OA. 2006. Adipocyte differentiation from the inside out. *Nat Rev Mol Cell Biol*, 7:885-896.

Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S*, et al.* 2004. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics*, 3:1154-1169.

Royle L, Radcliffe CM, Dwek RA, Rudd PM. 2006. Detailed structural analysis of N-glycans released from glycoproteins in SDS-PAGE gel bands using HPLC combined with exoglycosidase array digestions. *Methods Mol Biol*, 347:125-143.

Ruhaak LR, Miyamoto S, Kelly K, Lebrilla CB. 2012. N-glycan profiling of Dried Blood Spots. *Analytical Chemistry*, 84:396-402.

Ruhaak LR, Zauner G, Huhn C, Bruggink C, Deelder AM, Wuhrer M. 2010. Glycan labeling strategies and their use in identification and quantification. *Analytical and Bioanalytical Chemistry*, 397:3457-3481.

Saito S, Onuma Y, Ito Y, Tateno H, Toyoda M, Hidenori A, Nishino K, Chikazawa E, Fukawatase Y, Miyagawa Y*, et al.* 2011. Possible linkages between the inner and outer cellular states of human induced pluripotent stem cells. *BMC Syst Biol*, 5 Suppl 1:S17.

Sato Y, Takahashi M, Shibukawa Y, Jain SK, Hamaoka R, Miyagawa J, Yaginuma Y, Honke K, Ishikawa M, Taniguchi N. 2001. Overexpression of N-acetylglucosaminyltransferase III enhances the epidermal growth factor-induced phosphorylation of ERK in HeLaS3 cells by up-regulation of the internalization rate of the receptors. *The Journal of biological chemistry*, 276:11956.

Satomaa T, Heiskanen A, Mikkola M, Olsson C, Blomqvist M, Tiittanen M, Jaatinen T, Aitio O, Olonen A, Helin J*, et al.* 2009. The N-glycome of human embryonic stem cells. *BMC Cell Biol*, 10:42.

Schindler J, Nothwang HG. 2006. Aqueous polymer two-phase systems: Effective tools for plasma membrane proteomics. *Proteomics*, 6:5409-5417.

Schipper BM, Marra KG, Zhang W, Donnenberg AD, Rubin JP. 2008. Regional anatomic and age effects on cell function of human adipose-derived stem cells. *Ann Plast Surg*, 60:538-544.

Schmittgen TD, Livak KJ. 2008. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc*, 3:1101-1108.

Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*, 9:671-675.

Schofield R. 1978. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells*, 4:7-25.

Schwanhausser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M. 2011. Global quantification of mammalian gene expression control. *Nature*, 473:337-342.

Scott MA, Nguyen VT, Levi B, James AW. 2011. Current methods of adipogenic differentiation of mesenchymal stem cells. *Stem Cells Dev*, 20:1793-1804.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13:2498-2504.

Shen W, Chen J. 2008. Application of imaging and other noninvasive techniques in determining adipose tissue mass. *Methods Mol Biol*, 456:39-54.

Sillat T, Saat R, Pollanen R, Hukkanen M, Takagi M, Konttinen YT. 2012. Basement membrane collagen type IV expression by human mesenchymal stem cells during adipogenic differentiation. *J Cell Mol Med*, 16:1485-1495.

Singer SJ, Nicolson GL. 1972. The Fluid Mosaic Model of the Structure of Cell Membranes. *Science*, 175:720-731.

Slatter MA, Bhattacharya A, Flood TJ, Abinun M, Cant AJ, Gennery AR. 2006. Use of two unrelated umbilical cord stem cell units in stem cell transplantation for Wiskott-Aldrich syndrome. *Pediatr Blood Cancer*, 47:332-334.

Slatter MA, Gennery AR. 2006. Umbilical cord stem cell transplantation for primary immunodeficiencies. *Expert Opin Biol Ther*, 6:555-565.

Sprenger RR, Jensen ON. 2010. Proteomics and the dynamic plasma membrane: Quo Vadis? *Proteomics*, 10:3997-4011.

St Croix CM, Shand SH, Watkins SC. 2005. Confocal microscopy: comparisons, applications, and problems. *Biotechniques*, 39:S2-5.

Stavenhagen K, Kolarich D, Wuhrer M. 2015. Clinical Glycomics Employing Graphitized Carbon Liquid Chromatography-Mass Spectrometry. *Chromatographia*, 78:307-320.

Strem BM, Hicok KC, Zhu M, Wulur I, Alfonso Z, Schreiber RE, Fraser JK, Hedrick MH. 2005. Multipotential differentiation of adipose tissue-derived stem cells. *Keio J Med*, 54:132-141.

Struwe WB, Gough R, Gallagher ME, Kenny DT, Carrington SD, Karlsson NG, Rudd PM. 2015. Identification of O-glycan structures from chicken intestinal mucins provides insight into Campylobactor jejuni pathogenicity. *Mol Cell Proteomics*.

Sumer-Bayraktar Z, Kolarich D, Campbell MP, Ali S, Packer NH, Thaysen-Andersen M. 2011. N-Glycans Modulate the Function of Human Corticosteroid-Binding Globulin. *Molecular & Cellular Proteomics : MCP*, 10:M111.009100.

Sumer-Bayraktar Z, Nguyen-Khuong T, Jayo R, Chen DDY, Ali S, Packer NH, Thaysen-Andersen M. 2012. Micro- and macroheterogeneity of N-glycosylation yields size and charge isoforms of human sex hormone binding globulin circulating in serum. *Proteomics*, 12:3315-3327.

Szabo E, Feng T, Dziak E, Opas M. 2009. Cell adhesion and spreading affect adipogenesis from embryonic stem cells: the role of calreticulin. *Stem Cells*, 27:2092-2102.

Szabo Z, Guttman A, Rejtar T, Karger BL. 2010. Improved sample preparation method for glycan analysis of glycoproteins by CE-LIF and CE-MS. *Electrophoresis*, 31:1389-1395.

Takahashi M, Kuroki Y, Ohtsubo K, Taniguchi N. 2009. Core fucose and bisecting GlcNAc, the direct modifiers of the N-glycan core: their functions and target proteins. *Carbohydr Res*, 344:1387-1390.

Taniguchi N, Korekane H. 2011. Branched N-glycans and their implications for cell adhesion, signaling and clinical applications for cancer biomarkers and in therapeutics. *BMB Rep*, 44:772-781.

Tapp H, Hanley EN, Jr., Patt JC, Gruber HE. 2009. Adipose-derived stem cells: characterization and current application in orthopaedic tissue repair. *Exp Biol Med (Maywood)*, 234:1-9.

Tateno H, Toyota M, Saito S, Onuma Y, Ito Y, Hiemori K, Fukumura M, Matsushima A, Nakanishi M, Ohnuma K*, et al.* 2011. Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J Biol Chem*, 286:20345-20353.

Thanawiroon C, Rice KG, Toida T, Linhardt RJ. 2004. Liquid chromatography/mass spectrometry sequencing approach for highly sulfated heparin-derived oligosaccharides. *The Journal of biological chemistry*, 279:2608-2615.

Thaysen-Andersen M, Packer NH. 2014. Advances in LC-MS/MS-based glycoproteomics: Getting closer to system-wide site-specific mapping of the N- and O-glycoproteome. *Biochim Biophys Acta*, 1844:1437-1452.

Thaysen-Andersen M, Venkatakrishnan V, Loke I, Laurini C, Diestel S, Parker BL, Packer NH. 2015. Human neutrophils secrete bioactive paucimannosidic proteins from azurophilic granules into pathogen-infected sputum. *J Biol Chem*, 290:8789-8802.

Thomsson KA, Karlsson NG, Hansson GC. 1999. Liquid chromatography-electrospray mass spectrometry as a tool for the analysis of sulfated oligosaccharides from mucin glycoproteins. *Journal of Chromatography A*, 854:131-139.

Toyoda M, Yamazaki-Inoue M, Itakura Y, Kuno A, Ogawa T, Yamada M, Akutsu H, Takahashi Y, Kanzaki S, Narimatsu H*, et al.* 2011. Lectin microarray analysis of pluripotent and multipotent stem cells. *Genes Cells*, 16:1-11.

Tran DT, Ten Hagen KG. 2013. Mucin-type O-glycosylation during development. *J Biol Chem*, 288:6921-6929.

Trayhurn P, Wood IS. 2004. Adipokines: inflammation and the pleiotropic role of white adipose tissue. *Br J Nutr*, 92:347-355.

Urs S, Smith C, Campbell B, Saxton AM, Taylor J, Zhang B, Snoddy J, Jones Voy B, Moustaid-Moussa N. 2004. Gene expression profiling in human preadipocytes and adipocytes by microarray analysis. *J Nutr*, 134:762-770.

van Hoof D, Krijgsveld J, Mummery C. 2012. Proteomic analysis of stem cell differentiation and early development. *Cold Spring Harb Perspect Biol*, 4.

van Kooyk Y, Kalay H, Garcia-Vallejo JJ. 2013. Analytical Tools for the Study of Cellular Glycosylation in the Immune System. *Front Immunol*, 4:451.

Venkatakrishnan V, Thaysen-Andersen M, Chen SCA, Nevalainen H, Packer NH. 2015. Cystic fibrosis and bacterial colonization define the sputum N-glycosylation phenotype. *Glycobiology*, 25:88-100.

Voog J, Jones DL. 2010. Stem cells and the niche: a dynamic duo. *Cell Stem Cell*, 6:103-115.

Walker J. 1996. The Bicinchoninic Acid (BCA) Assay for Protein Quantitation. The Protein Protocols Handbook: Humana Press. p. 11-14.

Wall ME, Bernacki SH, Loboa EG. 2007. Effects of serial passaging on the adipogenic and osteogenic differentiation potential of adipose-derived human mesenchymal stem cells. *Tissue Eng*, 13:1291-1298.

Wang QA, Scherer PE, Gupta RK. 2014. Improved methodologies for the study of adipose biology: insights gained and opportunities ahead. *J Lipid Res*, 55:605-624.

Wang X, Gu J, Ihara H, Miyoshi E, Honke K, Taniguchi N. 2006. Core fucosylation regulates epidermal growth factor receptor-mediated intracellular signaling. *J Biol Chem*, 281:2572-2577.

Wang X, Wang H, Figueroa BE, Zhang WH, Huo C, Guan Y, Zhang Y, Bruey JM, Reed JC, Friedlander RM. 2005. Dysregulation of receptor interacting protein-2 and caspase recruitment domain only protein mediates aberrant caspase-1 activation in Huntington's disease. *J Neurosci*, 25:11645-11654.

Wang YC, Nakagawa M, Garitaonandia I, Slavin I, Altun G, Lacharite RM, Nazor KL, Tran HT, Lynch CL, Leonardo TR*, et al.* 2011. Specific lectin biomarkers for isolation of human pluripotent stem cells identified through array-based glycomic analysis. *Cell Res*, 21:1551-1563.

Wearne KA, Winter HC, Goldstein IJ. 2008. Temporal changes in the carbohydrates expressed on BG01 human embryonic stem cells during differentiation as embryoid bodies. *Glycoconj J*, 25:121-136.

Wearne KA, Winter HC, O'Shea K, Goldstein IJ. 2006. Use of lectins for probing differentiated human embryonic stem cells for carbohydrates. *Glycobiology*, 16:981-990.

Weerapana E, Imperiali B. 2006. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology*, 16:91R-101R.

Wongtrakul-Kish K, Kolarich D, Pascovici D, Joss J, Deane E, Packer N. 2013. Characterization of N- and O-linked glycosylation changes in milk of the tammar wallaby (Macropus eugenii) over lactation. *Glycoconjugate Journal*, 30:523-536.

Wuhrer M, Deelder AM, Hokke CH. 2005. Protein glycosylation analysis by liquid chromatography-mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci*, 825:124-133.

Wuhrer M, Stam JC, van de Geijn FE, Koeleman CA, Verrips CT, Dolhain RJ, Hokke CH, Deelder AM. 2007. Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum. *Proteomics*, 7:4070-4081.

Xie X, Yi Z, Bowen B, Wolf C, Flynn CR, Sinha S, Mandarino LJ, Meyer C. 2010. Characterization of the Human Adipocyte Proteome and Reproducibility of Protein Abundance by One-Dimensional Gel Electrophoresis and HPLC-ESI-MS/MS. *J Proteome Res*, 9:4521-4534.

Yates JR, Ruse CI, Nakorchevsky A. 2009. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu Rev Biomed Eng*, 11:49-79.

Ye F, Zhang H, Yang YX, Hu HD, Sze SK, Meng W, Qian J, Ren H, Yang BL, Luo MY*, et al.* 2011. Comparative proteome analysis of 3T3-L1 adipocyte differentiation using iTRAQ-coupled 2D LC-MS/MS. *J Cell Biochem*, 112:3002-3014.

Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, 13:134.

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L. 2006. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res*, 34:W293-297.

Yu J, Thomson JA. 2008. Pluripotent stem cell lines. *Genes Dev*, 22:1987-1997.

Zauner G, Deelder AM, Wuhrer M. 2011. Recent advances in hydrophilic interaction liquid chromatography (HILIC) for structural glycomics. *Electrophoresis*, 32:3456-3466.

Zavan B, Vindigni V, Gardin C, D'Avella D, Della Puppa A, Abatangelo G, Cortivo R. 2010. Neural potential of adipose stem cells. *Discov Med*, 10:37-43.

Zeeberg BR, Feng W, Wang G, Wang MD, Fojo AT, Sunshine M, Narasimhan S, Kane DW, Reinhold WC, Lababidi S*, et al.* 2003. GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol*, 4:R28.

Zhang Y, Marsboom G, Toth PT, Rehman J. 2013. Mitochondrial respiration regulates adipogenic differentiation of human mesenchymal stem cells. *PLoS One*, 8:e77077.

Zhao P, Yang L, Lopez JA, Fan J, Burchfield JG, Bai L, Hong W, Xu T, James DE. 2009. Variations in the requirement for v-SNAREs in GLUT4 trafficking in adipocytes. *J Cell Sci*, 122:3472-3480.

Zhao Y, Zhang W, Kho Y. 2004. Proteomic analysis of integral plasma membrane proteins. *Anal Chem*, 76:1817-1823.

Zimmerlin L, Donnenberg AD, Rubin JP, Basse P, Landreneau RJ, Donnenberg VS. 2011. Regenerative therapy and cancer: in vitro and in vivo studies of the interaction between adipose-derived stem cells and breast cancer cells from clinical isolates. *Tissue Eng Part A*, 17:93-106.

Zuk P. 2013. Adipose-Derived Stem Cells in Tissue Regeneration: A Review. *ISRN Stem Cells*, 2013:35.

Zuk PA. 2010. The adipose-derived stem cell: looking back and looking ahead. *Mol Biol Cell*, 21:1783-1787.

Zuk PA, Zhu M, Mizuno H, Huang J, Futrell JW, Katz AJ, Benhaim P, Lorenz HP, Hedrick MH. 2001. Multilineage cells from human adipose tissue: implications for cell-based therapies. *Tissue Eng*, 7:211-228.

Zybailov B, Mosley AL, Sardiu ME, Coleman MK, Florens L, Washburn MP. 2006. Statistical analysis of membrane proteome expression changes in Saccharomyces cerevisiae. *J Proteome Res*, 5:2339-2347.

Chapter 3 Supplementary Data

Table S3.1 Comprehensive analysis of MS/MS fragmentation spectra and knowledge of known PGC-LC elution order enabled the characterisation of 130 *N*-linked glycan structures in total from ADSCs, dADSCs and native adipocytes. The details of each structure are shown below and *N*-acetylneuraminic acid linkages are added where known. The final column contains extra information for some structures that helped in structural elucidation

Glycan #	Detected m/z	Theoretical Mass (Da)	Retention time (min)	Composition	Mean % ± SD			Proposed Structure	Comments
					ADSCs	dADSCs	Adipocytes		
	[57 2]	57 2	35	(HexNAc)2 (uc)	0 05 ± 0 02	-	0 07 ± 0 02		
2	[698 3]^{2} & [397 6]	397 5	40 4	(Hex)3 + (Ma)3(G cNAc)2	4 48 ± 83	5 02 ± 37	2 54 ± 0 70		
3	[75 3]^{3}	463 6	46 3	(HexNAc)2 (uc) + (Ma)3(G cNAc)2	0 32 ± 0 8	9 ± 0 9	0 9 ± 0		
4	[739 3]^{2}	479 5	30 6	(Hex) (HexNAc)2 + (Ma)3(G cNAc)2	-	-	0 06 ± 0 04		
5	[739 3]^{3}	479 5	36 5	(Hex) (HexNAc)2 + (Ma)3(G cNAc)2	0 0 ± 0 0	-	0 8 ± 0 03		
6	[749 3]	749 3	37 3	(Hex)2 (HexNAc)2	0 0 ± 0 3	0 26 ± 0 06	0 04 ± 0 02		
7	[759 8]^{3}	520 6	3 9	(HexNAc)3 + (Ma)3(G cNAc)2	-	-	0 39 ± 0 6		
8	[779 3]^{3}	559 5	38 9	(Hex)4 + (Ma)3(G cNAc)2	2 80 ± 0 55	2 97 ± 0 57	45 ± 0 34		
9	[779 3]^{3}	559 5	40 5	(Hex)4 + (Ma)3(G cNAc)2	03 ± 0 5	0 87 ± 0 29	0 22 ± 0 09		
0	[783 3]^{3}	567 6	5 2	(Hex) (HexNAc) (NeuAc) + (Ma)3(G cNAc)2	0 0 ± 0 09	-	0 42 ± 0 2	α2-6	NOT digested by NAN
	[79 8]^{3}	584 6	45 6	(Hex)2 (HexNAc) (uc) + (Ma)3(G cNAc)2	0 57 ± 0 2	0 38 ± 0 23	0 08 ± 0 07		

2	[799 7]^{2}	600 6	43 5	(Hex)3 (HexNAc) + (Ma)3(G cNAc)2	2 2 ± 0 29	97 ± 0 86	4 ± 0 9	
3	[8 2 4]^{2}	625 6	37 8	(Hex) (HexNAc)2 (uc) + (Ma)3(G cNAc)2	-	-	0 05 ± 0 0	
4	[8 2 4]^{2}	625 6	48 2	(Hex) (HexNAc)2 (uc) + (Ma)3(G cNAc)2	0 22 ± 0 03	0 9 ± 0 04	0 54 ± 0 27	
5	[8 2 4]^{2}	625 6	49	(Hex) (HexNAc)2 (uc) + (Ma)3(G cNAc)2	0 2 ± 0 05	0 6 ± 0 6	0 2 ± 0	
6	[820 3]^{2}	787 7	30 6	(Hex)2 (HexNAc)2 + (Ma)3(G cNAc)2	-	-	0 02 ± 0 0	2x
7	[820 3]^{2}	787 7	33 0	(Hex)2 (HexNAc)2 + (Ma)3(G cNAc)2	-	-	0 22 ± 0 0	
8	[820 3]^{2}	787 7	36 4	(Hex)2 (HexNAc)2 + (Ma)3(G cNAc)2	-	-	0 58 ± 0 28	
9	[820 3]^{2}	787 7	45 2	(Hex)2 (HexNAc)2 + (Ma)3(G cNAc)2	3 6 ± 0 6	2 04 ± 0 62	0 ± 0 49	
20	[832 8]^{2}	666 6	38 4	(HexNAc)3 (uc) + (Ma)3(G cNAc)2	-	0 09 ± 0 06	09 ± 0 32	
2	[832 8]^{2}	666 6	45	(HexNAc)3 (uc) + (Ma)3(G cNAc)2	0 6 ± 0 2	0 60 ± 0 24	-	
22	[832 8]^{2}	666 6	49 9	(HexNAc)3 (uc) + (Ma)3(G cNAc)2	0 05 ± 0 004	0 02 ± 0 03	-	
23	[840 8]^{2}	828 7	34 6	(Hex) (HexNAc)3 + (Ma)3(G cNAc)2	0 02 ± 0 04	0 08 ± 0 06	63 ± 0 29	

24	[856 4]^{2}	7 3 6	55 9	(Hex) (HexNAc) (uc) (NeuAc) + (Ma)3(G cNAc)2	0 53 ± 0 24	0 3 ± 0 2	0 3 ± 0 06	α2-6	L ke y kage acco d g o RT pa e
25	[856 4]^{2}	7 3 6	62 3	(Hex) (HexNAc) (uc) (NeuAc) + (Ma)3(G cNAc)2	0 34 ± 0 6	0 5 ± 0 4	0 09 ± 0 03	α2-3	L ke y kage acco d g o RT pa e
26	[860 4]^{2}	72 6	39	(Hex)5 + (Ma)3(G cNAc)2	8 74 ± 2 32	9 63 ± 2 90	4 49 ± 0 8		
27	[864 3]^{2}	875 7	4 7	(Hex)2 (HexNAc) (NeuAc) + (Ma)3(G cNAc)2	0 34 ± 0 36	0 2 ± 0 8	-	α2-6	L ke y kage acco d g o RT pa e
28	[864 3]^{2}	875 7	50 2	(Hex)2 (HexNAc) (NeuAc) + (Ma)3(G cNAc)2	0 93 ± 0 65	0 67 ± 0 25	0 57 ± 0 22	α2-3	L ke y kage acco d g o RT pa e ; D ges ed by NAN
29	[872 8]^{2}	746 6	4 4	(Hex)3 (HexNAc) (uc) + (Ma)3(G cNAc)2	0 7 ± 0 5	0 09 ± 0 6	-		
30	[872 8]^{2}	746 6	49	(Hex)3 (HexNAc) (uc) + (Ma)3(G cNAc)2	0 63 ± 0 36	0 25 ± 0 28	0 05 ± 0 08		
3	[893 4]^{2}	787 7	50 2	(Hex)2 (HexNAc)2 (uc) + (Ma)3(G cNAc)2	06 ± 0 9	37 ± 5 32	7 62 ± 2 22		
32	[895 4]	895 3	45 6	(Hex)2 (HexNAc)2 (uc)	0 56 ± 0 36	27 ± 0 63	0 24 ± 0 03		

33	[9 5]	9 3	43 9	(Hex)3 (HexNAc)2	0 3 ± 0 5	0 44 ± 0 20	0 9 ± 0 05	
34	[9 3 8]^{2}	828 7	37 7	(Hex) (HexNAc)3 (uc) + (Ma)3(G cNAc)2	-	-	0 05 ± 0 04	
35	[9 3 8]^{2}	828 7	40	(Hex) (HexNAc)3 (uc) + (Ma)3(G cNAc)2	-	0 3 ± 0 27	6 5 ± 03	
36	[92 5]^{2}	844 7	36 7	(Hex)2 (HexNAc)3 + (Ma)3(G cNAc)2	0 03 ± 0 0	0 5 ± 0 03	0 95 ± 0 27	
37	[937 4]^{2}	875 7	46 6	(Hex)2 (HexNAc) (uc) (NeuAc) + (Ma)3(G cNAc)2	0 ± 0 09	0 ± 0 2	-	L ke y kage acco d g o RT pa e
38	[937 4]^{2}	875 7	54 5	(Hex)2 (HexNAc) (uc) (NeuAc) + (Ma)3(G cNAc)2	34 ± 0 78	0 62 ± 0 26	0 23 ± 0 5	D ges ed by NAN
39	[94 4]^{2}	883 7	39 8	(Hex)6 + (Ma)3(G cNAc)2	5 2 ± 2 4	5 39 ± 3 7	5 24 ± 3 07	
40	[945 4]^{2}	89 7	44 9	(Hex)3 (HexNAc) (NeuAc) + (Ma)3(G cNAc)2	0 3 ± 0 08	0 49 ± 0 26	0 02 ± 0 02	L ke y kage acco d g o RT pa e
4	[945 4]^{2}	89 7	52 8	(Hex)3 (HexNAc) (NeuAc) + (Ma)3(G cNAc)2	2 06 ± 0 58	62 ± 0 29	54 ± 0 7	L ke y kage acco d g o RT pa e
42	[965 9]^{2}	932 7	46 6	(Hex)2 (HexNAc)2 (NeuAc) + (Ma)3(G cNAc)2	0 73 ± 0 46	0 69 ± 0 29	0 2 ± 0 5	L ke y kage acco d g o RT pa e
43	[965 9]^{2}	932 7	53 9	(Hex)2 (HexNAc)2 (NeuAc) + (Ma)3(G cNAc)2	38 ± 70	0 24 ± 0 6	0 9 ± 0 02	L ke y kage acco d g o RT pa e
44	[965 9]^{2}	932 7	54 3	(Hex)2 (HexNAc)2 (NeuAc) + (Ma)3(G cNAc)2	3 7 ± 0 86	2 02 ± 0 74	27 ± 0 8	L ke y kage acco d g

									o RT pa e
45	[966 3]^{2}	933 7	48 4	(Hex)2 (HexNAc)2 (uc)2 + (Ma)3(G cNAc)2	0 87 ± 0 62	0 47 ± 0 23	0 04 ± 0 04		
46	[966 3]^{2}	933 7	49 0	(Hex)2 (HexNAc)2 (uc)2 + (Ma)3(G cNAc)2	0 37 ± 0 4	0 28 ± 0 08	0 03 ± 0 04		
47	[974 3]^{2}	949 7	50	(Hex)3 (HexNAc)2 (uc) + (Ma)3(G cNAc)2	-	-	0 07 ± 0 03		
48	[974 3]^{2}	949 7	53	(Hex)3 (HexNAc)2 (uc) + (Ma)3(G cNAc)2	0 9 ± 0 04	0 3 ± 0	0 08 ± 0 06		
49	[974 3]^{2}	949 7	56 0	(Hex)3 (HexNAc)2 (uc) + (Ma)3(G cNAc)2	0 2 ± 0 07	0 04 ± 0 07	-		
50	[994 9]^{2}	990 7	4 2	(Hex)2 (HexNAc)3 (uc) + (Ma)3(G cNAc)2	0 ± 0	04 ± 0 22	5 23 ± 0 0		
5	[002 9]^{2}	2006 7	48 3	(Hex)3 (HexNAc)3 + (Ma)3(G cNAc)2	0 47 ± 0 32	0 42 ± 0 27	0 2 ± 0 06		
52	[002 9]^{2}	2006 7	55 8	(Hex)3 (HexNAc)3 + (Ma)3(G cNAc)2	0 9 ± 0 0	0 08 ± 0 08	0 02 ± 0 002		
53	[008]^{2}	3026	63	(Hex)3 (HexNAc)3 (uc) (NeuAc)3 + (Ma)3(G cNAc)2	0 09 ± 0 6	0 07 ± 0 3	0 0 ± 0 07		
54	[0 8 4]^{2}	2037 7	56 0	(Hex)3 (HexNAc) (uc) (NeuAc) + (Ma)3(G cNAc)2	0 56 ± 0 4	0 08 ± 0 3	0 3 ± 0 002	α2-6	L ke y kage acco d g o RT pa e

55	[0 8 4]^{2}	2037 7	64 5	(Hex)3 (HexNAc) (uc) (NeuAc) + (Ma)3(G cNAc)2	0 06 ± 0 06	-	0 04 ± 0 005	α2-3	L ke y kage acco d g o RT pa e
56	[022 4]^{2}	2045 7	42 0	(Hex)7 + (Ma)3(G cNAc)2	6 ± 0 98	0 56 ± 0 27	0 69 ± 0 25		
57	[038 9]^{2}	2078 8	5	(Hex)2 (HexNAc)2 (uc) (NeuAc) + (Ma)3(G cNAc)2	2 06 ± 0 73	3 53 ± 0 77	0 56 ± 0 05	α2-6	L ke y kage acco d g o RT pa e
58	[038 9]^{2}	2078 8	57 5	(Hex)2 (HexNAc)2 (uc) (NeuAc) + (Ma)3(G cNAc)2	66 ± 0 02	36 ± 0 23	33 ± 0 44	α2-3	L ke y kage acco d g o RT pa e
59	[038 9]^{2}	2078 8	58 2	(Hex)2 (HexNAc)2 (uc) (NeuAc) + (Ma)3(G cNAc)2	9 09 ± 0 77	8 56 ± 2 8	8 52 ± 2 48	α2-3	L ke y kage acco d g o RT pa e
60	[057 4]	057 4	5 4	(Hex)3 (HexNAc)2 (uc)	64 ± 0 25	79 ± 37	0 96 ± 0 29		
6	[059 4]^{2}	2 9 8	46	(Hex) (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	-	-	0 36 ± 0 2	α2-3	D ges ed by NAN
62	[059 4]^{2}	2 9 8	47 4	(Hex) (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	-	-	0 09 ± 0 04	α2-3	D ges ed by NAN
63	[073 5]^{2}	073 4	32 7	(Hex)4 (HexNAc)2	0 02 ± 0 02	-	0 03 ± 0 0		
64	[073 5]^{2}	073 4	36	(Hex)4 (HexNAc)2	-	-	0 02 ± 0 03		
65	[075 9]^{2}	2 52 8	48 9	(Hex)3 (HexNAc)3 (uc) + (Ma)3(G cNAc)2	-	-	0 39 ± 0 2		
66	[075 9]^{2}	2 52 8	52 2	(Hex)3 (HexNAc)3 (uc) + (Ma)3(G cNAc)2	29 ± 0 86	2 22 ± 0 84	2 ± 0 69		

67	[075 9]^{2}	2 52 8	56 5	(Hex)3 (HexNAc)3 (uc) + (Ma)3(G cNAc)2	0 26 ± 0 03	0 54 ± 0 23	0 07 ± 0 09	
68	[075 9]^{2}	2 52 8	58 4	(Hex)3 (HexNAc)3 (uc) + (Ma)3(G cNAc)2	0 22 ± 0 9	0 2 ± 0 8	-	
69	[075 9]^{2}	2 52 8	60 2	(Hex)3 (HexNAc)3 (uc) + (Ma)3(G cNAc)2	0 49 ± 0 06	0 37 ± 0 38	0 24 ± 0 9	
70	[078 9]^{2}	2 58 7	56 9	(Hex)2 (HexNAc)2 (uc) (NeuAc) (Su p) + (Ma)3(G cNAc)2	-	-	0 08 ± 0 02	
7	[078 9]^{2}	2 58 7	60 9	(Hex)2 (HexNAc)2 (uc) (NeuAc) (Su p) + (Ma)3(G cNAc)2	-	-	0 02 ± 0 03	
72	[096 4]^{2}	2396 9	43 2	(Hex)2 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	-	-	0 3 ± 0 03	
73	[096 4]^{2}	2396 9	45	(Hex)2 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	-	-	0 9 ± 0 03	
74	[096 4]^{2}	2396 9	49 5	(Hex)2 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	-	-	0 4 ± 0 03	
75	[5]^{2}	2369 8	47 5	(Hex)2 (HexNAc)2(NeuAc)2 + (Ma)3(G cNAc)2	0 07 ± 0 03	0 57 ± 0 80	0 24 ± 0 8	α2-6 α2-6 L ke y kage acco d g o RT pa e
76	[5]^{2}	2369 8	54 9	(Hex)2 (HexNAc)2(NeuAc)2 + (Ma)3(G cNAc)2	0 40 ± 0 0	05 ± 07	0 04 ± 0 02	α2-3 α2-6 L ke y kage acco d g o RT pa e

77	[5)[+]	2369 8	60 8	(Hex)2 (HexNAc)2(NeuAc)2 + (Ma)3(G cNAc)2	32 ± 0 5	0 59 ± 0 2	0 28 ± 0 06	α2-3 α2-3	Likely kage acco d g o RT ga e
78	[9)[+]	2224 8	49 8	(Hex)2 (HexNAc)2 (uc)2 (NeuAc) + (Ma)3(G cNAc)2	0 4 ± 0 05	0 2 ± 0 4	-		
79	[9)[+]	2224 8	54 7	(Hex)2 (HexNAc)2 (uc)2 (NeuAc) + (Ma)3(G cNAc)2	-	-	0 08 ± 0 07		
80	[9)[+]	2224 8	55 9	(Hex)2 (HexNAc)2 (uc)2 (NeuAc) + (Ma)3(G cNAc)2	0 43 ± 0 8	0 2 ± 0 9	0 03 ± 0 03		
8	[9)[+]	2224 8	57 2	(Hex)2 (HexNAc)2 (uc)2 (NeuAc) + (Ma)3(G cNAc)2	0 53 ± 0 28	0 27 ± 0 08	0 05 ± 0 005		
82	[27 9)[+]	2256 8	6 8	(Hex)4 (HexNAc)2 (NeuAc) + (Ma)3(G cNAc)2	0 9 ± 0 02	0 05 ± 0 08	-		
83	[40 5)[+]	228 9	48 2	(Hex)2 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	-	0 0 ± 0 09	8 ± 0 04	α2-6	NOT d ges ed by NAN
84	[40 5)[+]	228 8	48 9	(Hex)2 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	0 0 ± 0 03	0 4 ± 0 04	2 45 ± 0 70	α2-6	NOT d ges ed by NAN
85	[48 6)[+]	2297 8	48 8	(Hex)3 (HexNAc)3 (NeuAc) + (Ma)3(G cNAc)2	0 06 ± 0	0 09 ± 0 03	-	α2-3	D ges ed by NAN
86	[48 6)[+]	2297 8	54	(Hex)3 (HexNAc)3 (NeuAc) + (Ma)3(G cNAc)2	0 7 ± 0 2	0 24 ± 0	0 03 ± 0 0	α2-3	D ges ed by NAN
87	[77 4)[+]	2355 9	44 5	(Hex)3 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	-	-	0 7 ± 0 03		
88	[77 4)[+]	2355 9	45 6	(Hex)3 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	-	- -	9 ± 0 05		

89	[77 4]^{2}	2355 9	49 0	(Hex)3 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	-	-	0 ± 0 03	
90	[77 4]^{2}	2355 9	50 6	(Hex)3 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	-	-	0 ± 0 0	
9	[84 5]^{2}	2369 8	57 9	(Hex)2 (HexNAc)2 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	0 67 ± 0 7	9 ± 0 43	0 0 ± 0 02	α2-6 α2-6
92	[84 5]^{2}	2369 8	59 3	(Hex)2 (HexNAc)2 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	0 2 ± 0 05	0 03 ± 0 05	0 02 ± 0 02	α2-6 α2-3
93	[84 5]^{2}	2369 8	62 6	(Hex)2 (HexNAc)2 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	3 6 ± 0 97	2 43 ± 0 20	2 25 ± 0 58	α2-3 α2-6
94	[84 5]^{2}	2369 8	65 5	(Hex)2 (HexNAc)2 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	0 28 ± 0 05	0 07 ± 0	0 20 ± 0 02	α2-3 α2-3
95	[22 5]^{2}	2443 9	56 3	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	0 26 ± 0 45	-	0 07 ± 0 02	
96	[22 5]^{2}	2443 9	58 7	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	0 33 ± 0 05	0 50 ± 0 25	0 53 ± 0 06	
97	[22 5]^{2}	2443 9	60	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	0 36 ± 0 0	0 2 ± 0 20	0 6 ± 0 07	
98	[22 5]^{2}	2443 9	6 0	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	-	0 45 ± 0 46	0 03 ± 0 02	

99	[22 5]^{2}	2443 9	6 5	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	0 4 ± 0 6	0 8 ± 0 5	0 22 ± 0 08	
00	[22 5]^{2}	2443 9	62 3	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	0 24 ± 0 3	0 03 ± 0 06	0 2 ± 0 05	
0	[22 5]^{2}	2443 9	62 5	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	0 32 ± 0 32	0 23 ± 0 2	0 02 ± 0 04	
02	[22 5]^{2}	2443 9	63 4	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	0 3 ± 0 25	0 27 ± 0 33	0 7 ± 0 03	
03	[22 5]^{2}	2443 9	64 9	(Hex)3 (HexNAc)3 (uc) (NeuAc) + (Ma)3(G cNAc)2	-	-	0 04 ± 0 02	
04	[235 5]	235 4	46 9	(Hex)2 + (Ma)3(G cNAc)2	2 90 ± 2 20	3 65 ± 00	4 84 ± 07	2x
05	[25 5]^{3}	3756 3	64 7	(Hex)5 (HexNAc)5 (uc) (NeuAc)3 + (Ma)3(G cNAc)2	0 07 ± 0 06	0 05 ± 0 05	0 2 ± 0 04	3x 3x
06	[25 5]^{3}	3756 3	65 2	(Hex)5 (HexNAc)5 (uc) (NeuAc)3 + (Ma)3(G cNAc)2	0 05 ± 0 04	0 002 ± 0 004	0 02 ± 0 03	3x 3x
07	[25 5]^{3}	3756 3	65 7	(Hex)5 (HexNAc)5 (uc) (NeuAc)3 + (Ma)3(G cNAc)2	0 06 ± 0 06	0 03 ± 0 05	0 06 ± 0 03	3x 3x
08	[258 5]^{2}	25 7 9	57 0	(Hex)4 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	0 4 ± 0 09	0 5 ± 0 26	0 08 ± 0 0	2x
09	[258 5]^{2}	25 7 9	59	(Hex)4 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	0 08 ± 0 03	0 03 ± 0 05	0 02 ± 0 03	2x
0	[258 5]^{2}	25 7 9	6 7	(Hex)4 (HexNAc)4 (uc) + (Ma)3(G cNAc)2	0 03 ± 0 0	0 05 ± 0 04	0 03 ± 0 02	2x
	[260 4]	260 5	44 0	(HexNAc) (uc) + (Ma)3(G cNAc)2	0 09 ± 0 06	0 26 ± 0 09	0 5 ± 0 09	

2	[260 4]	260 5	54 5	(HexNAc) (uc) + (Ma)3(G cNAc)2	0 9 ± 0 06	0 54 ± 0 34	0 0 ± 0 05	
3	[286]^{2}	2572 9	56 7	(Hex)2 (HexNAc)3 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	-	-	0 28 ± 0 06	
4	[323 0]^{2}	2647 0	48 0	(Hex)3 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2	-	-	0 32 ± 0	
5	[323 0]^{2}	2647 0	49 4	(Hex)3 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2	-	-	0 8 ± 0 09	
6	[323 0]^{2}	2647 0	57 8	(Hex)3 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2	-	-	0 3 ± 0 06	
7	[348 5]^{3}	4047 4	64 7	(Hex)5 (HexNAc)5 (uc) (NeuAc)4 + (Ma)3(G cNAc)2	0 0 ± 0 02	0 02 ± 0 03	0 07 ± 0 04	
8	[348 5]^{3}	4047 4	65 0	(Hex)5 (HexNAc)5 (uc) (NeuAc)4 + (Ma)3(G cNAc)2	0 07 ± 0 04	0 02 ± 0 03	0 05 ± 0 02	
9	[367 0]^{2}	2735 0	6 6	(Hex)3 (HexNAc)3 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	0 24 ± 0 08	0 30 ± 0 09	0 22 ± 0 02	
20	[367 0]^{2}	2735 0	62 6	(Hex)3 (HexNAc)3 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	0 4 ± 0 0	0 3 ± 0 09	0 3 ± 0 07	

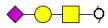
2	[367 0]^{2}	2735 0	63 2	(Hex)3 (HexNAc)3 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	0 20 ± 0 09	0 3 ± 0 3	0 2 ± 0 04	
22	[367 0]^{2}	2735 0	64 6	(Hex)3 (HexNAc)3 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	0 6 ± 0 04	0 07 ± 0	0 09 ± 0 02	
23	[440 0]^{2}	2880 0	56	(Hex)3 (HexNAc)3 (NeuAc)3 + (Ma)3(G cNAc)2	0 0 ± 0 02	0 3 ± 0 38	0 0 ± 0 0	α2-6 α2-3 α2-6
24	[440 0]^{2}	2880 0	6 8	(Hex)3 (HexNAc)3 (NeuAc)3 + (Ma)3(G cNAc)2	0 3 ± 0 05	0 50 ± 0 60	0 0 ± 0 0	α2-3 α2-3 α2-6
25	[468 6]^{2}	2938	52 3	(Hex)3 (HexNAc)4 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	-	-	0 05 ± 0 04	
26	[468 6]^{2}	2938	52 7	(Hex)3 (HexNAc)4 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	-	-	0 04 ± 0 03	
27	[468 6]^{2}	2938	59 7	(Hex)3 (HexNAc)4 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	-	-	0 03 ± 0 0	
28	[468 6]^{2}	2938	60 2	(Hex)3 (HexNAc)4 (uc) (NeuAc)2 + (Ma)3(G cNAc)2	-	-	0 02 ± 0 02	
29	[470 2]^{3}	44 2 6	66 2	(Hex)6 (HexNAc)6 (uc) (NeuAc)4 + (Ma)3(G cNAc)2	0 002 ± 0 004	0 0 ± 0 0	0 0 ± 0 004	
30	[5 3]^{2}	3026	63 9	(Hex)3 (HexNAc)3 (uc) (NeuAc)3 + (Ma)3(G cNAc)2	0 2 ± 0 07	0 7 ± 0 9	0 07 ± 0 07	

NAN1 = α2-3 neuraminidase from *Streptococcus pneumonia*
RT = Retention time
"-" = below detection

Table S3.2 Low abundance glycan compositions (no MS2)

Detected *m/z*	Theoretical Mass (Da)	Retention time (min)	Composition
[962 4]^{2}	925 7	42	(Hex)6 (Ace a e) + (Ma)3(G cNAc)2
[962 4]^{2}	925 7	42 4	(Hex)6 (Ace a e) + (Ma)3(G cNAc)2
[962 4]^{2}	925 7	43 7	(Hex)6 (Ace a e) + (Ma)3(G cNAc)2
[962 4]^{2}	925 7	44 4	(Hex)6 (Ace a e) + (Ma)3(G cNAc)2
[962 4]^{2}	925 7	46 3	(Hex)6 (Ace a e) + (Ma)3(G cNAc)2
[962 4]^{2}	925 7	47 2	(Hex)6 (Ace a e) + (Ma)3(G cNAc)2
[962 4]^{2}	925 7	47 7	(Hex)6 (Ace a e) + (Ma)3(G cNAc)2
[9 9]^{2}	2240 8	60 5	(Hex)3 (HexNAc)2 (uc) (NeuAc) + (Ma)3(G cNAc)2
[92 4]^{2}	253 9	62 5	(Hex)3 (HexNAc)2 (uc) (NeuAc)2 + (Ma)3(G cNAc)2
[373 0]^{3}	4 2 5	65	(Hex)6 (HexNAc)6 (uc) (NeuAc)3 + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	55 4	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	58 3	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	59 0	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	60	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	60 8	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	6 7	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	62 6	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	62 7	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	63 4	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	63 6	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	64 0	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	64 7	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[404 5]^{2}	2809 0	65 2	(Hex)4 (HexNAc)4 (uc) (NeuAc) + (Ma)3(G cNAc)2
[549 6]^{2}	3 00 0	6 2	(Hex)4 (HexNAc)4 (uc) (NeuAc)2 + (Ma)3(G cNAc)2

Table S3.3 Comprehensive analysis of MS/MS fragmentation spectra led to the characterisation of seven *O*-linked glycan structures in adipocytes, ADSCs and dADSCs. The details of each structure are shown below and *N*-acetylneuraminic acid linkages are added where known.

Glycan #	Detected *m/z*	Theoretical Mass (Da)	Retention time (min)	Composition	Mean % ± SD			Proposed Structure
					ADSCs	dADSCs	Adipocytes	
O	[675 3]	675 2		(Hex) (HexNAc) (NeuAc)	-	0 7 ± 0 62	2 ± 0 54	
O2	[675 3]	675 2	2 6	(Hex) (HexNAc) (NeuAc)	23 50 ± 4 47	29 25 ± 7 26	62 70 ± 3 85	
O3	[749 2]	749 3	23 8	(Hex)2 (HexNAc)2	3 ± 0 87	2 39 ± 2 5	0 53 ± 0 5	
O4	[966 4]	966 3	22 7	(Hex) (HexNAc) (NeuAc)2	53 5 ± 3 29	50 97 ± 82	30 93 ± 3 08	
O5	[040 5]	040 4	27	(Hex)2 (HexNAc)2 (NeuAc)	5 4 ± 2 94	7 86 ± 2 52	2 38 ± 23	
O6	[040 5]	040 4	29	(Hex)2 (HexNAc)2 (NeuAc)	89 ± 7	83 ± 2 02	0 45 ± 0 79	
O7	[33 4]	33 5	30 6	(Hex)2 (HexNAc)2 (NeuAc)2	3 0 ± 24	6 99 ± 2 79	45 ± 0 73	

Chapter 4 Supplementary Data

Table S4.1 Functional annotation of all proteins identified in native adipocytes, ADSCs and dADSCs according to cellular component at GO level 3. Four hundred and four proteins were successfully categorised by DAVID and categories are listed in order of enrichment with respect to the total human proteome (lowest to highest *P*-value).

Cellular Component	Number of Proteins	*P*-value	Benjamini-Hochberg *P*-value	% of categorised proteins
cytop asm	316	6.30E-36	5.70E-34	78.2
organe e membrane	105	7.40E-32	3.40E-30	26.0
organe e enve ope	77	3.80E-30	1.40E-28	19.1
organe e nner membrane	57	1.60E-29	4.70E-28	14.1
m tochondr a membrane	61	6.90E-29	1.80E-27	15.1
m tochondr a enve ope	62	2.90E-28	6.50E-27	15.3
ntrace u ar organe e	317	5.00E-17	9.10E-16	78.5
membrane-bounded ves c e	55	1.70E-16	3.70E-15	13.6
membrane	271	3.00E-15	4.50E-14	67.1
cytop asm c ves c e	55	3.90E-14	5.40E-13	13.6
ntrace u ar	350	3.90E-12	4.50E-11	86.6
nso ub e fract on	59	1.60E-11	1.70E-10	14.6
ntrace u ar membrane-bounded organe e	275	1.30E-10	1.20E-09	68.1
organe e outer membrane	19	2.30E-10	2.10E-09	4.7
outer membrane	19	4.50E-10	3.90E-09	4.7
co agen	12	9.30E-10	7.70E-09	3.0
p asma membrane	147	1.70E-07	1.20E-06	36.4
endomembrane system	47	3.20E-07	2.20E-06	11.6
nuc eo d	9	1.00E-06	7.00E-06	2.2
ce surface	27	2.10E-06	1.30E-05	6.7
nuc ear enve ope-endop asm c ret cu um network	24	2.10E-06	1.30E-05	5.9
basement membrane	12	6.20E-06	3.60E-05	3.0
organe e umen	79	1.00E-05	5.90E-05	19.6
proton-transport ng ATP synthase comp ex	7	1.30E-05	7.10E-05	1.7
extrace u ar matr x	25	1.70E-05	9.10E-05	6.2
proton-transport ng two-sector ATPase comp ex	9	2.10E-05	1.10E-04	2.2
resp ratory cha n	11	2.70E-05	1.40E-04	2.7
ntrace u ar organe e umen	76	2.80E-05	1.40E-04	18.8
endop asm c ret cu um membrane	21	3.40E-05	1.60E-04	5.2
prote naceous extrace u ar matr x	23	4.60E-05	2.10E-04	5.7
ntrace u ar non-membrane-bounded organe e	100	6.90E-05	3.00E-04	24.8
proton-transport ng ATP synthase comp ex, cata yt c core F(1)	4	3.50E-04	1.50E-03	1.0
ce project on	35	4.60E-04	1.80E-03	8.7
nuc ear membrane	9	6.80E-04	2.70E-03	2.2
proton-transport ng two-sector ATPase comp ex, cata yt c doma n	5	1.10E-03	4.30E-03	1.2
nuc ear enve ope	15	1.20E-03	4.40E-03	3.7
mono ayer-surrounded p d storage body	4	1.40E-03	5.10E-03	1.0

Cellular Component	Number of Proteins	*P*-value	Benjamini-Hochberg *P*-value	% of categorised proteins
o gosacchary transferase comp ex	4	1.90E-03	7.00E-03	1.0
proton-transport ng two-sector ATPase comp ex, proton-transport ng doma n	5	3.40E-03	1.20E-02	1.2
ce ead ng edge	11	3.70E-03	1.30E-02	2.7
membrane raft	11	4.80E-03	1.60E-02	2.7
nuc ear outer membrane	4	5.40E-03	1.80E-02	1.0
nuc eosome	7	6.30E-03	2.10E-02	1.7
proton-transport ng ATP synthase comp ex, coup ng factor F(o)	4	6.60E-03	2.10E-02	1.0
myos n comp ex	7	7.40E-03	2.30E-02	1.7
ruff e	7	8.50E-03	2.60E-02	1.7
receptor comp ex	9	1.20E-02	3.60E-02	2.2
externa s de of p asma membrane	11	1.50E-02	4.50E-02	2.7
neuron project on	17	2.00E-02	6.00E-02	4.2
am n n comp ex	3	2.20E-02	6.40E-02	0.7
NADH dehydrogenase comp ex	5	2.50E-02	6.90E-02	1.2
ves c e umen	5	3.30E-02	8.90E-02	1.2
ce soma	10	3.50E-02	9.30E-02	2.5
hemog ob n comp ex	3	3.90E-02	1.00E-01	0.7
extr ns c to membrane	21	4.00E-02	1.00E-01	5.2
heterotr mer c G-prote n comp ex	4	4.10E-02	1.00E-01	1.0
m crobody membrane	4	4.80E-02	1.20E-01	1.0

Table S4.2 101 proteins were unchanged during *in vitro* adipogenic differentiation (dADSCs relative to ADSCs)

Identifier	Description	Average logNSAF ADSCs	Average logNSAF dADSCs	*P*-value
sp|P25705|ATPA_HUMAN	ATP synthase subunit alpha, mitochondrial	-5.99	-4.99	5.01E-02
sp|P05023|AT1A1_HUMAN	Sodium/potassium-transporting ATPase subunit alpha-1	-5.65	-6.80	5.02E-02
sp|Q6NZI2|PTRF_HUMAN	Polymerase I and transcript release factor	-6.53	-5.20	5.28E-02
sp|P38646|GRP75_HUMAN	Stress-70 protein, mitochondrial	-6.94	-5.69	5.37E-02
sp|P08473|NEP_HUMAN	Neprilysin; CD10	-7.03	-6.30	5.46E-02
sp|P35232|PHB_HUMAN	Prohibitin	-6.35	-5.56	6.04E-02
sp|P23396|RS3_HUMAN	40S ribosomal protein S3	-5.41	-6.51	6.06E-02
sp|P11142|HSP7C_HUMAN	Heat shock cognate 71 kDa protein	-5.86	-5.68	6.08E-02
sp|P00338|LDHA_HUMAN	L-lactate dehydrogenase A chain	-6.44	-7.28	6.09E-02
sp|P07437|TBB5_HUMAN	Tubulin beta chain	-4.67	-5.42	6.46E-02
sp|Q04941|PLP2_HUMAN	Proteolipid protein 2; Differentiation-dependent protein A4	-4.57	-5.56	6.76E-02
sp|P08195|4F2_HUMAN	4F2 cell-surface antigen heavy chain; CD98	-5.98	-7.36	6.78E-02
sp|O00159|MYO1C_HUMAN	Unconventional myosin-Ic	-5.87	-7.27	7.38E-02
sp|P13645|K1C10_HUMAN	Keratin, type I cytoskeletal 10	-3.81	-4.29	7.61E-02
sp|P68371|TBB4B_HUMAN	Tubulin beta-4B chain	-4.59	-5.44	8.26E-02
sp|P63261|ACTG_HUMAN	Actin, cytoplasmic 2	-2.57	-3.10	8.27E-02
sp|Q15758|AAAT_HUMAN	Neutral amino acid transporter B(0)	-5.41	-6.23	8.59E-02
sp|P60033|CD81_HUMAN	CD81 antigen	-6.11	-4.91	9.13E-02
sp|P11021|GRP78_HUMAN	78 kDa glucose-regulated protein	-5.78	-5.10	9.35E-02
sp|Q6DD88|ATLA3_HUMAN	Atlastin-3	-6.54	-7.61	9.82E-02
sp|P10809|CH60_HUMAN	60 kDa heat shock protein, mitochondrial	-6.43	-5.22	9.88E-02
sp|P08670|VIME_HUMAN	Vimentin	-3.44	-2.55	1.03E-01
sp|P18206|VINC_HUMAN	Vinculin	-6.43	-7.72	1.04E-01

Identifier	Description	Average logNSAF ADSCs	Average logNSAF dADSCs	*P*-value
sp|P04264|K2C1_HUMAN	Keratin, type II cytoskeletal 1	-3.50	-4.06	1.04E-01
sp|O00264|PGRC1_HUMAN	Membrane-associated progesterone receptor component 1	-5.37	-4.75	1.10E-01
sp|Q99715|COCA1_HUMAN	Collagen type XII, alpha-1 chain	-6.36	-8.09	1.10E-01
sp|P06576|ATPB_HUMAN	ATP synthase subunit beta, mitochondrial	-5.15	-4.36	1.14E-01
sp|Q02413|DSG1_HUMAN	Desmoglein-1; Cadherin family member 4	-7.29	-8.12	1.16E-01
sp|Q7KZF4|SND1_HUMAN	Staphylococcal nuclease domain-containing protein 1	-7.01	-7.65	1.16E-01
sp|O60664|PLIN3_HUMAN	Perilipin-3; 47 kDa mannose 6-phosphate receptor-binding protein	-6.69	-7.18	1.22E-01
sp|P60709|ACTB_HUMAN	Actin, cytoplasmic 1; Beta-actin	-2.55	-3.04	1.23E-01
sp|P16671|CD36_HUMAN	Platelet glycoprotein 4; CD36	-4.53	-5.41	1.38E-01
sp|Q9H0U4|RAB1B_HUMAN	Ras-related protein Rab-1B	-5.09	-5.30	1.56E-01
sp|Q8TCT9|HM13_HUMAN	Minor histocompatibility antigen H13	-5.75	-6.57	1.61E-01
sp|P10620|MGST1_HUMAN	Microsomal glutathione S-transferase 1	-5.36	-4.52	1.68E-01
sp|Q9Y3E0|GOT1B_HUMAN	Vesicle transport protein GOT1B	-5.31	-5.96	1.70E-01
sp|Q9Y639|NPTN_HUMAN	Neuroplastin	-6.46	-7.18	1.73E-01
sp|P29966|MARCS_HUMAN	Myristoylated alanine-rich C-kinase substrate	-5.63	-6.64	2.25E-01
sp|O75369|FLNB_HUMAN	Filamin-B	-8.12	-8.85	2.35E-01
sp|P54289|CA2D1_HUMAN	Voltage-dependent calcium channel subunit alpha-2/delta-1	-7.55	-7.00	2.35E-01
sp|Q16853|AOC3_HUMAN	Membrane primary amine oxidase	-5.06	-5.74	2.36E-01
sp|P35613|BASI_HUMAN	Basigin; CD147	-5.55	-6.27	2.40E-01
sp|Q9BUF5|TBB6_HUMAN	Tubulin beta-6 chain	-5.56	-6.22	2.47E-01
sp|Q96S97|MYADM_HUMAN	Myeloid-associated differentiation marker	-5.52	-6.51	2.51E-01
sp|P02545|LMNA_HUMAN	Prelamin-A/C	-5.58	-5.08	2.53E-01
sp|P48047|ATPO_HUMAN	ATP synthase subunit O, mitochondrial	-6.27	-5.60	2.63E-01

Identifier	Description	Average logNSAF ADSCs	Average logNSAF dADSCs	*P*-value
sp|Q9HDC9|APMAP_HUMAN	Adipocyte plasma membrane-associated protein	-6.85	-6.36	2.81E-01
sp|P04843|RPN1_HUMAN	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	-6.63	-6.30	2.85E-01
sp|Q9NX63|CHCH3_HUMAN	Coiled-coil-helix-coiled-coil-helix domain-containing protein 3, mitochondrial	-6.38	-5.78	2.89E-01
sp|O15173|PGRC2_HUMAN	Membrane-associated progesterone receptor component 2	-5.36	-4.88	2.90E-01
sp|P49327|FAS_HUMAN	Fatty acid synthase	-8.08	-7.25	3.00E-01
sp|P62820|RAB1A_HUMAN	Ras-related protein Rab-1A	-4.94	-5.51	3.12E-01
sp|P05556|ITB1_HUMAN	Integrin beta-1; Fibronectin receptor subunit beta; CD29	-5.85	-6.37	3.22E-01
sp|P16070|CD44_HUMAN	CD44 antigen; Heparan sulfate proteoglycan	-5.38	-5.94	3.40E-01
sp|Q13740|CD166_HUMAN	CD166 antigen; Activated leukocyte cell adhesion molecule	-6.42	-7.36	3.45E-01
sp|Q7Z794|K2C1B_HUMAN	Keratin, type II cytoskeletal 1b	-5.22	-5.63	3.49E-01
sp|P00387|NB5R3_HUMAN	NADH-cytochrome b5 reductase 3	-4.98	-4.54	3.54E-01
sp|P06756|ITAV_HUMAN	Integrin alpha-V; Vitronectin receptor subunit alpha; CD51	-6.68	-6.95	3.58E-01
sp|Q9NQC3|RTN4_HUMAN	Reticulon-4	-7.07	-6.46	3.63E-01
sp|P35527|K1C9_HUMAN	Keratin, type I cytoskeletal 9; Cytokeratin-9	-4.09	-4.40	3.68E-01
sp|Q9UHG3|PCYOX_HUMAN	Prenylcysteine oxidase 1	-6.42	-6.85	3.92E-01
sp|Q99623|PHB2_HUMAN	Prohibitin-2	-6.00	-5.60	3.95E-01
sp|P21589|5NTD_HUMAN	5--nucleotidase; CD73	-5.40	-6.23	3.97E-01
sp|Q562R1|ACTBL_HUMAN	Beta-actin-like protein 2	-3.98	-4.19	4.49E-01
sp|Q15149|PLEC_HUMAN	Plectin	-7.35	-8.04	4.64E-01
sp|P36957|ODO2_HUMAN	Dihydrolipoyllysine-residue succinyltransferase component of 2-oxoglutarate dehydrogenase complex, mitochondrial	-7.18	-6.99	4.74E-01
sp|P05186|PPBT_HUMAN	Alkaline phosphatase, tissue-nonspecific isozyme	-6.91	-7.27	4.77E-01
sp|Q00325|MPCP_HUMAN	Phosphate carrier protein, mitochondrial	-6.10	-6.40	4.79E-01

Identifier	Description	Average logNSAF ADSCs	Average logNSAF dADSCs	*P*-value
sp|P22695|QCR2_HUMAN	Cytochrome b-c1 complex subunit 2, mitochondrial	-6.51	-6.13	5.04E-01
sp|P61019|RAB2A_HUMAN	Ras-related protein Rab-2A	-5.84	-5.55	5.25E-01
sp|P02533|K1C14_HUMAN	Keratin, type I cytoskeletal 14	-4.39	-4.88	5.28E-01
sp|Q9P0S3|ORML1_HUMAN	ORM1-like protein 1; Adoplin-1	-6.15	-5.99	5.53E-01
sp|P05091|ALDH2_HUMAN	Aldehyde dehydrogenase, mitochondrial	-7.15	-6.91	5.60E-01
sp|P51149|RAB7A_HUMAN	Ras-related protein Rab-7a	-5.80	-5.51	5.62E-01
sp|P15144|AMPN_HUMAN	Aminopeptidase N; CD13	-4.80	-5.15	5.67E-01
sp|P11279|LAMP1_HUMAN	Lysosome-associated membrane glycoprotein 1	-6.08	-6.29	6.14E-01
sp|Q9H4G4|GAPR1_HUMAN	Golgi-associated plant pathogenesis-related protein 1	-5.11	-5.33	6.42E-01
sp|Q8WZ42|TITIN_HUMAN	Titin	-10.63	-11.02	6.43E-01
sp|P23284|PPIB_HUMAN	Peptidyl-prolyl cis-trans isomerase B	-5.83	-5.69	6.54E-01
sp|Q9Y5M8|SRPRB_HUMAN	Signal recognition particle receptor subunit beta	-6.51	-6.32	6.75E-01
sp|P14625|ENPL_HUMAN	Endoplasmin; 94 kDa glucose-regulated protein	-6.02	-5.89	6.79E-01
sp|Q9NZN4|EHD2_HUMAN	EH domain-containing protein 2	-6.69	-6.93	6.92E-01
sp|P04792|HSPB1_HUMAN	Heat shock protein beta-1	-5.59	-5.42	6.96E-01
sp|P13647|K2C5_HUMAN	Keratin, type II cytoskeletal 5	-5.62	-5.87	7.23E-01
sp|P50454|SERPH_HUMAN	Serpin H1	-5.90	-5.67	7.26E-01
sp|Q14108|SCRB2_HUMAN	Lysosome membrane protein 2; CD36 antigen-like 2	-6.47	-6.14	7.32E-01
sp|Q13423|NNTM_HUMAN	NAD(P) transhydrogenase, mitochondrial	-7.40	-7.21	7.33E-01
sp|P02786|TFR1_HUMAN	Transferrin receptor protein; CD71	-7.15	-7.40	7.50E-01
sp|P08779|K1C16_HUMAN	Keratin, type I cytoskeletal 16	-4.71	-4.97	7.54E-01
sp|P04406|G3P_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase; GAPDH	-5.78	-5.69	8.37E-01
sp|Q03135|CAV1_HUMAN	Caveolin-1	-4.53	-4.60	8.48E-01
sp|O00116|ADAS_HUMAN	Alkyldihydroxyacetonephosphate synthase, peroxisomal	-7.24	-7.19	8.66E-01

Identifier	Description	Average logNSAF ADSCs	Average logNSAF dADSCs	*P*-value
sp|P07355|ANXA2_HUMAN	Annexin A2	-5.28	-5.35	8.68E-01
sp|O75396|SC22B_HUMAN	Vesicle-trafficking protein SEC22b; ER-Golgi SNARE of 24 kDa	-5.41	-5.28	8.70E-01
sp|O15260|SURF4_HUMAN	Surfeit locus protein 4	-6.17	-6.12	9.01E-01
sp|P43121|MUC18_HUMAN	Melanoma cell adhesion molecule; CD146	-6.83	-6.90	9.10E-01
sp|P51148|RAB5C_HUMAN	Ras-related protein Rab-5C	-5.23	-5.25	9.37E-01
sp|Q9BVC6|TM109_HUMAN	Transmembrane protein 109	-5.67	-5.71	9.44E-01
sp|P80723|BASP1_HUMAN	Brain acid soluble protein 1	-4.66	-4.68	9.51E-01
sp|Q09666|AHNK_HUMAN	Neuroblast differentiation-associated protein AHNAK	-8.49	-8.46	9.69E-01
sp|Q07065|CKAP4_HUMAN	Cytoskeleton-associated protein 4; p63	-5.21	-5.20	9.79E-01

Table S4.3 75 proteins were unchanged during *in vivo* adipogenic differentiation (Adipocytes relative to ADSCs)

Identifier	Description	Average logNSAF ADSCs	Average logNSAF Adipocytes	*P*-value
sp|Q9BVC6|TM109_HUMAN	Transmembrane protein 109	-5.66	-6.62	5.00E-02
sp|P35527|K1C9_HUMAN	Keratin, type I cytoskeletal 9	-4.08	-5.07	5.43E-02
sp|Q14697|GANAB_HUMAN	Neutral alpha-glucosidase AB	-7.07	-7.74	5.66E-02
sp|Q9HDC9|APMAP_HUMAN	Adipocyte plasma membrane-associated protein	-6.84	-5.74	5.68E-02
sp|P62987|RL40_HUMAN	Ubiquitin-60S ribosomal protein L40	-4.81	-5.43	7.09E-02
sp|P04843|RPN1_HUMAN	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1	-6.62	-5.90	7.24E-02
sp|Q16853|AOC3_HUMAN	Membrane primary amine oxidase	-5.06	-3.62	7.25E-02
sp|P22695|QCR2_HUMAN	Cytochrome b-c1 complex subunit 2, mitochondrial	-6.50	-5.30	7.30E-02
sp|P54289|CA2D1_HUMAN	Voltage-dependent calcium channel subunit alpha-2/delta-1	-7.54	-6.54	7.33E-02
sp|Q8TCT9|HM13_HUMAN	Minor histocompatibility antigen H13	-5.74	-6.99	8.37E-02
sp|P07437|TBB5_HUMAN	Tubulin beta chain	-4.66	-4.17	8.49E-02
sp|Q9Y3E0|GOT1B_HUMAN	Vesicle transport protein GOT1B	-5.30	-6.42	8.59E-02
sp|P60033|CD81_HUMAN	CD81	-6.10	-4.87	8.66E-02
sp|P35579|MYH9_HUMAN	Myosin-9	-4.92	-5.94	9.47E-02
sp|P00387|NB5R3_HUMAN	NADH-cytochrome b5 reductase 3	-4.97	-4.00	9.58E-02
sp|Q96S97|MYADM_HUMAN	Myeloid-associated differentiation marker	-5.51	-6.72	1.02E-01
sp|Q9BSJ8|ESYT1_HUMAN	Extended synaptotagmin-1	-7.56	-6.79	1.03E-01
sp|P68104|EF1A1_HUMAN	Elongation factor 1-alpha 1	-5.39	-5.81	1.03E-01
sp|Q6NZI2|PTRF_HUMAN	Polymerase I and transcript release factor	-6.52	-5.50	1.08E-01
sp|P35908|K22E_HUMAN	Keratin, type II cytoskeletal 2 epidermal	-4.18	-5.01	1.08E-01
sp|P21796|VDAC1_HUMAN	Voltage-dependent anion-selective channel protein 1	-5.88	-4.62	1.09E-01
sp|P06576|ATPB_HUMAN	ATP synthase subunit beta, mitochondrial	-5.14	-3.84	1.11E-01

Identifier	Description	Average logNSAF ADSCs	Average logNSAF Adipocytes	*P*-value
sp|Q9NZM1|MYOF_HUMAN	Myoferlin	-6.38	-8.06	1.13E-01
sp|Q00610|CLH1_HUMAN	Clathrin heavy chain 1	-7.01	-8.52	1.16E-01
sp|P51149|RAB7A_HUMAN	Ras-related protein Rab-7a	-5.79	-6.59	1.17E-01
sp|P02533|K1C14_HUMAN	Keratin, type I cytoskeletal 14	-4.38	-6.15	1.24E-01
sp|P35232|PHB_HUMAN	Prohibitin	-6.34	-5.49	1.35E-01
sp|Q03135|CAV1_HUMAN	Caveolin-1	-4.52	-3.80	1.38E-01
sp|P10620|MGST1_HUMAN	Microsomal glutathione S-transferase 1	-5.35	-4.36	1.39E-01
sp|O00264|PGRC1_HUMAN	Membrane-associated progesterone receptor component 1	-5.36	-4.64	1.43E-01
sp|P13645|K1C10_HUMAN	Keratin, type I cytoskeletal 10	-3.81	-4.61	1.45E-01
sp|Q9BUF5|TBB6_HUMAN	Tubulin beta-6 chain	-5.55	-4.81	1.49E-01
sp|P11279|LAMP1_HUMAN	Lysosome-associated membrane glycoprotein 1; CD107a	-6.07	-6.56	1.50E-01
sp|O15173|PGRC2_HUMAN	Membrane-associated progesterone receptor component 2	-5.35	-4.53	1.64E-01
sp|Q14108|SCRB2_HUMAN	Lysosome membrane protein 2; CD36 antigen-like 2	-6.46	-7.40	1.65E-01
sp|P04792|HSPB1_HUMAN	Heat shock protein beta-1	-5.58	-6.48	1.97E-01
sp|Q9H4G4|GAPR1_HUMAN	Golgi-associated plant pathogenesis-related protein 1	-5.10	-5.85	2.01E-01
sp|P06756|ITAV_HUMAN	Integrin alpha-V; Vitronectin receptor subunit alpha; CD51	-6.67	-7.17	2.05E-01
sp|P25705|ATPA_HUMAN	ATP synthase subunit alpha, mitochondrial	-5.99	-5.26	2.08E-01
sp|P08473|NEP_HUMAN	Neprilysin; CD10	-7.03	-7.62	2.21E-01
sp|P68371|TBB4B_HUMAN	Tubulin beta-4B chain	-4.58	-4.22	2.39E-01
sp|Q9H0U4|RAB1B_HUMAN	Ras-related protein Rab-1B	-5.08	-5.56	2.47E-01
sp|Q01082|SPTB2_HUMAN	Spectrin beta chain, non-erythrocytic 1	-7.67	-8.28	2.49E-01
sp|Q9Y639|NPTN_HUMAN	Neuroplastin	-6.45	-6.88	2.66E-01
sp|P38646|GRP75_HUMAN	Stress-70 protein, mitochondrial	-6.93	-6.34	2.72E-01
sp|P05997|CO5A2_HUMAN	Collagen type V, alpha-2 chain	-7.72	-8.41	2.87E-01

Identifier	Description	Average logNSAF ADSCs	Average logNSAF Adipocytes	*P*-value
sp|Q8WZ42|TITIN_HUMAN	Titin	-10.63	-11.56	2.95E-01
sp|P08754|GNAI3_HUMAN	Guanine nucleotide-binding protein G(k) subunit alpha	-5.90	-6.24	3.05E-01
sp|P02452|CO1A1_HUMAN	Collagen type I, alpha-1 chain	-5.82	-6.24	3.15E-01
sp|Q07954|LRP1_HUMAN	Prolow-density lipoprotein receptor-related protein 1	-8.49	-9.04	3.42E-01
sp|P16671|CD36_HUMAN	Platelet glycoprotein 4; CD36	-4.53	-3.80	3.44E-01
sp|Q99623|PHB2_HUMAN	Prohibitin-2	-5.99	-6.41	3.59E-01
sp|Q09666|AHNK_HUMAN	Neuroblast differentiation-associated protein AHNAK	-8.48	-9.23	3.69E-01
sp|P27105|STOM_HUMAN	Erythrocyte band 7 integral membrane protein; Stomatin	-4.63	-4.37	4.21E-01
sp|P49748|ACADV_HUMAN	Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	-7.08	-6.49	4.28E-01
sp|P05556|ITB1_HUMAN	Integrin beta-1; Fibronectin receptor subunit beta; CD29	-5.85	-5.43	4.32E-01
sp|P11021|GRP78_HUMAN	78 kDa glucose-regulated protein	-5.78	-6.16	4.47E-01
sp|P08123|CO1A2_HUMAN	Collagen type I, alpha-2 chain	-6.13	-6.54	4.65E-01
sp|P04083|ANXA1_HUMAN	Annexin A1	-6.49	-6.07	4.67E-01
sp|Q00325|MPCP_HUMAN	Phosphate carrier protein, mitochondrial	-6.09	-5.76	4.74E-01
sp|Q6DD88|ATLA3_HUMAN	Atlastin-3	-6.53	-6.15	4.86E-01
sp|P16615|AT2A2_HUMAN	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	-6.55	-7.12	5.06E-01
sp|O15260|SURF4_HUMAN	Surfeit locus protein 4	-6.16	-6.74	5.07E-01
sp|P04899|GNAI2_HUMAN	Guanine nucleotide-binding protein G(i) subunit alpha-2	-5.14	-5.26	5.55E-01
sp|P23634|AT2B4_HUMAN	Plasma membrane calcium-transporting ATPase 4	-6.57	-6.84	5.89E-01
sp|P18206|VINC_HUMAN	Vinculin; Metavinculin	-6.43	-6.87	6.06E-01
sp|Q99536|VAT1_HUMAN	Synaptic vesicle membrane protein VAT-1 homolog	-5.79	-6.10	6.18E-01
sp|P35613|BASI_HUMAN	Basigin; CD147	-5.55	-5.36	7.31E-01
sp|P07355|ANXA2_HUMAN	Annexin A2	-5.28	-5.04	7.48E-01
sp|P00338|LDHA_HUMAN	L-lactate dehydrogenase A chain	-6.43	-6.61	7.50E-01

Identifier	Description	Average logNSAF ADSCs	Average logNSAF Adipocytes	*P*-value
sp|P12111|CO6A3_HUMAN	Collagen type VI, alpha-3 chain	-6.88	-6.74	8.03E-01
sp|P10809|CH60_HUMAN	60 kDa heat shock protein, mitochondrial	-6.43	-6.26	8.16E-01
sp|Q92629|SGCD_HUMAN	Delta-sarcoglycan	-6.58	-6.84	6.90E-01
sp|Q16851|UGPA_HUMAN	UTP--glucose-1-phosphate uridylyltransferase	-7.15	-7.30	6.99E-01
sp|Q9H3K2|GHITM_HUMAN	Growth hormone-inducible transmembrane protein	-6.73	-7.02	3.05E-01

Appendix A

1. Ethics Approval

MACQUARIE UNIVERSITY

KATHERINE WONGTRAKUL-KISH <katherine.wongtrakul-kish@students.mq.edu.au>

Final Approval- Ethics application reference-5201100385

Ethics Secretariat <ethics.secretariat@mq.edu.au> Tue, Jun 28, 2011 at 3:24 PM
To: Mr Benjamin Herbert <benjamin.herbert@mq.edu.au>
Cc: Miss Sinead Pamela Blaber <sinead.blaber@students.mq.edu.au>, Mr Cameron John Hill <cameron.hill@students.mq.edu.au>, Miss Rebecca Anne Webster <rebecca.webster@students.mq.edu.au>, Mr Michael John Medynskyj <michael.medynskyj@students.mq.edu.au>, Miss Katherine Wongtrakul-Kish <katherine.wongtrakul-kish@students.mq.edu.au>

Dear Dr Herbert

Re: "Investigation of the properties and secretions of adipose tissue and cells isolated from adipose tissue" (Ethics Ref: 5201100385)

Thank you for your recent correspondence. Your response has addressed the issues raised by the Human Research Ethics Committee and you may now commence your research.

Regarding your query in your response to point 4, the committee recommends that for completeness of research the dataset can ethically be retained (unless specifically requested by the participant when they withdraw from the study)

The following personnel are authorised to conduct this research:

Dr Benjamin Herbert- Chief Investigator/Supervisor
Mr Cameron Hill, Mr Michael Medynskyj, Mr Sinead Blaber, Ms Katherine Wongtrakul-kish
& Ms Rebecca Webster- Co-investigators

NB. STUDENTS: IT IS YOUR RESPONSIBILITY TO KEEP A COPY OF THIS APPROVAL EMAIL TO SUBMIT WITH YOUR THESIS.

Please note the following standard requirements of approval:

1. The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Human Research (2007).

2. Approval will be for a period of five (5) years subject to the provision of annual reports. Your first progress report is due on 28 June 2012.

If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report for the project.

Progress reports and Final Reports are available at the following website:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/forms

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit

on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. All amendments to the project must be reviewed and approved by the Committee before implementation. Please complete and submit a Request for Amendment Form available at the following website:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/forms

5. Please notify the Committee immediately in the event of any adverse effects on participants or of any unforeseen events that affect the continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at the following websites:

http://www.mq.edu.au/policy/

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics/policy

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have final approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

If you need to provide a hard copy letter of Final Approval to an external organisation as evidence that you have Final Approval, please do not hesitate to contact the Ethics Secretariat at the address below.

Please retain a copy of this email as this is your official notification of final ethics approval.

Yours sincerely
Dr Karolyn White
Director of Research Ethics
Chair, Human Research Ethics Committee

2. Contents of Accompanying DVD

- **Additional Chapter 4 Supplementary Data**

Table S4.4 - Normalized Spectral Abundance Factor (NSAF) of total proteins detected, including a spectral fraction of 0.5 to all

Table S4.5. - Spectral counts of total proteins detected

Table S4.6:
Sheet 1: Proteins found only in ADSCs
Sheet 2: Proteins found only in dADSCs
Sheet 3: Proteins found only in Adipocytes
Sheet 4: Proteins found in both dADSCs and Adipocytes

- **Additional Chapter 5 Supplementary Data**

Table S5.1. Proteins identified in each de-*N*-glycosylated protein gel band