Investigation of the extent and role of
N-linked glycosylation in the human scavenger
receptor CD36

Sarah Jayne Hoosdally

A thesis submitted for the degree of
Doctor of Philosophy

Imperial College London
MRC Clinical Sciences Centre
2009
Abstract

Human CD36 is a class B scavenger receptor expressed in a variety of cell types such as macrophage and adipocytes. This plasma membrane glycoprotein has a wide range of ligands including oxidised low density lipoprotein (oxLDL) and long chain fatty acids which involves the receptor in diseases such as atherosclerosis and insulin resistance. CD36 is heavily modified post-translationally by N-linked glycosylation and ten putative N-linked glycosylation sites situated in the large extracellular loop of the protein have been identified, however their utilisation and role in the folding and function of the protein have not been characterised. Using mass spectrometry on purified and PNGaseF-deglycosylated CD36, and also by comparing the electrophoretic mobility of different glycosylation-site mutants, this study determined that nine of the ten sites can be modified by glycosylation. Flow cytometric analysis of the different glycosylation mutants expressed in mammalian cells, established that glycosylation is necessary for trafficking to the plasma membrane. Minimally-glycosylated mutants that supported trafficking were identified and indicated the importance of carboxy-terminal sites N247, N321 and N417 and amino-terminal sites N102 and N205. However, unlike the related mouse scavenger receptor SR-BI, no individual site was found to be essential for proper trafficking of CD36. Surprisingly, these minimally-glycosylated mutants appear to be predominantly core glycosylated indicating that mature glycosylation is not necessary for surface expression in mammalian cells. The data also show that neither the nature nor the pattern of glycosylation is relevant to binding of modified LDL.
Declaration of own work

I hereby declare that all experiments presented in this thesis are my own work, except for the following:-

- Automated DNA sequencing was performed by Lisa Lowery and Ivan Andrew in MRC Clinical Sciences Centre Genomics Core Laboratory.
- Plasmid vector pCD36-12His was provided by Dr Kenneth Linton.
- CD36nong was generated by Mr Edward Andress.
- Q-ToF mass spectrometry was carried out by MRC Clinical Sciences Centre Proteomics and Peptide Synthesis facility.
- FT-ICR mass spectrometry was carried out by Dr Helen Cooper, Functional Genomics and Proteomics Unit, School of Biosciences, University of Birmingham.

Sarah J Hoosdally

PhD Student

Dr Kenneth Linton

Supervisor
Contents

Abstract ........................................................................................................................................... 2
Declaration of own work .................................................................................................................... 3
Contents .............................................................................................................................................. 4
List of figures ...................................................................................................................................... 8
List of tables ....................................................................................................................................... 11
Acknowledgements .......................................................................................................................... 12

1. Introduction .................................................................................................................................. 14
  1.1. Identification of Class B scavenger receptor CD36 ................................................................. 14
  1.2. Structure of CD36 ................................................................................................................... 16
    1.2.1. Extracellular domain of CD36 ....................................................................................... 18
    1.2.2. Ligands of CD36 and localisation of binding sites ........................................................ 19
    1.2.3. Post-translational modifications .................................................................................... 22
      1.2.3.1. Phosphorylation .................................................................................................... 22
      1.2.3.2. N-linked glycosylation ......................................................................................... 23
      1.2.3.3. Disulphide bonds ................................................................................................. 24
      1.2.3.4. Palmitoylation ...................................................................................................... 27
  1.3. Expression and Regulation of CD36 ......................................................................................... 28
    1.3.1. Regulation of CD36 .................................................................................................... 28
    1.3.2. Naturally occurring CD36 mutations and variants ....................................................... 29
  1.4. CD36 in normal and disease pathology .................................................................................... 32
    1.4.1. Innate Immunity ........................................................................................................ 32
    1.4.2. CD36 and atherosclerosis ......................................................................................... 33
    1.4.3. Role of CD36 in pathogenesis of Plasmodium falciparum malaria and sickle cell anaemia ......................................................................................................................... 35
    1.4.4. CD36 and angiogenesis ................................................................................................. 36
    1.4.5. CD36 and fatty acid transport ..................................................................................... 37
    1.4.6. CD36, insulin resistance and metabolic syndrome ..................................................... 38
  1.5. Biosynthesis of N-linked glycosylation .................................................................................... 40
    1.5.1. N-linked glycosylation in the ER ............................................................................... 41
    1.5.2. N-linked glycosylation in the Golgi apparatus ............................................................. 43
1.6. Objectives of thesis ........................................................................................................44

2. Material and methods ..................................................................................................46

2.1. Bacterial culture medium .....................................................................................46

2.2. Bacterial culture and storage ...............................................................................46

2.3. Molecular Biology .................................................................................................47

2.3.1. Plasmids ..........................................................................................................47

2.3.2. Preparation of plasmid DNA ...........................................................................48

2.3.3. Site-directed mutagenesis ...............................................................................50

2.3.4. Agarose gel electrophoresis ...........................................................................52

2.3.5. Purification of DNA fragment from agarose gels ...........................................52

2.3.6. Subcloning of plasmid DNA ..........................................................................53

2.4. Insect cell culture ....................................................................................................53

2.4.1. Insect cell culture medium and reagents .......................................................53

2.4.2. Insect cell co-transfection ...............................................................................53

2.4.3. Plaque assay to isolate recombinant baculovirus ...........................................55

2.4.4. Amplification of baculovirus ..........................................................................56

2.4.4.1. Large-scale infection for protein production ..............................................58

2.4.4.2. Harvesting insect cell membrane ...............................................................58

2.5. Mammalian Tissue Culture ..................................................................................59

2.5.1. Mammalian tissue culture medium and reagents .........................................59

2.5.2. Transient transfection of HEK293T cells .......................................................59

2.5.3. Tunicamycin Treatment ..................................................................................60

2.6. Flow cytometry ......................................................................................................60

2.6.1. Cell surface expression of CD36 ....................................................................61

2.7. Protein Biochemistry ...........................................................................................62

2.7.1. Materials used for protein biochemistry .......................................................62

2.7.2. Purification of CD36 by affinity chromatography ..........................................62

2.7.2.1. Wildtype CD36 .........................................................................................62

2.7.2.2. Non-glycosylated CD36-12 Histidine .......................................................63

2.7.2.3. CD36N8-10-12 Histidine .........................................................................63

2.7.3. Trichloroacetic acid precipitation .................................................................64

2.7.4. Concentration of protein ..............................................................................64

2.7.5. Protein quantitation .......................................................................................64
2.7.6. Preparation of protein samples for SDS-Polyacrylamide gel electrophoresis (SDS-PAGE) ...........................................................................................................65
2.7.7. SDS-PAGE ...................................................................................................65
2.7.8. Detection of protein by colloidal blue staining ............................................67
2.7.9. Detection of proteins by western blotting ....................................................67
2.7.10. Solid-Phase Ligand binding assay ..........................................................68
2.7.11. Glycosidase digestion .............................................................................69
2.8. Mass Spectrometry ..........................................................................................70
2.8.1. Quadropole Time of Flight MS/MS .............................................................70
2.8.2. Fourier-transform ion cyclotron resonance (FT-ICR) MS ...........................70

3. Expression, purification and ligand binding of CD36 ........................................71

3.1. Introduction .....................................................................................................71
3.1.1. Detergents .................................................................................................72
3.2. Expression of CD36 in E.coli strains ............................................................74
3.3. N-linked glycosylation in insect cells ............................................................75
3.3.1. The baculovirus expression system ...........................................................76
3.3.2. The baculovirus transfer vectors ...............................................................77
3.3.3. Generation of CD36-expressing baculoviruses by homologous recombination ...........................................................................................................78
3.4. Results ...........................................................................................................80
3.4.1. Baculovirus amplification and protein expression .......................................81
3.4.2. Protein purification ...................................................................................82
3.4.2.1. Wild type CD36 ...............................................................................82
3.4.2.2. CD36nong ....................................................................................84
3.4.2.3. CD36N8-10 ..................................................................................86
3.4.3. Concentrating the purified protein ............................................................88
3.4.4. Deglycosylation of native purified CD36 .................................................90
3.4.5. Affinity of acetylated low-density lipoprotein for insect cell purified wild type CD36 ........................................................................................................92
3.5. Discussion ....................................................................................................94

4. Identifying the occupancy of 10 putative glycosylation sites in CD36 .................97

4.1. Introduction ...................................................................................................97
4.1.1. Mass Spectrometry ..................................................................................97
4.1.2. Liquid Chromatography Tandem mass spectrometry using Quadrupole Time-of-Flight (LC MS/MS Q-ToF) ................................................................. 99
4.1.3. Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometry introduction........................................................................................................... 100
4.2. Enzymatic deglycosylation of wild type CD36 ........................................ 102
4.3. Preparation of protein for mass spectrometry ........................................ 104
4.4. LC/MS/MS Q-ToF analysis results .......................................................... 106
4.5. FT-ICR analysis .................................................................................................... 113
4.6. Detecting glycosylation sites by different electrophoretic mobility of mutant proteins 115
4.7. Discussion ............................................................................................................. 118

5. The role of glycans in trafficking and folding of CD36 ........................................ 123
5.1. Introduction ........................................................................................................... 123
5.1.1. The mammalian expression vector, pCI-neo............................................. 123
5.1.2. Expression of CD36 using a mammalian cell system .............................. 124
5.1.3. Polyethyleneimine transfection ................................................................. 125
5.1.4. Post-transfection use of butyric acid ...................................................... 125
5.1.5. Flow cytometric analysis ........................................................................... 125
5.2. Results ................................................................................................................... 128
5.2.1. Cell surface expression of CD36 .............................................................. 128
5.2.1.1. Titration of primary antibodies ............................................................. 130
5.2.1.2. Titration of secondary antibody ........................................................... 132
5.2.2. Mutation of N-linked glycosylation sites in CD36 .................................. 134
5.2.2.1. Generation and cell surface expression of CD36nong .......................... 134
5.2.2.2. Determination of the minimal glycosylation required for cell surface expression ......................................................................................... 138
5.2.2.3. Characterization of the amino-terminal glycosylation set ............. 139
5.2.2.4. Characterization of the carboxy-terminal glycosylation set .......... 143
5.2.2.5. The effect of individual carboxy-terminal glycosylation sites on the trafficking of CD36N1-7 ................................................................. 146
5.2.2.6. The effect of amino-terminal glycosylation sites on the trafficking of CD36N8-10 ................................................................................. 148
5.2.2.7. Glycosylation and not reintroduction of asparagines is the important determinant of trafficking efficiency ............................................. 151
5.2.2.8. Analysis of the nature of the glycosylation in transiently expressed CD36 isoforms.................................................................156

5.2.2.9. Affinity CD36 glycosylation mutants for acetylated low-density lipoprotein .................................................................161

5.3. Discussion ...........................................................................................................................................................................164

6. General Discussion...............................................................................................................................................................169

6.1. Purification of wild type CD36 and mass spectrometry........................................................169

6.2. Comparison of electrophoretic mobility of CD36 mutants.........................................................171

6.3. Establishing the minimal level of glycosylation to allow cell surface expression and function of CD36 ....................................................................................................................................172

6.4. Purification of mutant CD36 isoforms ....................................................................175

6.5. The nature of mutant CD36 glycans..............................................................................177

6.6. Summary and perspective for future research.................................................................................179

References........................................................................................................181

Appendix........................................................................................................196

Abbreviations..................................................................................................198
Table of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Cartoon representation of the predicted topology of human CD36.</td>
<td>18</td>
</tr>
<tr>
<td>1.2</td>
<td>Alignment of CD36 homologues demonstrating extracellular cysteine conservation</td>
<td>26</td>
</tr>
<tr>
<td>1.3</td>
<td>The core N-linked glycan.</td>
<td>41</td>
</tr>
<tr>
<td>1.4</td>
<td>N-linked glycan processing in the ER.</td>
<td>43</td>
</tr>
<tr>
<td>1.5</td>
<td>N-linked glycan processing in the Golgi.</td>
<td>44</td>
</tr>
<tr>
<td>2.1</td>
<td>pCD36-12His plasmid used in this study.</td>
<td>47</td>
</tr>
<tr>
<td>3.1</td>
<td>Detergent solubilization of a membrane protein.</td>
<td>73</td>
</tr>
<tr>
<td>3.2</td>
<td>Homologous recombination between AcMNPV genome in pBlueBAC4.5 transfer vector</td>
<td>79</td>
</tr>
<tr>
<td>3.3</td>
<td>Immunoblot autoradiograph showing expression of CD36 from baculovirus clones</td>
<td>80</td>
</tr>
<tr>
<td>3.4</td>
<td>A time course assay to monitor infection and expression of CD36 in insect cells at x 5 M.O.I.</td>
<td>81</td>
</tr>
<tr>
<td>3.5</td>
<td>SDS polyacrylamide gels stained with colloidal blue and an immunoblot probed with anti-His showing protein and wild type</td>
<td>85</td>
</tr>
<tr>
<td>3.6</td>
<td>SDS polyacrylamide gel stained with colloidal blue, showing total protein CD36N8-10 from each fraction during purification</td>
<td>86</td>
</tr>
<tr>
<td>3.7</td>
<td>Immunoblot of Sf21 membrane preparations, starting and soluble fraction after incubation with detergent</td>
<td>88</td>
</tr>
<tr>
<td>3.8</td>
<td>SDS-PAGE gel stained with colloidal blue showing increasing BSA concentrations and purified wild type CD36-12His.</td>
<td>89</td>
</tr>
<tr>
<td>3.9</td>
<td>Typical BSA standard curve measured by the density of pixels from SDS-PAGE stained with colloidal blue.</td>
<td>90</td>
</tr>
</tbody>
</table>
Figure 3.10 – SDS polyacrylamide gel stained with colloidal blue and immunoblot of CD36 treated with PNGaseF. ................................................................. 91
Figure 3.11 - Interaction of wild type CD36 is with BODIPY Ac-LDL. ........................................ 93
Figure 4.1 - General components of mass spectrometer.......................................................... 98
Figure 4.2 - Electrospray quadrupole time of flight analyser............................................... 100
Figure 4.3 - Fourier transform ion cyclotron analyser......................................................... 101
Figure 4.4 - PNGaseF enzyme reaction and deglycosylation of CD36.................................... 103
Figure 4.5 - Following tryptic digestion, protein fragments are ionized into related y and b-series protonated ions................................................................. 108
Figure 4.6 - Q-ToF MS spectra showing that N1, N5, N7, N8, N9 and N10 of CD36 are glycosylated. ........................................................................................................... 109
Figure 4.7 – FT-ICR MS spectra showing that N6 of CD36 both occupied and unoccupied by N-glycan.................................................................................................................. 114
Figure 4.8 - Electrophoretic mobility shift analysis to test glycosylation status of CD36 probed with mAb1955. ........................................................................................................ 116
Figure 4.9 - Electrophoretic mobility shift analysis to test occupancy of N4 in wild type CD36 probed with mAb1955. ........................................................................................................ 117
Figure 4.9 – Primary sequence alignment of CD36 orthologues............................................. 120
Figure 5.1 - Schematic diagram of a FACScan flow cytometer (BD Biosciences).................... 126
Figure 5.2 - Flow cytometric analysis of cells........................................................................ 129
Figure 5.3 - Titration of mAb1955 and mAb1258 binding to HEK293T cells expressing wild type CD36. ........................................................................................................ 131
Figure 5.4 - Titration of secondary antibody binding to HEK293T cells expressing wild type CD36. ........................................................................................................ 133
Figure 5.5 - Analysis of HEK293T cells transfected with pCD36nong. ................................. 136
Figure 5.6 - Flow cytometric analysis of HEK293T cells transiently expressing CD36N7-10 and CD36N1-7.

Figure 5.7 Analysis of HEK293T cells transiently expressing CD36N1-6, CD36N2-7, CD36N1-5 and CD36N3-7.

Figure 5.8 - Analysis of HEK293T cells transiently expressing CD36N1,2,6,7.

Figure 5.9 - Analysis of CD36 carboxy-terminal glycosylation sites.

Figure 5.10 - Flow cytometric analysis of HEK293T cells describing the influence of the carboxy-terminal sites N8, N9 and N10 on the trafficking of CD36N1-7.

Figure 5.11 - Flow cytometric analysis of HEK293T cells expressing CD36min-g into which amino-terminal glycosylation sites have been reintroduced.

Figure 5.12 - Flow cytometric analysis of HEK293T cells expressing wild type CD36 and proteins with sites N2 and N4 removed.

Figure 5.13 - Analysis of HEK293T cells expressing CD36 isoforms and Pgp in the presence and absence of tunicamycin treatment.

Figure 5.14 – Glycopeptide structures that are sensitive to EndoH.

Figure 5.15 - Electrophoretic mobility of wild type CD36 and CD36N7,9,10 before and after treatment with 1U EndoH or 1U PNGaseF.

Figure 5.16 - Electrophoretic mobility of various cell surface expressing CD36 isoforms before and after treatment with 1U EndoH or 1U PNGaseF.

Figure 5.17 - Interaction of wild type CD36-12His expressed on the surface of HEK293T mammalian cells with Ac-LDL.

Figure 5.18 - Interaction of CD36 mutants expressed on the surface of HEK293T mammalian cells with Ac-LDL.
List of Tables

Table 1.1 – Most common mutations in introns and exons of CD36 .........................31
Table 2.1 – Oligonucleotides for site-directed mutagenesis of CD36 ..........................51
Table 3.1 - A List of detergents with the cmc at room temperature and pH7.0 .............74
Table 4.1 – Complete tryptic digestion of histidine-tagged wild type CD36 results in 46
fragments ........................................................................................................................................106
Table 4.2 - The conservation of human putative N-linked glycosylation sites in members of
CD36 family including mouse, rat, hamster and bovine .........................................................119
Table 5.1 - The cell surface expression of mutant CD36 compared to wild type, measured by
flow cytometry ..........................................................................................................................137
Acknowledgments

I would like to thank my supervisor Kenny Linton, who has supported and advised me throughout my PhD, although allowing me to win more often at table tennis would have been appreciated. Thanks to all past and present members of Membrane Transport Proteins and Disease Group, in particular Catherine, who was a great help and kept my spirits lifted when the project was not going as well as hoped.

Finally, I would like to thank my family and friends for their support, in particular my Mum and Dad, who encouraged me to take the first step toward studying again after a few years break. Finally, to Joe who despite being the other side of the world for the past year, has given me motivation and support, and has been my rock during the last four years.
1. Introduction

1.1. Identification of Class B scavenger receptor CD36

Macrophage scavenger receptors, which were first identified by Brown and Goldstein, are membrane proteins that mediate endocytosis of chemically modified low density lipoproteins (LDL) such as the natural and clinically important oxidized LDL (ox-LDL) (Gerrity, 1981) and the synthetic acetylated LDL (ac-LDL) (Goldstein et al., 1979). The first scavenger receptors isolated and cloned were type I and type II (Via et al., 1985), which are generated by alternative splicing of a single gene and designated the class A scavenger receptors (Emi et al., 1993; Freeman et al., 1990). These receptors each have six domains, five of which are identical (the carboxy-terminal being different) with the collagenuous extracellular domain allowing the binding of polyanionic ligands (Acton et al., 1993; Doi et al., 1993).

CD36 was first identified on the surface of monocytes by the monoclonal antibody OKM5 in 1984 (Talle et al., 1983), but it was not identified as a scavenger receptor until 1991 when it was discovered that CD36 expressed on the surface of macrophage facilitates the recognition and uptake of apoptotic cells (Savill 1991). In addition, using a mouse macrophage cDNA expression library, transfected COS-7 cells incubated with fluorescent ox-LDL identified CD36 as a receptor for modified lipoproteins (Endemann et al., 1993). Expression of CD36 (initially known as glycoprotein IV) had also been demonstrated on the surface of platelets (Asch et al., 1987) and endothelial cells (Swerlick et al., 1992), and was already predicted to be a receptor for thrombospondin (Asch et al., 1987), collagen (Tandon et al., 1989) and mediate the cytoadherence of erythrocytes infected with Plasmodium falciparum (Oquendo et al., 1989). CD36 was classified as a class B scavenger receptor when hamster SR-BI was cloned and identified as a new member of the CD36 family (Acton et al. 1994). The ligands
of CD36 and SR-BI, as shown by direct binding and competition assays, were similar and included modified lipoproteins but not other polyanionics, which are ligands for class A scavenger receptors. Due to this observation, SR-BI and CD36 were designated as the new class B scavenger receptors (Acton et al., 1994). Other members of the CD36 family have since been identified and include the human membrane glycoproteins, lysosomal integral membrane protein II (LIMPII) (Vega et al., 1991), and CD36 and LIMPIIA Analogous-1 (CLA-1), and Drosophila proteins epithelial membrane protein (emp) and Croquemort.

Rat LIMPII is one of the few characterized lysosomal proteins, and northern blot analysis has confirmed its wide spread cellular distribution. It is 74kDa in size and the primary amino acid sequence shares a 32% identity to CD36. The difference is especially evident in the carboxy-terminal cytoplasmic tail which is considerably longer in LIMPII and may be used to determine the sub-cellular localisation of the proteins (Calvo et al., 1995).

CLA-1 is a 88kDa human protein that also shares a 32% primary amino acid identity to human CD36. CLA-1 is mainly located on the plasma membrane of cells involved in storage, secretion or steroid hormone synthesis (ovary, testis and placenta) (Calvo et al., 1995). The ligands are similar to CD36 including high-density lipoprotein (HDL), LDL, very low-density lipoprotein (VLDL) and modified lipoprotein. CLA-1 is the human homologue of SR-BI, the main difference being that SR-BI is not expressed in the hamster placenta.

The non-mammalian homologues identified include the Drosophila proteins emp (Hart and Wilcox, 1993) and Croquemort (Franc et al., 1996), and sensory neuron membrane protein (SNMP) also found in drosophila, but originally identified in Lepidoptera (Nichols and Vogt, 2008). Croquemort has 23% amino acid identity to CD36 and is expressed on macrophage and hemocytes and is essential for phagocytosis of apoptotic corpses (Franc et al., 1996).

Emp has 32% identity to the amino acid sequence of CD36 and is expressed in precursor cells for adult epidermal structures (Hart and Wilcox, 1993). SNMPs are membrane bound proteins thought to play a role in odour detection and many candidate homologues have been
identified in different species (Nichols and Vogt, 2008). To date no members of the CD36 family have been identified in prokaryotes, but the study of the function of these CD36 homologues may provide insights into the evolution of CD36 and elucidate how the protein acquired its broad ligand specificity (see below).

1.2. Structure of CD36

One feature that is thought to distinguish a class B scavenger receptor from other scavenger receptors is the two hydrophobic putative transmembrane domains, which are separated by a large hydrophilic extracellular region (Figure 1.1). This was demonstrated in LIMPII where both hydrophobic domains serve to anchor the protein to the lysosomal membrane (Vega et al., 1991). Since the primary amino acid sequence of CD36 contains two hydrophobic regions adjacent to the carboxy and amino termini (Oquendo et al., 1989), the ditopic topological model was proposed with two small cytoplasmic tails at the amino-terminus (GCDRNC-) and the carboxy-terminus (-CACRSKTIK). The identification of palmitoylation of the four amino and carboxy-terminal cysteines of CD36 by Tao et al. agreed with this proposed topology (Tao et al., 1996) (see section 1.2.3.4). Palmitoylation of proteins occur in a variety of positions along the protein backbone, but generally the residues are situated on the cytoplasmic side of the membrane, supporting the predicted two cytoplasmic tails of CD36. Although this topology is generally agreed upon, it is not universally accepted, and another model has been proposed with one transmembrane domain. Deletion of the carboxy-terminal putative transmembrane domain by Pearce et al. resulted in secretion of CD36 rather than membrane anchoring, leading to the hypothesis that only the carboxy-terminal is transmembrane and the amino-terminal, an uncleaved signal peptide, is extracellular (Pearce et al., 1994). To resolve this controversy Gruarin et al. reinvestigated the two topological
models of CD36. In this later study, truncation of the carboxy-terminal domain did not result in CD36 being secreted; this required deletion of both the carboxy and amino-terminal domains. In addition, a FLAG epitope which was attached to the amino-terminus of wild type CD36 could not be detected by an anti-FLAG antibody in transiently transfected COS-7 cells unless the cells were permeabilized (Gruarin et al., 2000). Palmitoylation was also confirmed in COS-7 cells, suggesting that the modification is not cell specific and consistent with the proposed ditopic topology of CD36.
Figure 1.1 – Cartoon representation of the predicted topology of human CD36.

The proposed topology is ditopic, with the small cytoplasmic tails palmitoylated. The extracellular loop has 10 putative N-linked glycosylation sites shown as dark blue hexagons, a phosphorylation site shown as a dark green pentagon and mapped ligand binding sites in pale blue, green and black. Disulphide bonds between the extracellular cysteines are shown based upon the pattern observation in bovine CD36.

1.2.1. Extracellular domain of CD36

Human CD36 consists of 471 amino acids and has a predicted weight of 53kDa (Oquendo et al., 1989). The majority of the protein forms a large extracellular hydrophilic loop, predicted
to be 46kDa. The transmembrane domains and small cytoplasmic residues make up a small percentage of total protein (Figure 1.1). The extracellular domain is heavily modified post-translationally by glycosylation, phosphorylation and disulphide bridging. It is also the site of ligand binding.

1.2.2. Ligands of CD36 and localisation of binding sites

Many different ligands are now known to bind to CD36, and Pearce et al identified both ox-LDL and thrombospondin-1 (TSP-1) binding domains on the extracellular loop of the protein using Glutathione S-Transferase/CD36 fusion proteins (Frieda et al., 1995; Pearce et al., 1998). TSP-1 is a large adhesive glycoprotein that is involved in a number of important cellular processes such as tumourogenesis, development and vascular biology (Bornstein, 1992). Much of TSP-1 function occurs on the cell surface by interaction with receptors including CD36. GST/CD36 fusion proteins were generated for overlapping fragments of the entire extracellular domain and incubated with radiolabelled TSP-1. The binding of TSP-1 to four fusion proteins, each with a twenty-seven amino acid sequence in common, identified amino acids 93-120 as the likely CD36 binding site for TSP-1, however this binding site may be modulated by a downstream sequence of amino acids 139-155 (Pearce et al., 1994). The discovery that LIMPII also binds TSP-1 led to identification of a conserved TSP-1 binding sequence found in the CD36 family, called CD36 LIMPII Emp sequence homology (CLESH). It is actually composed of three conserved motifs, a protein kinase C phosphorylation sequence and two TSP-1 binding motif blocks (Crombie and Silverstein, 1998). Interestingly, other non-CD36 related proteins have also been identified with the CLESHE domain and also bind TSP-1.
The binding site for ox-LDL was identified by using the same method but incubating the fusion proteins with radiolabelled ox-LDL. This showed that the ox-LDL binding site is not the same as that for TSP-1 and has been mapped to amino acids 120-155 with a second region with less affinity between amino acids 28-93 (Pearce et al., 1998).

*Plasmodium falciparum* is one of the four species of *Plasmodium* that naturally infects humans, and is responsible for most malarial disease and almost all ensuing mortality. An important survival mechanism of *falciparum* malaria is the sequestration of *P. falciparum* infected erythrocytes (PE) in the microvasculature of vital organs (Baruch et al., 1999). A number of receptors have been implicated in the cytoadherence of PE to endothelial cells, including CD36. The binding of PEs to endothelial cells expressing CD36 can be inhibited by various CD36 antibodies that recognise the sequence between amino acids 155-183 (Daviet et al., 1995; Daviet et al., 1997). This suggested tentatively a binding domain for PE, however, some of these antibodies also inhibit the binding of TSP-1. Since TSP-1 and the antibodies have different CD36 binding domains there is a possibility that the antibodies act allosterically causing a conformational change in the PE binding site. Using CD36 peptides to block the binding of PE to CD36, the sequence between amino acids 139-184 was identified as the likely PE binding site on human CD36, in particular the sequences between the amino acids 145 and 171, and 146 to 164 (Baruch et al., 1999). *Plasmodium falciparum* erythrocyte membrane protein 1 (Pfemp1) has since been identified on the surface of PEs as the ligand for CD36 (Yipp et al., 2003).

Other important ligands that bind to the extracellular loop of CD36 include anionic phospholipids and long chain fatty-acids. The binding of anionic phospholipids (which also bind to SR-BI) was determined by the incubation of radiolabelled liposomes containing different phospholipids, to COS cells transiently expressing CD36 or SR-BI.

Phosphatidylserine and phosphatidylinositol were identified as the ligands (Rigotti et al., 1995). The recognition of these phospholipids is reminiscent of the scavenger activity of
CD36 and links CD36 and SR-BI to the function of Croquemort (known as the “death catcher” which probably recognises apoptotic corpses via the anionic phospholipids within the outer leaflet of the membrane (Franc et al., 1996)).

Abumrad et al identified an 88kDa membrane protein on the surface of rat adipocytes that when labelled with a derivative of long chain fatty-acids inhibits long chain fatty-acid transport into adipocytes (Abumrad et al., 1993). The protein was designated the fatty acid translocase (FAT) but amino-terminal sequencing of the protein showed it to be very similar to human CD36. A synthetic oligonucleotide derived from this amino-terminal sequence was used to isolate the cDNA from the rat adipocyte library and has since been identified as the rat homologue of human CD36 (Abumrad et al., 1993).

Ligands more recently identified for CD36 include fibrillar β-amyloid (here, CD36 is a co-receptor with Toll Like Receptor 2 (TLR-2)) which is a major constituent of Alzheimer plaques (Bamberger et al., 2003), the thyroid hormone T₃ (van der Putten et al., 2003) and hexarelin, which is a growth hormone protein, and its derivative EP80317, which has no growth hormone releasing properties (Bodart et al., 2002). This ligand may interfere with the binding of ox-LDL to CD36 possibly having a protective effect against atherosclerosis (Demers et al., 2004). CD36 also interacts with advanced glycation end products (AGE) (Ohgami et al., 2001) which are generated during inflammation and diabetes (Basta et al., 2004). These are protein adducts that may accumulate resulting in motifs recognised by CD36.
1.2.3. Post-translational modifications

1.2.3.1. Phosphorylation

Phosphorylation and dephosphorylation are known to regulate a number of physiological processes, and kinase consensus sequences are present on many extracellular and cell surface proteins (Ehrlich et al., 1990). One phosphorylation site has been identified on human CD36 by Asch et al. CD36 domains were randomly expressed in a recombinant expression system, and the library was screened using a peptide (CSVTCG) corresponding to the TSP-1 sequence that binds to CD36. The sequence identified in CD36 that binds to TSP-1 (RGPYTYRVRFLA) is also a protein kinase C (PKC) consensus substrate sequence. When the CD36 peptide was incubated with PKC and phosphorylated, the binding to the TSP-1 peptide was inhibited by about 60%, suggesting TSP-1 binding may be regulated by phosphorylation of CD36. Site-directed mutagenesis confirmed that threonine 92 was the site of phosphorylation, and the model proposed is that CD36 is constitutively phosphorylated, but can be dephosphorylated extracellularly leading to increases TSP-1 binding. This can then be reversed by PKC mediated phosphorylation (Asch et al., 1993). PKC, however, is an intracellular kinase, so the PKC-like kinase responsible for phosphorylation of CD36 has yet to be identified. This phosphorylation site is also important for the regulation of cytoadherence of Plasmodium falciparum-infected erythrocytes to the microvascular endothelium. CD36 expressed endogenously on the surface of human dermal microvascular endothelial cells (HDMECs) supports the adhesion of PEs and this is regulated by the Src family of kinases and alkaline phosphatase (AP). The ectodomain of CD36 is constitutively phosphorylated and is recognised by a phospho-specific antibody, and although PE can bind to phosphorylated CD36, optimal binding is achieved following dephosphorylation by AP (Ho et al., 2005).
1.2.3.2. N-linked glycosylation

Different Mr values have been reported for CD36 from different cell types, for example, in human platelets (88kDa), mammary epithelial cells (85kDa) and erythroblasts (78kDa) (Greenwalt et al., 1990; Kieffer et al., 1989; Tandon et al., 1989). The primary amino acid sequences of CD36 in these cell types are identical, therefore it is likely that these size differences are caused by cell specific post-translational modifications, particularly N-linked glycosylation.

Human CD36 has ten putative N-linked glycosylation sites (recognized by the consensus sequence Asn-X-Ser/Thr, where X is any amino acid except proline), and by comparing the predicted size of the peptide backbone (53kDa) to the aforementioned observed sizes, some, if not all, of these sites are occupied with oligosaccharides. In CD36 purified from platelets both alkali-labile O-glycosidic linkages (Tandon et al., 1989) and N-glycosidic linkages have been found (Nakata et al., 1993), although occupancy of specific sites has yet to be identified. It is likely that N-linked glycosylation is responsible for the protease resistance first observed with CD36 expressed on the surface of mammary epithelial cells (Kobylka and Carraway, 1973), resistant membrane-bound CD36 becomes sensitive to proteases after treatment with the endoglycosidase PNGaseF, which cleaves N-linked glycans from the protein. Also, when cells expressing LIMPII are treated with tunicamycin, thereby preventing glycosylation, the protein only has a 5% half life of a fully glycosylated protein, consistent with a protective role for the added glycans (Barriocanal et al., 1986). Since CD36 and LIMPII are likely to be exposed to proteases in the blood and in the lysosome, respectively, this may be the reason why both proteins are heavily glycosylated.

Some of the putative N-linked glycosylation sites in human CD36 are shared between CD36 homologues. LIMPII has in total eleven N-linked glycosylation sites, three of which are
conserved within human CD36 (N205, N249 and N417 using human CD36 nomenclature) and CLA-1 has ten N-linked glycosylation sites, four of which are conserved (N102, N205, N220 and N249 using hCD36 nomenclature). Rat Cd36 and bovine Cd36 (which have 85% and 82% amino acid identity to human CD36 respectively) also share glycosylation sites. Rat Cd36 has nine putative N-linked glycosylation sites, eight of which are conserved in the human protein and bovine Cd36 has eight putative N-linked glycosylation sites, seven of which are conserved in human CD36 (see Figure 4.9). The structure of bovine Cd36 has been studied and characterized using purified protein from milk fat globule membrane (MFGM). Berglund et al; discovered that all eight of the putative N-linked glycosylation sites were occupied by fragmenting the purified protein and analysing the amino sugars in the fractions that had been separated by reverse phase separation. The carbohydrate composition of the glycans was also studied by high pH anion-exchange chromatography. The glycans showed structural heterogeneity, for example the glycans belonging to Asn101 were probably complex oligosaccharides as they had low mannose content, but the glycans associated with Asn205, Asn247 and Asn 417 had high mannose content so maybe sensitive to Endoglycosidase H (Berglund et al., 1996).

Further analysis of the occupancy of the putative N-linked glycosylation in human CD36 is required with a view to understanding the role of these sugars in relation to the function of the protein. Further analysis of biosynthesis of the N-linked glycosylation is discussed in section 1.5.

1.2.3.3. Disulphide bonds

Disulphide bonds are formed by the cross-linking of cysteine residues in the protein. They are highly conserved features which play a role in stabilization, folding and structure of proteins.
Human CD36 has ten cysteines in total, however, four of these are intracellular and are palmitoylated (Tao et al., 1996). The remaining six residues (C243, 272, 311, 313, 322 and 333) are all situated in the carboxy-terminal half of the extracellular loop of the protein. These cysteines are highly conserved within the CD36 family, only C311 is not conserved in SR-BI, CLA-1, LIMPII and emp and C313 is not conserved in emp (Figure 1.2).

The electrophoretic mobility of native CD36 under reducing and non-reducing conditions is very similar, which led to the assumption that the protein does not have any intrachain or interchain disulphide bonds (Oquendo et al., 1989). This is in contrast with the observation that most membrane and secreted proteins have disulphide bonds, which are formed in the ER (Hwang et al., 1992), where chaperones including disulphide isomerases create an optimal environment for the folding and trafficking of proteins (Helenius, 1994).

In disagreement with previous studies Gruarin et al, found that disulphide bonds do form in human CD36, and the reason that the change in electrophoretic mobility cannot be visualised under reducing conditions is due to extensive glycosylation. If the protein is previously deglycosylated using PNGaseF, the change in electrophoretic mobility under reducing conditions can be observed. Furthermore, in the presence of the reducing agent DTT, newly formed CD36 is retained and degraded in the ER suggesting that disulphide bonds are required for CD36 to mature and traffic to the plasma membrane (Gruarin et al., 1997).

The pattern of disulphide bonding has been elucidated in bovine Cd36. The protein was purified from MFGM and no free cysteines were identified (by the lack of incorporation of $^{14}$C labelled iodoacetic acid). Following CnBr degradation of Cd36, the resulting fragments were separated and analysed by amino acid sequencing and mass spectrometry, with the expectation that the disulphide bonds would be still be intact. This method identified the following bonding pattern C243-C311, C272-C333 and C313-C322 (Rasmussen et al., 1998).
Due to the conservation of cysteines in human CD36, it is likely that the pattern of disulphide bonding is also conserved.

Figure 1.2 – Alignment of CD36 homologues demonstrating extracellular cysteine conservation.

The conserved cysteines are highlighted in yellow, the non-conserved substitutions highlighted in red, and non-conserved cysteines highlighted in green. The loop formed by disulphide bond between C4 and C5 in bovine CD36 may not be present in CLA-1 or SR-BI due to the close proximity of these residues. Emp may only have one conserved disulphide bond between C2 and C6.
1.2.3.4. Palmitoylation

Palmitoylation is a post-translational attachment of the sixteen carbon saturated fatty-acid palmitate through thioester linkage to a cysteine (although there is no known consensus sequence). Site-directed mutagenesis on the ten cysteines in CD36, followed by incubation with radio-labelled palmitate, identified the two most carboxy-terminal and amino-terminal cysteines as palmitoylation sites (C3, C7, C463 and C466) (Tao et al., 1996).

Palmitoylation is thought to be involved in targeting the protein to the membrane, however, site-directed mutagenesis of the cysteines in the carboxy-terminus demonstrated that modification of these residues is unnecessary for plasma membrane localisation (Malaud et al., 2002). Investigation into the role of palmitoylation in lipid raft targeting of CD36 and binding of ac-LDL has been carried out in our lab (Wharton and Linton, personal communication). Two CD36 palmitoylation mutants were generated by site-directed mutagenesis, in which the carboxy-terminal and amino-terminal pairs of cysteines were mutated to serine. These mutants were combined to generate the quadruple mutants devoid of the target cysteines. Flow cytometric analysis confirmed that palmitoylation is not required for cell surface expression or binding of ac-LDL. However, there did appear to be a reduction in CD36 localisation to lipid rafts in the membrane, and uptake of ac-LDL was markedly reduced in the mutant without palmitoylation. This suggests that although palmitoylation is not required for trafficking to the cell membrane, it is required for the localisation to lipid rafts in the membrane which appears to be important for the uptake of ac-LDL. It is hypothesised that expression of CD36 in lipid rafts may allow the co-localisation of different proteins necessary for receptor-mediated ox-LDL internalisation.

Palmitoylation is also known to be reversible, which may prove to be important in the translocation of CD36 from intracellular stores to the plasma membrane (see below).
1.3. Expression and Regulation of CD36

1.3.1. Regulation of CD36

CD36 is 28kb long and is located on chromosome 7 q11.2 (Fernandez-Ruiz et al., 1993). The gene has 15 exons, generally exons 1, 2 and 15 are non-coding and exons 3 and 14 encode the amino and carboxy-terminal of the protein respectively (Armesilla and Vega, 1994). However, alternative splicing of mRNA transcripts is known to occur, for example exons 4 and 5 can be skipped resulting in a protein that lacks amino acid 43 to 143. This encompasses some of the ligand binding domain and three N-linked glycosylation sites in the extracellular loop, and interestingly similar exon skipping occurs in CLA-1 (Tang et al., 1994). There is also alternative splicing in exon 1 that lacks TATA boxes and CpG islands, leading to different transcripts which are regulated tissue specifically (Andersen et al., 2006). Overall, it appears that the molecular mechanisms that regulate CD36 are complex which possibly reflects the multifunctional role of the protein. The protein can also be regulated post-translationally by phosphorylation and N-linked glycosylation. These again are probably tissue and/or ligand specific and are described in more details in section 1.2.3.1 and 1.2.3.2.

Regulation at the level of trafficking of CD36 from intracellular pools to the plasma membrane has also been described. This was demonstrated in skeletal muscle following different stimuli such as muscle contraction and insulin release and identification of the different signalling pathways involved may led to better understanding of CD36 regulation. Candidate kinases are protein kinase A, adenosine monophosphate activated protein-kinase, protein kinase C and mitogen-activated protein kinase (Koonen et al., 2004; Luiken et al., 2002).
Upregulation of CD36 expression has been demonstrated in response to different CD36 ligands. For example, CD36 is upregulated by ox-LDL via a signalling pathway involving indirect dependence on PPAR (Sato et al., 2002).

1.3.2. Naturally occurring CD36 mutations and variants

Mutations in the coding sequence of CD36 have been identified that range from small and gross deletions, insertions, rearrangements, variant repeats and duplications (Rac et al., 2007) (Table 1). Two types of CD36 deficiencies have been identified from mutations, Type I and Type II. Type I has no CD36 expression on the surface of monocytes or platelets and Type II has no CD36 expression on the surface of platelets. There are over twenty types of mutations that lead to Type I deficiency (Rac et al., 2007) and most of the mutations are homozygous or compound heterozygous. One of the most common is C268T (Kashiwagi et al., 2001). This results in a 81kDa protein, possibly with a N-linked glycosylation impairment and the protein is degraded (Kashiwagi et al., 1993; Kashiwagi et al., 1995). Observed consequences of Type I CD36 deficiency include abnormal glucose and lipid metabolism; the involvement of CD36 in metabolic processes is described further in section 1.4.6.

The molecular basis for type II deficiency is largely unknown, but it has been observed in individuals who have a heterozygous mutation in CD36, most notably C268T. Interestingly, Kashiwagi et al., found both C268 and T268 forms of CD36 in monocytes from two patients with Type II deficiency, but only T268 in platelets from these same individuals. This suggests that other platelet-specific factors contribute to this phenotype (Kashiwagi et al., 1995).

A high degree of variability has been observed in CD36 particularly in individuals from African and Asian descent, and the naturally occurring polymorphisms have been associated with different diseases. These include five intronic SNPs in African-American population
which show a significantly increase in the odds for the development of metabolic syndrome (Love-Gregory et al., 2008), another five SNPs associated with increased cardiovascular risk (see section 1.4.2) and three polymorphisms in Thai patients with \textit{Plasmodium falciparum} malaria which are associated with protection against cerebral malaria, 14T-C, 53G-T and 3TG$_{12}$ (Omi et al., 2003).
<table>
<thead>
<tr>
<th>Exon number</th>
<th>Amino acids encoded</th>
<th>Change in nucleotide sequence</th>
<th>Change in amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None</td>
<td>Deletion exons 1-3</td>
<td>No expression CD36</td>
</tr>
<tr>
<td>2</td>
<td>None</td>
<td>Deletion exons 1-3</td>
<td>No expression CD36</td>
</tr>
<tr>
<td>3</td>
<td>1-40</td>
<td>Deletion exons 1-3</td>
<td>No expression CD36</td>
</tr>
<tr>
<td>4</td>
<td>41-94</td>
<td>C268T</td>
<td>Pro90Ser</td>
</tr>
<tr>
<td>5</td>
<td>94-143</td>
<td>319-324 del</td>
<td>Inframe deletion 107-108</td>
</tr>
<tr>
<td></td>
<td></td>
<td>329-330 del</td>
<td>Frameshift at 110</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G367A</td>
<td>Glu123Lys</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C380T</td>
<td>Ser127Lue</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T411C</td>
<td>Ala137Val</td>
</tr>
<tr>
<td>6</td>
<td>144-203</td>
<td>560 ins T</td>
<td>Frameshift at 187</td>
</tr>
<tr>
<td>7</td>
<td>204-234</td>
<td>619-624 del</td>
<td>Frameshift at 207</td>
</tr>
<tr>
<td></td>
<td></td>
<td>691-696 del</td>
<td>Inframe del 231-232</td>
</tr>
<tr>
<td>9</td>
<td>250-273</td>
<td>T760C</td>
<td>Phe254Lue</td>
</tr>
<tr>
<td>10</td>
<td>273-336</td>
<td>845-849 del</td>
<td>Frameshift at 282</td>
</tr>
<tr>
<td></td>
<td></td>
<td>949 ins</td>
<td>Frameshift at 317</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T975G</td>
<td>Tyr325Term</td>
</tr>
<tr>
<td>11</td>
<td>336-375</td>
<td>T1079G</td>
<td>Leu360Term</td>
</tr>
<tr>
<td>12</td>
<td>376-400</td>
<td>Del tttagAT</td>
<td>Skipping exon 12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1140-1146 del</td>
<td>Frameshift at 380</td>
</tr>
<tr>
<td></td>
<td></td>
<td>G1150C</td>
<td>Ala384Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1155del</td>
<td>Frameshift at 385</td>
</tr>
<tr>
<td>13</td>
<td>400-418</td>
<td>Del tattacagAG</td>
<td>Skipping exon 13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dupl. 1204-1246</td>
<td>Frameshift at 416</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1218-1224 del</td>
<td>Frameshift at 406</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1228-1239 del</td>
<td>Del Ile-Val-Pro-Ile</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A1237C</td>
<td>Ile413Leu</td>
</tr>
</tbody>
</table>

**Table 1.1 – Most common mutations in introns and exons of CD36.**

*Introns are indicated in lower case. Modified from Rac et al., (2007)*
1.4. CD36 in normal and disease pathology

1.4.1. Innate Immunity

The first identification of CD36 as a scavenger receptor was reported by Savill et al.; who was investigating the uptake of apoptotic cells by mononuclear phagocytes (Savill et al., 1991). A BLAST search confirmed the presence of CD36 homologues from primitive to more immunologically advanced organisms, and by comparing the functions of the proteins it was suggested that the uptake of apoptotic cells during normal homeostasis is the most ancient function of CD36 (Febbraio and Silverstein, 2007). When a cell is apoptotic, membrane changes occur and the outer leaflet becomes enriched in anionic phospholipids, such as phosphatidylserine (Pittoni and Valesini, 2002). Phosphatidylserine in apoptotic cells can undergo oxidative changes which are also recognised by CD36, so it maybe the oxidated species that trigger clearance by macrophage via CD36 (Greenberg et al., 2006). An example of routine apoptotic cell clearance is the removal of shed photo-receptor rod outer segments (ROS). CD36 expression on the surface of retinal pigment epithelial cells (RPEs) (recognised by immunofluorescence) mediates the binding and uptake of ROS which are shed on a daily basis and need to be cleared to maintain normal vision. Use of a CD36 antibody inhibits the uptake of ROS by RPEs by 20% suggesting there may be a co-receptor. This could possibly be integrin αυβ5 as inhibition of both CD36 and integrin αυβ5 reduced the uptake of ROS by 80% (Ryeom et al., 1996).

In addition to apoptotic cell recognition, CD36 may also recognise foreign pathogens. For example, CD36 KO mice are less efficient at phagocytosis of Gram positive Staphylococcus aureus leading to a higher incidence of abscesses and death (Stuart et al., 2005), but the specific ligand on the bacterium that enables recognition and uptake is unclear.
In addition to this, CD36 is also known to recognise modified lipoproteins which are by-products of the inflammatory response, however, due to our modern western diet, heavy in lipids, this function is inadequate and can lead to complications such as atherosclerosis.

1.4.2. CD36 and atherosclerosis

Formation of atherosclerotic lesions is a highly complex inflammatory process. During atherogenesis lipid laden macrophage called foam-cells accumulate within the arterial neo-intima and become a major contributor to plaque formation. Although ox-LDL was identified as a ligand for CD36 by Endemann \textit{et al}., its role in atherogenesis was unclear before the development of mouse models (Endemann \textit{et al}., 1993). Now there are several lines of \textit{in vitro} and \textit{in vivo} evidence that implicate the binding and endocytosis of ox-LDL in atherosclerotic lesion formation. In particular, the development and characterization of the ApoE/CD36 double knockout mouse has been informative. ApoE single knockout mice have very high circulating cholesterol levels and develop atherosclerotic lesions more rapidly than wild type mice (van Ree \textit{et al}., 1994). When fed a western diet, ApoE/CD36 double knockout mice showed a significant reduction (76.5\%) of aortic tree lesions at 12 weeks old when compared to single ApoE knockout, and macrophages, isolated from these mice also displayed a 60\% reduction in internalization of ox-LDL (Febbraio \textit{et al}., 2000). This phenotype could be reversed when CD36 was reintroduced using bone marrow transfer approach (Febbraio \textit{et al}., 2004). The reduced lesion size in ApoE/CD36 knockout mice was not only observed in young animals but also in 35 week old mice (Febbraio \textit{et al}., 2000). In addition, the CD36 ligand EP80317, a hexarelin derivative was shown to reduce lesion size in ApoE single knockout mice, probably by competing with the binding site of modified lipoproteins (Marleau \textit{et al}., 2005).
However, the role of CD36 in atherosclerosis is not without controversy with results from a study conducted by Moore et al., appearing to contradict Febbraio’s conclusion, that the loss of CD36 in an ApoE null background can reduce the atherosclerotic lesion size. In the more recent publication, ApoE/CD36 double knockout mice were fed a western diet for 8 weeks, and the size of the lesion were determined over the aortic sinus and tree. This time the results differed according to the sex of the animal, reporting that male mice showed no significant reduction in lesion sizes and female mice showing reduced lesion area through the aortic tree but increased lesion area in the aortic sinus. This led to the conclusion that the absence of CD36 was pro-atherogenic, and suggested that endocytosis of ox-LDL was independent of the scavenger receptor (Moore et al., 2005). These seemingly contradictory results have been re-examined with questions raised about the later study; for example the lesion size in the aortic sinus may have a non-linear relationship with the aortic tree lesion area (Curtiss, 2006), and since the aortic sinus undergoes remodelling, only the aortic tree lesions should be measured. This reappraisal makes the results from the two studies similar and consistent with a role for CD36 in atherogenesis. Slight variations in the congenic strains of mice used in the studies may also have resulted in altered function in macrophage and endothelial cells and finally, exposure to certain pathogens may increase the size of lesions by neutralizing the protective effect in the absence of CD36, which may be the reason for different lesion sizes (Burnett et al., 2001).

CD36 deficiency in macrophage (Type I deficiency) has been described in a small percentage of the Japanese population. This population display a 40-50% reduction in ox-LDL binding and uptake in macrophage, but whether this correlates with a reduced risk of atherosclerotic lesions is unclear (Febbraio et al., 2001). In addition CD36 polymorphisms have been associated with increased cardiovascular risk. Ma et al., selected five common haplotypes for analysis; 33137A-G, 31118G-A, 25444G-A, 27665del and 30294G-C, and one (33137A, 31118G, 35444G, 27665 no deletion, 30294G) was found that effected lipid and glucose
metabolism. The fasting free fatty-acid plasma levels were higher than normal (Ma et al., 2004) reminiscent of mouse Cd36 knockout phenotype (Febbraio et al., 1999), however, curiously, this was stronger in men than women. It is not clear why this is the case, but may be due to differences in deposition of body fat and the hormonal environment. Increased circulating fatty acids has been reported to be toxic to endothelial cells which may lead to the increase in cardiovascular risk, however, it is likely that additional factors contribute to the phenotype.

1.4.3. Role of CD36 in pathogenesis of Plasmodium falciparum malaria and sickle cell anaemia

A particularly severe complication of Plasmodium falciparum infection is the development of cerebral malaria (Miller et al., 1994) where parasitized erythrocytes adhere to the microvasculature of the brain. CD36 has been implicated as a receptor involved in the sequestration in the brain via its ligand Pfemp-1, or possibly due to the membrane changes that occur in PE’s (Ockenhouse et al., 1991). This is despite there being little evidence that CD36 is expressed in the cerebrum. The suggestion that the sequestration of PEs allows the parasite to avoid spleen-dependent immune intervention was tested by Aitman et al. The study found common frameshift or deletion mutations amongst the African population (nucleotides T1264G, G1439C and 1888delA) which all result in production of a truncated CD36 without the carboxy-terminal half of the protein. The frequency of these mutations was significantly higher in patients with severe cerebral malaria compared to the control group which, in disagreement with the previous hypothesis, suggested that CD36 deficient patients had more severe symptoms (Aitman et al., 2000). In contrast, a study investigating CD36 in Thai patients found that the nonsense mutation T188G protected the patient against severe
malaria (Pain et al., 2001). Both these studies suggest that the outcome of infection is not solely dependent on CD36-mediated parasite sequestration or immune response. However, given the mortality of *Plasmodium falciparum* infection the protective mutation should be more frequently observed. Why other non-protective mutations are maintained is unclear but maybe due to protection against other coincidental diseases. One candidate for such a disease is sickle-cell anaemia, and interestingly, patients with this disease are more resistant to malaria. Sickle-cell anaemia is prevalent in Africa and is caused by mutation in the haemoglobin gene. This results in changes in erythrocytes and slower transit of the “sickle”-shaped cells through the vascular system, promoting ligand/receptor recognition and complexes which contribute to the vascular pathology (Hebbel et al., 1980). Although CD36 expressed on microvascular endothelial cells is thought to be one such receptor, the ligands are less clear. These may include TSP-1 (Sugihara et al., 1992), and due to the membrane changes membranes are asymmetric and anionic phospholipids are exposed on the cell surface (Setty et al., 2002). There may therefore be a selective pressure to develop CD36 deficiency in sickle-cell anaemia (resulting in less severe pathology), which would suggest that malaria, sickle-cell anaemia and CD36-deficiency may have co-evolved.

### 1.4.4. CD36 and angiogenesis

The CD36/TSP-1 receptor/ligand complex has also been implicated in the antiangiogenesis of tumour cells and is effective against most angiogenic inducers such as bFGF and VEGF (Good et al., 1990). The suspected pathway involves CD36 physically associating with the tyrosine kinase fyn (which is localised to caveolae) after exposure to TSP-1. The kinase becomes activated and leads to downstream activation of MAPK p38, triggering a kinase cascade. Both fyn and p38 are both shown by the use of inhibitors to be essential in this
pathway, which ultimately leads to caspase DNA cleavage and apoptosis of the cell (Jimenez et al., 2000). Unravelling the pathway that contains the switch between pro-angiogenic and anti-angiogenic responses has the potential to lead to powerful therapeutic targets.

1.4.5. CD36 and fatty acid transport

When studying rat adipocytes, Abumrad et al., identified Cd36 as a fatty-acid translocase protein. This was a controversial finding as long-chain fatty acid (LCFA) were previously believed to diffuse across the plasma membrane, however the distribution of Cd36 favours such a transport function as it is expressed in tissues with high metabolic capabilities such as adipose tissue (Abumrad et al., 1993), heart muscle (Luiken et al., 1999a) and skeletal muscle (Bonen et al., 1999) and the kinetics of LCFA uptake appear saturatable. Consistent with this hypothesis, LCFA transport in Cd36 KO mice is dramatically reduced in heart and skeletal muscle (Coburn et al., 2000). When a perfused heart system from this mouse strain was isolated by Kuang et al, the heart was found to rely heavily on glucose as an energy source (Kuang et al., 2004). Additionally, transgenic mice which overexpress Cd36 in skeletal muscle show enhanced fatty acid oxidation, a decrease in circulating fatty acids and triglycerides, and decreased fat deposition in agreement with a role for Cd36 in LCFA uptake (Ibrahimi et al., 1999).

The mechanism for LCFA uptake via CD36 and other translocase proteins is unclear, however, it is thought that CD36 does not act like a standard transporter. CD36 (and possibly the plasma membrane fatty acid binding protein (FABPpm)) may assist in the disassociation of the fatty acid from albumin (Glatz et al., 2003) (Stremmel et al., 2001) and promote facilitated diffusion across the bilayer by mediating the integration of protonized LCFA into the outer leaflet. In doing so, a gradient could form across the membrane allowing flipping of
LCFA into the inner leaflet of the membrane where it may be activated by addition of coenzyme A (Stremmel et al., 2001). The transport proteins may also act together as a complex because inhibition of CD36 or FABPpm leads to 80% reduction in LCFA uptake, but inhibiting both proteins does not cause any additional reduction (Luiken et al., 1999b).

As the uptake of LCFA maybe a rate limiting step in fatty acid oxidation, the regulation of cell surface CD36 expression is important. This regulation can be as a consequence of acute stimuli such as muscle contraction or insulin release, or a longer term response. Elevated plasma LCFA and triacylglycerol concentrations have been associated with insulin resistant states, therefore this may link CD36 with metabolic syndromes (Shulman, 2000).

1.4.6. CD36, insulin resistance and metabolic syndrome

Insulin resistance is a feature of many complex human disorders that, in part may have a genetic basis. These include type 2 diabetes, obesity, combined hyperlipidaemia and essential hypertension (Aitman et al., 1997a; Aitman et al., 1997b; Groop et al., 1989; Reaven, 1988; Reaven et al., 1996). The spontaneous hypertensive rat (SHR) has been proposed as a model for these human disorders due to similar symptoms of insulin resistance, hypertriglyceridaemia, abdominal obesity and hypertension (Iritani et al., 1977; Reaven et al., 1989). Using a combination of quantitative trait loci mapping, cDNA microarrays, and radiation hybrid mapping, Aitman et al., identified that Cd36 was a defective gene in some SHR strains (Aitman et al., 1999). Replacement of the mutant allele with the wild type version allowed reversal of glucose intolerance and also improved the insulin responsiveness, however, had no effect on hypertension. This led to the conclusion that Cd36 deficiency underlies the insulin resistance in SHR (Pravenec et al., 2001). This is not an unreasonable assumption as CD36 was already identified as a possible transporter of LCFA, and some of
the features of insulin resistance syndromes could be attributed to the defects in fatty acid
metabolism (such as raised blood fatty acids)(Aitman et al., 1997a). However, further studies
into different SHR strains by Gotoda et al., showed that some of these strains did not have
mutations in Cd36, and suggested that Cd36 mutations arose de novo and did not necessarily
contribute to the symptoms presented (Gotoda et al., 1999).

Interestingly, Cd36 KO mice do not show insulin resistance, confirming that metabolic
syndromes are complex conditions (Febbraio et al., 1999). When compared to wild type
mice, the Cd36 KO mice showed a higher whole body glucose uptake and lower glucose
storage at basal conditions, and under hyperinsulinemic conditions the whole body glucose
and uptake was higher than in wild type (Goudriaan et al., 2003). This is consistent with the
previous study by Kuang et al., confirming the use of glucose as an alternative energy source
especially for the heart in absence of LCFA uptake (Kuang et al., 2004). However, under
these conditions the hepatic glucose production was not inhibited indicating some hepatic
insulin resistance (Goudriaan et al., 2003).

The differences in genetic background and nutritional intake may influence the results of
these studies, accounting for the differences between the mouse and rat models. Although
blood pressure was not measured in Cd36 knock-out mouse, the observation from transgenic
rescue in SHR suggests that Cd36 deficiency does not cause hypertension, but it is likely that
Cd36 does play a role in glucose tolerance and insulin sensitivity.

Studies into insulin resistance syndromes in humans has proved to be difficult with
inconclusive results have been reported. There are conflicting reports about whether the
Japanese population with type I CD36 deficiency show insulin resistance (Furuhashi et al.,
2003; Furuhashi et al., 2004; Kajihara et al., 2001; Kuwasako et al., 2003; Miyaoka et al.,
2001; Yanai et al., 2000), however, polymorphisms in CD36 have been associated with the
condition (Corpeleijn et al., 2006; Lepretre et al., 2004). Assuming that CD36 plays a role in
insulin resistance, the mechanism in which it affects insulin resistance remains undefined,
however given that LCFA uptake in muscle is clearly affected by CD36 deficiency, the may be a contributing factor (Hwang et al., 1998). Other influences such as diet, exercise and other genetic factors are also likely to be important.

1.5. **Biosynthesis of N-linked glycosylation**

N-linked glycosylation is the most common post-translational modification and occurs on the majority of secreted and membrane expressed proteins that transit through the ER. The importance of N-linked glycosylation is demonstrated by a group of inherited metabolic diseases caused by congenital disorders of glycosylation (CDG), of which there are eleven different disorders resulting in different clinical phenotypes (Marquardt and Denecke, 2003). Glycans are flexible hydrated branches of oligosaccharides (Wormald et al., 2002) and are semi-independent appendages. This suggests that they can be removed or modified without effecting the protein structure allowing the heterogeneity which characterizes glycoproteins (Rudd et al., 1999), with alternative glycans presented on the same protein derived from the same cell during different stages of differentiation, from different tissues types or from diseased tissues (Lowe and Marth, 2003).

During N-linked glycosylation there is a division of labour between the endoplasmic reticulum (ER) and the Golgi apparatus. The ER is responsible for the initial synthesis, addition and trimming of the core oligosaccharide to the growing nascent polypeptide chain, and at this stage the glycans are homogeneous. The folded proteins are then transported to the Golgi apparatus, where the oligosaccharides are trimmed further and then elongated generating complex heterogeneous glycosylation.
1.5.1. **N-linked glycosylation in the ER**

The process of attaching the 14 saccharide core unit to the N-linked glycosylation consensus sequence Asn-X-Ser/Thr begins with presynthesis of the oligosaccharide on a lipid carrier molecule, dolichol-pyrophosphate (DPP). In eukaryotes, this core glycan has a predefined structure of Glc$_3$Man$_9$GlcNAc$_2$ (Kornfeld and Kornfeld, 1985) (Figure 1.3).

![Diagram of core N-linked glycan](image)

**Figure 1.3 – The core N-linked glycan.**

This glycan is synthesised on a lipid carrier before being transferred to the polypeptide. The arrows indicate the sites of cleavage which occur as part of quality control in the ER. The core oligosaccharide contains two N-acetylglucosamines, nine mannoses and three glucoses. The inclusion of the fucose, galactose and sialic acid in the key is relevant for figure 1.5..

The saccharides are added to DPP by monosaccharide transferases (Burda and Aebi, 1999) in the ER membrane, and the first seven sugars are added on the cytoplasmic side of the membrane. The sugars are then “flipped” to the luminal side of the membrane by an ATP independent bi-directional flippase (Hirschberg and Snider, 1987). The remaining sugars are then added linearly to DPP (Figure 1.4), and the final glucose added is required for
oligosaccharyltransferase (OST) to recognise the oligosaccharide, and transfer it to the Asn of the nascent polypeptide (Burda and Aebi, 1998; Spiro, 2000). Once the glycans have been transferred to the polypeptide they are thought to play a role in folding of the protein either directly, by helping promote and stabilise local protein structure by contacting the polypeptide (Petrescu et al., 2004) or indirectly via the calnexin and calreticulin cycle. In this cycle, immediately after the oligosaccharide is transferred to the polypeptide the first glucose is cleaved by glucosidase I and then the second by glucosidase II (Hammond and Helenius, 1994) . The monoglucosylated glycan can then bind either to calnexin (a transmembrane ER protein) or calreticulin (a soluble ER protein), which facilitate the folding of the protein (Ou et al., 1993). Glucosidase II then cleaves the remaining glucose releasing the protein from the chaperone. If the protein is misfolded it is recognised by UDP-Glc:glycoprotein glucosyltransferase (GT), which reglucosylates the oligosaccharide and allows binding to the chaperone (Sousa et al., 1992). This cycle continues until the protein either folds properly or is targeted for degradation, and allows a quality control of proteins. Finally, before the proteins are exported to the Golgi apparatus, one mannose is cleaved by ERα-mannosidase.
Figure 1.4 – N-linked glycan processing in the ER.

The N-linked glycan is synthesised as a precursor initially in the cytosol before being flopped into the lumen. The oligosaccharide is transferred onto the asparagine within the consensus sequences of a growing polypeptide before being trimmed and entering the calnexin/calreticulin cycle. After removal of one mannose the correctly folded protein can be transported to the Golgi apparatus (Helenius and Aebi, 2001).

1.5.2. N-linked glycosylation in the Golgi apparatus

When proteins leave the ER they pass through the Golgi apparatus where they receive further saccharides before being trafficked to their final destination; but there are many differences between how the ER and Golgi apparatus operates. The Golgi apparatus elongates the existing core glycan in the lumen and makes use of soluble precursors such as nucleotide sugars (Figure 1.5)(Hirschberg et al., 1998).

Unlike the ER, the Golgi apparatus does not have a rigorous quality control process, and glycoproteins with defective moieties can readily be exported (Elbein, 1991; Stanley, 1984), although unfinished products cannot. The proteins are progressively modified as they move from cis to trans through the multicompartamental architecture of the Golgi and the process usually takes 5-15 minutes (Pelham and Rothman, 2000). This ensures a high degree of
completeness, but the lack of quality control allows a wide diversity of glycan structures, which is a characteristic of glycoproteins. Once the glycans have been added to the glycoprotein, it is transported to its final localisation.

Figure 1.5 – N-linked glycan processing in the Golgi.

The N-linked glycan undergoes further trimming and then elongation, with the addition of different sugars depending on the protein or the tissue. This represents one pathway with the addition of three N-acetylglucosamines, one fucose, three galactoses and three sialic acids (see Figure 1.3 for key (Helenius and Aebi, 2001).

1.6. Objectives of thesis

Although much is known about the function of CD36, solving the structure would assist in elucidating the structure function relationships within the protein.

The main objective of this project is to identify the role that the N-linked glycans play in folding, trafficking and binding of ligand (ac-LDL) to human CD36. In chapter 3 the purification procedure used to isolate the membrane expressed CD36 from Sf21 insect cell is described and the sensitivity to endoglycosidases for purified wild type and mutant CD36 isoforms is explored.
The occupancy of the ten putative N-linked glycosylation sites had to be determined by comparing the electrophoretic mobility of mammalian whole cell lysates expressing various CD36 mutants, and by mass spectrometry using CD36 purified from insect cells (described in chapter 4).

A combination of site-directed mutagenesis, flow cytometric assays to analyse cell surface expression and solid phase ligand binding assays using BODIPY ac-LDL were used. This is described in chapter 5. By determining the role of N-linked glycosylation with regard to the function of CD36 will help to elucidate a further understanding into the regulation of the protein, and also identify CD36 isoforms that will be suitable for further structural studies.
2. Material and methods

2.1. Bacterial culture medium

Lysogeny broth (LB): 1% NaCl, 1% tryptone, 0.5% yeast extract

LB agar: 1% NaCl, 1% tryptone, 0.5% yeast extract, 2% agar

SOC Medium: 2% tryptone, 0.5% yeast extract, 0.05% NaCl, 0.4% glucose, 0.01M MgCl$_2$, 0.01M MgSO$_4$

Where appropriate, solid or liquid growth media were supplemented with ampicillin to a final concentration of 100µg/ml.

2.2. Bacterial culture and storage

Bacteria were grown in liquid media with shaking (250 rpm), or on plates in an incubator, at 37°C. Bacteria were frozen for long-term storage at -80°C in a mixture of LB medium plus 20% glycerol.
2.3. Molecular Biology

2.3.1. Plasmids

The mammalian expression plasmid pCD36-12His, which encodes wild type CD36 with a 12 histidine tag (derived from pCIneo; Figure 2.1) and the cloning vector pGEM3 were supplied by Promega; Madison, WI and have been described earlier (Blott et al., 1999; Taylor, et al., 2001). The baculovirus transfer vectors; pBacPAK9 and pBlueBAC4.5 were supplied by BD Biosciences and invitrogen respectively. All plasmids encoding different CD36 forms have a 12 histidine tag.

![Diagram of pCD36-12His plasmid]

Figure 2.1 – pCD36-12His plasmid used in this study.
2.3.2. Preparation of plasmid DNA

Small-scale (µg-quantities) plasmid DNA preparations were generated using the GenElute HP Plasmid Miniprep Kit as described by the manufacturer (Sigma-Aldrich). The protocol is based upon the procedure of rapid alkaline lysis, as first described by Birnboim and Doby (Birnboim and Doly, 1979). Bacterial cells from a 5ml overnight culture were harvested by centrifugation for 5mins at 14,000rpm (EBA 12, Hettich; Tuttlingen, Germany). The cell pellet was then resuspended in 200µl Resuspension solution (50mM Tris-HCl pH8.0, 10mM EDTA, 100µg/ml RNase A). Lysis Buffer (200µl: 1% SDS, 0.2M NaOH) was then added and the solution was mixed by inversion. The alkalinity of the Lysis Buffer causes the denaturation of the nucleic acids and protein within the lysate. After 5mins incubation at room temperature, Neutralization Buffer (350µl; 3M Potassium Acetate, pH5.5) was added, causing the aggregation of insoluble renatured genomic DNA and high molecular weight RNA, and the precipitation of protein-SDS complexes. The lysate was then centrifuged for 10mins at 14,000rpm. GenElute HP Miniprep Binding Columns were prepared by adding 700µl Column Preparation Solution (750mM NaCl, 50mM MOPS, 15% ethanol, 0.15% Triton X-100, pH7.0) and centrifuging for 1min at 14,000rpm. The supernatant from the cell lysate was then added to the column and incubated for 1min at room temperature. Following incubation, the supernatant was passed through the columns by centrifugation at 14,000rpm for 1min. The columns were washed by adding Wash Solution (500µl; 1M NaCl, 50mM MOPS, 15% ethanol, pH8.5) and centrifuging for 1min at 14,000rpm. The eluted Wash Solution was discarded and the columns were centrifuged as before in order to remove any traces of Wash Solution. Plasmid DNA was eluted by adding pre-warmed 40µl dH2O to the column and incubating for 5min at room temperature, followed by centrifugation for 1min at 14,000rpm.
Large-scale (mg-quantities) plasmid DNA preparations were generated following a similar protocol, using the Plasmid Mega Kit (Qiagen). Bacterial cells from 400ml of an overnight culture were harvested by centrifugation for 15mins at 6000rpm, 4°C in a Sorvall RC-5B centrifuge (GS-3 rotor; Sorvall LLC, New Castle, DE). The cell pellet was resuspended in 50ml Resuspension Solution at 4°C. Lysis Buffer (50ml) was then added before mixing by inversion and incubating for 5min at room temperature. Following addition of Neutralization Buffer (50ml), the solution was then mixed by inversion, and incubated at 4°C for 30mins. The resulting lysate was cleared by centrifugation for 30min at 11,000rpm, 4°C in a Sorvall RC-5B centrifuge (GS-3 rotor). The supernatant was poured onto a Qiagen 2500 column, pre-equilibrated with 35ml Column Preparation Solution. The column was washed with 200ml Wash Solution before the DNA was eluted using 35ml Buffer QF (1.25M NaCl, 50mM Tris-HCl, 15% ethanol, pH8.5). Isopropanol (24.5ml) was added to the eluate and the DNA precipitate was pelleted by centrifugation for 30min at 11,000rpm, 4°C in a Sorvall RC-5B centrifuge (SS-34 rotor). The pellet was washing using 70% ethanol (7ml), centrifuged as before, air-dried, and resuspended in 100µl dH₂O.

Recovered plasmid DNA was quantitated by diluting the sample 1:100 in a final volume of 100µl dH₂O and measuring the optical density at a wavelength of 260nm (OD260) using a spectrophotometer (Beckman DU 640; Beckman Coulter, High Wycombe, UK). The concentration of DNA in the sample was then calculated using the following equation:

\[ \text{OD260} \times 50\text{ng/µl} \times \text{dilution factor} \]

Estimation of the purity of the plasmid DNA was carried out by measuring the ratio of OD260:OD280. A ratio of 1.8-2.0 indicated that the sample was pure.
2.3.3. Site-directed mutagenesis

Oligonucleotide-directed mutations were introduced into pCD36-wt or derivatives thereof (Table 2.1) using the ‘Quikchange II XL Site-Directed Mutagenesis kit’ or ‘Quikchange Multi Site-Directed Mutagenesis kit’ as described by the manufacturer (Stratagene). Briefly, the ‘template’ DNA is PCR amplified using a pair of mutagenic primers, each complementary to opposite strands of the vector when using Quikchange II XL Site-Directed Mutagenesis kit, or a single mutagenic primer complementary to either strand when using Quikchange Multi Site-Directed Mutagenesis kit, to generate mutation-containing synthesised DNA. Following PCR amplification, endonuclease digestion of methylated and hemi-methylated DNA is carried out using \textit{DpnI}. Since the template plasmid DNA is generated using dam+ \textit{E.coli} strain, whereas PCR-synthesised DNA is unmethylated, the template is susceptible to \textit{DpnI} digestion. This enables only the newly synthesised and mutation containing DNA to be transformed into XL-10 Gold \textit{E.Coli}.

In addition to generating codon changes, each mutagenic primer also introduced or removed a restriction endonuclease recognition site, without any further changes to the amino acid sequence. Thus, successful incorporation of mutations into the target plasmid could be determined using restriction enzyme digest on purified plasmid DNA. Each mutant \textit{CD36} gene was confirmed by sequencing and to verify no additional mutations were inadvertently introduced.
<table>
<thead>
<tr>
<th>Amino acid change</th>
<th>Nucleotide sequence (5’ → 3’)</th>
<th>Restriction enzyme site introduced</th>
</tr>
</thead>
<tbody>
<tr>
<td>N79Q</td>
<td>CACAGGAAGTGATGATGcAgAGCtcCAACATT CAAGTTAAGC</td>
<td>+SacI</td>
</tr>
<tr>
<td>N102Q</td>
<td>TCGTTTTCTAGCCAAaGAaAGcGAATaCCCCAG GACGCTG</td>
<td>-StyI</td>
</tr>
<tr>
<td>N134Q</td>
<td>TGGAAACAGAGGGCTGAtcAgTTCAAGGTTCTCA ATC</td>
<td>+BclI</td>
</tr>
<tr>
<td>N163Q</td>
<td>GTTCAATGATCCTCAATCTaATTecAgAA GTCTAAATTCTATTGTTCC</td>
<td>+AseI</td>
</tr>
<tr>
<td>N205Q</td>
<td>GGTCTGTTTTATCCTTAceAgAATACgcGCAGA TGGAGTTTATAAAG</td>
<td>+BslI</td>
</tr>
<tr>
<td>N220Q</td>
<td>GTTTTCAATGGAAAGATcATeATtcTAAAGGTGCCAATACTGC</td>
<td>+BglI</td>
</tr>
<tr>
<td>N235Q</td>
<td>CATATAAGGTAAGAAGGcAGCTGCTCATTCTTGGGAAG</td>
<td>+PvuII</td>
</tr>
<tr>
<td>N247Q</td>
<td>CACTGCGACATGATTcAgGTACAGATGCGAG GC</td>
<td>-AseI</td>
</tr>
<tr>
<td>N321Q</td>
<td>GAAAAAATTATCTCAAAAgATGTTGATCATATGC</td>
<td>+BsrDI</td>
</tr>
<tr>
<td>N417Q</td>
<td>TGTGCTTTTTTCTTGGCTTecAgGAGACTGGGA CATGGG</td>
<td>+BstXI</td>
</tr>
<tr>
<td>Q79N</td>
<td>CACAGGAAGTGATGATGcAgcAGCTCAACATT CAAGTTAAGC</td>
<td>-SacI</td>
</tr>
<tr>
<td>Q102N</td>
<td>TCGTTTTCTAGCCAAaGAaAGcGAATaCCCCAG GACGCTG</td>
<td>+StyI</td>
</tr>
<tr>
<td>Q134N</td>
<td>TGGAAACAGAGGGCTGAcAcTtcCAGTTCTCA ATC</td>
<td>-BclI</td>
</tr>
<tr>
<td>Q163N</td>
<td>GTTCAATGATCCTCAATCTCAATTaAcAA GTCTAAATTCTCTCA</td>
<td>+PacI</td>
</tr>
<tr>
<td>Q205N</td>
<td>GGTCTGTTTTATCCTTAACaAcATACGCAG ATGGAGTTTATAAAG</td>
<td>-BslI</td>
</tr>
<tr>
<td>Q220N</td>
<td>GTTTTCAATGGAAAGATaAcATCTCATAAG TGCCATAATCG</td>
<td>+PacI</td>
</tr>
</tbody>
</table>

Table 2.1 – Oligonucleotides for site-directed mutagenesis of CD36
2.3.4. **Agarose gel electrophoresis**

DNA fragments were separated by electrophoresis through a submerged gel (0.8-1% agarose, 0.3µg/ml ethidium bromide, in TBE buffer), prepared as described previously (Sambrook, 1989). Samples were mixed with 0.1 volume loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol, 30% glycerol and dH₂O), loaded into sample wells and electrophoresed at 2-10V/cm. The DNA was visualised by a UV transilluminator and images were recorded using the Gel Doc system (BioRad Laboratories, Hercules, CA).

2.3.5. **Purification of DNA fragment from agarose gels**

DNA fragments generated by restriction enzyme digestion were separated on an agarose gel as described in section 2.3.4. The DNA was visualised using a longwave UV transilluminator (VWR International, West Chester, PA) and gel fragments were excised using a scalpel. DNA was extracted from the gel fragments using the GenElute Agarose Spin Column Kit (Sigma-Aldrich), as described by the manufacturer. Briefly, the gel fragments were placed into 3 volumes of gel solubilisation solution and incubated at 55°C for 20 mins, or until the fragments have dissolved. The solution was added to a GenElute binding column and centrifuged for 1 min at 14,000rpm (Hettich EBA 12). The column was washed by adding 700µl wash solution and centrifuged for 1 min. The eluate was discarded and the column was centrifuged for a further 1 min to eliminate traces of the wash solution. The DNA was eluted from the column by the addition of 40µl pre-warmed ddH₂O, and centrifugation for 1 min at 14,000rpm.
2.3.6. **Subcloning of plasmid DNA**

Following mutagenesis, the mutated DNA was subcloned into the relevant plasmid using restriction enzyme sites and plasmids were tested for the presence of the mutation using a diagnostic restriction enzyme digest (Table 2.1). Ligations were carried out using T4 DNA ligase (400U) in ligation buffer (50mM Tris-HCl pH7.5, 10mM MgCl₂, 1mM ATP, 10mM dithiothreitol) overnight at 16°C.

2.4. **Insect cell culture**

2.4.1. **Insect cell culture medium and reagents**

SF900II and TC100 insect cell media were purchased from Invitrogen. Penicillin-Streptomycin Solution (Invitrogen) was added to all insect cell tissue culture media to a final concentration of 100 units/ml penicillin G, and 100µg/ml streptomycin sulphate, unless otherwise stated. In addition, TC100 was also routinely supplemented with 10% foetal calf serum (FCS; Invitrogen). All manipulations were performed in a sterile environment, with disposable plasticware and glassware reserved specifically for the purpose. All suspension of cultures of insect cells were grown at 27°C with shaking 100rpm.

2.4.2. **Insect cell co-transfection**

SF21 cells were co-transfected using *flash*BAC (Oxford Expression Technologies) or Bac-N-Blue transfection kit (Invitrogen) as instructed by the manufacturer. 1.5x10⁶ cells were
plated onto a 35mm tissue culture dish (BD Biosciences, San Jose, CA) in TC100 medium without FCS or antibiotics.

For each co-transfection with flashBAC, the following reagents were mixed:

- 100ng flashBAC DNA
- 500ng Transfer plasmid - pBacPAK-based vector
- 5µl Lipofectamin transfection reagent (Invitrogen)
- 1ml TC100 without serum or antibiotics

The solution was mixed and incubated for 15 minutes at room temperature. The medium was then aspirated and the transfection mixture was carefully added to the monolayer on insect cells. Each dish was then incubated overnight at 27°C in a sealed box containing a layer of paper wetted with 10mM EDTA to prevent the cells drying. Following this, 1ml TC100 (with 10%FCS and antibiotics) was added to each dish and the dishes were placed inside sealed box as before which was incubated at 27°C for 120 hours. For Bac-N-Blue co-transfection, the 100ng of flashBAC DNA was substituted with 250ng Bac-N-Blue DNA.

The protocol for Bac-N-Blue transfection was the same as for flashBAC except: for each co-transfection, the following reagents were mixed:

- 250ng Bac-N-Blue DNA
- 0.5µg Transfer plasmid (pBlueBAC-based vector)
- 500µl TC100 medium without FCS or antibiotics
- 10µl Cellfectin Transfection reagent
and after the addition of the transfection mixture the dishes were incubated for 4 hours at room temperature before adding TC100 containing 10% FCS and antibiotics. The dishes were placed in a sealed box as described previously and incubated for 72 hours. Following incubation, the medium was removed and retained as ‘Co-T’ baculovirus.

**2.4.3. Plaque assay to isolate recombinant baculovirus**

Plaque assays were used in order to isolate recombinant baculovirus clones (when using Bac-N-Blue), and also to assess the viral titre when using both Bac-N-Blue and flashBAC. In both cases, the following protocol was used.

For each baculovirus, six 60mm tissue culture dishes (BD Biosciences) were seeded with 3x10^6 Sf21 cells in TC100 medium, and left to settle on a flat surface. Whilst the cells were settling, a series of dilutions were made using baculovirus stock (10^{-2}, 10^{-3}, 10^{-4} dilutions for Co-T baculovirus, 10^{-4}, 10^{-5}, 10^{-6} dilutions for other stages of baculovirus production) to a final volume of 2ml in TC100. The culture medium was aspirated from the plates, and 1ml of each baculovirus was added. Plates were infected in duplicate. The cells were incubated for 2 hours at room temperature, during which time 4% low melting point agarose (LMP-Ag; Sigma-Aldrich) was melted using a microwave oven and placed in a 37°C water bath. The TC100 medium was also preheated to 37°C in a water bath prior to the next step.

Following the 2-hour incubation, the 4% LMP-Ag was diluted to 1% using the preheated TC100. The medium covering the Sf21 was aspirated and each dish was given a 3ml layer of 1% LMP-Ag. The dishes were placed on a flat surface and incubated at room temperature for 1-2 hours, until the agarose solidified. Following this, 2ml TC100 was added onto the agarose ‘plug’ and the dishes were placed inside a sealed box containing a layer of paper wetted with 10mM EDTA, and incubated at 27°C for 72 hours. A solution containing 0.03% Neutral red
(Sigma-Aldrich) and 50mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal; Bio-Rad) in PBS was then added to each dish (2.5ml) and incubated for 2-4 hours. The liquid medium on the surface of the agarose plug was then aspirated and the dishes were inverted and left at 27°C overnight in the dark. The neutral red stains live cells, allowing the visualisation of virus-induced ‘plaques’ of lysed or dying cells as clear regions of the monolayer.

Recombinant blueBAC baculovirus can be identified as blue plaques following the release of β-galactosidase from recombinant baculovirus-infected cells, and the subsequent chromogenic cleavage of X-gal. Non-recombinant baculovirus do not produce β-galactosidase, and so generate clear plaques.

To estimate the titre of the initial virus, the number of plaques on the highest dilution dish were counted. The initial titre (in plaque forming units/ml; pfu/ml) is equal to the plaque count multiplied by the dilution factor. Using Bac-N-Blue if the infection was done using Co-T baculovirus, pure recombinant virus must be isolated. A plugged Pasteur pipette was used to extract, by gentle suction provided by a sterile rubber bulb, the agarose above a well-isolated blue plaque which is then expelled into 1ml TC100 and incubated at 4°C overnight. The is referred to as ‘plaquepick’ baculovirus.

### 2.4.4. Amplification of baculovirus

Bac-N-Blue baculovirus amplification takes place over a number of stages. The premise is the same for each stage; with a small ratio of virus particles to insect cells (multiplicity of infection; MOI) used to allow cell growth to continue whilst multiple cycles of viral infection and release occur.
**Plaquepick-to-Passage 1 amplification**

Sf21 cells \((3 \times 10^6)\) in TC100 plus 10% FCS were seeded in a T25 tissue culture flask (BD Biosciences). Following removal of the medium, 0.5ml plaquepick baculovirus was added to the cells, and incubated at room temperature for 1 hour. Fresh TC100 (5ml) was added to the cells, which were incubated for a further 72 hours at 27°C. The medium was removed from the flask and centrifuged at 500G for 5min (Hettich rotanta 46R) to pellet the cells. The supernatant was removed and stored at 4°C as ‘passage 1’ (P1) baculovirus.

**P1-to-Intermediate stock amplification**

A suspension culture containing containing 25ml SF900 cells at a density of 1x106 cells/ml in SF900 II media was infected with 2.5ml P1 baculovirus. The cells were incubated overnight, before being diluted with 25ml fresh SF900 II media. After 5 days, the cells were removed by centrifugation at 500G for 5min (Hettich rotanta 46R), and the supernatant was stored at 4°C as ‘intermediate stock’ baculovirus. At this stage the virus titre was calculated by plaque assay.

**Intermediate stock-to-Working stock amplification**

A suspension culture containing 100ml SF900 cells at a density of 2x10⁶ cells/ml (totalling 2x10⁸ cells) in SF900 II media was infected with 2x10⁷ pfu intermediate stock baculovirus (MOI=0.1). The cells were incubated overnight, before being diluted with 100ml fresh SF900 II media. After 5 days, the cells were removed by centrifugation at 500G for 5min (Hettich rotanta 46R), and the supernatant was stored at 4°C as ‘working stock’ baculovirus.

For baculovirus generated from flashBAC expression system, amplification can take place in one step from ‘Co-T’ to Working stock. The procedure is the similar to Intermediate stock-to-Working stock amplification for Bac-N-Blue baculovirus, except the cells were infected with 0.5ml of ‘Co-T’ instead of the intermediate stock and incubated for 5 days without dilution.
2.4.4.1. Large-scale infection for protein production

A suspension culture containing 200ml SF21 serum free cells at a density of 2x10^6 cells/ml (totalling 4x10^8 cells) in SF900II media is infected with 1.2x10^9 pfu working stock baculovirus (MOI=5). The cells were incubated overnight, before being diluted with 200ml fresh SF900II media. The cells were harvested after a total of 72 hours following initial infection.

2.4.4.2. Harvesting insect cell membrane

SF21 cell suspension cultures 72 hours post-infection were separated into eight 50ml centrifuge tubes (BD Biosciences) and centrifuged at 1000G for 10 minutes (4°C; Hettich rotanta 46R). The supernatant was discarded and each of the cell pellets were washed with 5ml Buffer A (10mM Tris pH7.4, 250mM sucrose, 0.2μM CaCl$_2$, 40μM leupeptin, 2mM benzamidine, 2μM pepstatin A). The cells were centrifuged as before and all eight pellets were resuspended in a total of 16ml Buffer A. The cells suspension was homogenised on ice using an Ultra-Turrax T25 (IKA Labortechnik, Staufen, Germany). The homogeniser was used at full speed in ten 20-second bursts, with a 30 second gap on ice between bursts. The homogenate was then centrifuged at 500G for 10 minutes (4°C; Hettich rotanta 46R) to pellet any residual whole cells and large organelles. After retrieval of the supernatant, the pellet was homogenised as before and centrifuged. The crude membrane fraction was isolated from the resulting supernatant by ultracentrifugation in TLA120.2 rotor (Beckman Optima TLX) at 40,000G for 1 hour, and resuspended in 1ml ice-cold Buffer B (10mM Tris-HCl, pH7.4, 250mM sucrose, 40μM leupeptin, 2mM benzamidine, 2μM pepstatin A). The membrane
preparation was homogenized by repeatedly passing through a 25G syringe. Membrane preparations were stored by adding 10% glycerol and freezing at –80°C.

2.5. Mammalian Tissue Culture

2.5.1. Mammalian tissue culture medium and reagents

Dulbecco’s Modified Eagle Medium (DMEM) with glutamax, Trypsin-EDTA, TrypLE Express, FCS and the transfection reagent Lipofectamine 2000 were purchased from Invitrogen.

The cationic polymer Polyethyleneimine (PEI), used as a transfection reagent, was supplied by Sigma-Aldrich.

Human embryonic kidney (HEK293T) cells (Imperial Cancer Research Fund, cell production unit) were grown as monolayers and maintained by regular passage in DMEM supplemented with 10% FCS under 5% CO\textsubscript{2} at 37°C with a water vapour-saturated atmosphere. All manipulations were performed in a sterile environment, with disposable plasticware and glassware reserved specifically for the purpose.

2.5.2. Transient transfection of HEK293T cells

For transfection in a T25 tissue culture flask (BD Biosciences), HEK293T cells were grown to 80% confluency as a monolayer. A solution containing 10µg DNA at a concentration of 0.5µg/µl in 5% glucose was mixed with 3µl PEI and incubated at room temperature for 10 minutes. DMEM (4.5ml) with 10% FCS was then added before applying to the cells. After 24 hours, butyric acid (Sigma-Aldrich) was added to a final concentration of 33µM. The cells
were harvested after a further 24 hours by brief incubation with TrypLE-Express (0.5ml). The trypsin was neutralised by the addition of 4.5ml DMEM supplemented with 10% FCS, and centrifuged (Hettich rotanta 46R, 100G, 10 minutes, 4°C).

For transfection in a 96 well plate, 50µl Poly-L-lysine (Sigma-Aldrich) was added to each well and incubated at 37°C for 30 minutes. The wells were then washed with PBS, and 1x10⁵ HEK293T cells were added to each well in DMEM, to a final volume of 150µl. 24 hours later, 0.6µg DNA in 5% glucose was mixed with 0.09µl PEI and incubated at room temperature for 10 minutes. 300µl DMEM with 10% FCS was then added before applying to the cells. The cells were used for a ligand binding assay (see section 2.7.10) 48 hours later.

2.5.3. Tunicamycin Treatment

500ng/ml of tunicamycin was added to 5 hours post-transfection where indicated.

2.6. Flow cytometry

All flow cytometry was performed using a FACScan benchtop analyser (BD Biosciences). The FACScan possesses an argon laser with an emission at 488nm and five detectors: two for light scatter and three (FL1, FL2 and FL3) for fluorescence. FACSflow sheath fluid and 5ml polypropylene FACS tubes were purchased from BD Biosciences.
2.6.1. Cell surface expression of CD36

HEK293T cells, transfected transiently with plasmid DNA encoding wild type CD36 (pCD36-wt) or mutant CD36 (in T25 tissue culture flask) were harvested (along with untransfected cells). The resulting cell pellets were then resuspended in FACS buffer (PBS with 0.1% Essentially Fatty Acid Free BSA (Sigma Aldrich)) in 50µl to a concentration of 1x10^7 cells/ml, and incubated with 2µg mAb1258 or μg mAb1955 for 30 mins at 4°C. The cells were recovered by centrifugation (Hettich rotanta 46R, 150G, 2 minutes, 4°C), and resuspended in 500µl FACS buffer. This washing step was repeated a further two times, before finally resuspending the cells in 50µl to a concentration of 1x10^7 cells/ml. The cell suspension was then incubated with 4µg of goat, anti-mouse IgG secondary antibody conjugated to R-phycoerythrin or goat, anti-rat IgG secondary antibody conjugated to R-phycoerythrin for 30 minutes at 4°C in the dark. The cells were then pelleted and washed in FACS buffer as before, and resuspended in 300µl for analysis.
2.7. Protein Biochemistry

2.7.1. Materials used for protein biochemistry

Leupeptin and Benzamidine protease inhibitors were purchased from Sigma Aldrich, along with N-Dodecylmaltoside C12 (DDM) and N-decylmatoside (DM). The detergents Fos-Choline 10 (FC10®) and Fos-Choline 12 (FC12®) were from Anachem, and Qiagen supplied the Nickel-Nitrilotriacetic acid (NTA) Agarose. Pepstatin A and Octaethylene glycol monododecyl ether (C<sub>12</sub>E<sub>8</sub>) was purchased from Calbiochem and n-Octyl-β-D-glucopyranoside (OG) was purchased from Merck Biosciences.

2.7.2. Purification of CD36 by affinity chromatography

2.7.2.1. Wildtype CD36

Insect cell membrane preparations were pelleted by centrifugation (Beckman Optima TLX; 100,000G, 50 min, 4°C), and then solubilized at 4°C with 2% (w/v) OG at a concentration of 5 mg protein ml<sup>-1</sup> in Buffer 1 (150mM NaCl, 20mM Tris, 1.5mM MgCl<sub>2</sub>, 5% (v/v) glycerol, 40μM leupeptin, 2mM benzamidine, 1μM pepstatin, pH6.8), homogenized by extrusion in a 21G needle and constantly mixed for 90 minutes at 4°C. The insoluble fraction was then pelleted by centrifugation (Beckman Optima TLX; 100,000G, 30 min, 4°C.) Ni-NTA Agarose was washed in dH<sub>2</sub>O (16μl of 50% slurry/1mg protein) and pre-equilibrated in Buffer 2 (Buffer 1 where 2% OG is replaced with 1% OG) in the presence of 20mM imidazole, at a volumetric ratio of 1:10 (packed resin:Buffer 2). To prevent non-specific interactions between proteins and resin, imidazole was added to the soluble fraction of CD36 at a final concentration of 20mM. The protein and resin were incubated with continuous
mixing for 1 hour at 4°C to ensure the binding of (His12)-tagged CD36. The non-bound proteins were decanted (flow-through), and the resin bed was washed four times with 20 bed volumes of ice-cold 60 mM imidazole, then once with 20 bed volumes of 60 mM imidazole, 80 mM imidazole, 100 mM imidazole and 120 mM imidazole, all in Buffer 3 (Buffer 2 but at pH 8), to remove non-specifically bound proteins (washes 1–4). The resin was further washed in 20 bed volume of Buffer 2, in the presence of 5 mM imidazole, to alter the pH for optimum elution of (His12)-tagged CD36 (final wash). The resin was further washed four times in Buffer 2, in the presence of 250 mM imidazole to elute CD36 (elutions 1–4).

2.7.2.2. Non-glycosylated CD36-12 Histidine

To purify non-glycosylated CD36 (CD36nong), the same protocol was used as wildtype CD36 except, the 2% OG in Buffer 1 was substituted for 0.6% SDS and 1% OG in Buffer 2 was substituted for 0.3% SDS. The solubilized protein was incubated with Ni-NTA with continual agitation for 30 minutes at room temperature, rather than for 1 hour at 4°C.

2.7.2.3. CD36N8-10-12 Histidine

To solubilise CD36N8-10, 2% OG in Buffer 1 was substituted with each of the following: 1% SDS, 4% OG, 2% DDM, 2% DM, 2% FC9®, 2% C12E8 and 1% Triton X-100 (TX100). To monitor purification, samples were removed from each stage, subjected to SDS-PAGE, and protein content visualized by colloidal blue staining. Diluted protein samples were concentrated by trichloroacetic acid precipitation prior to electrophoresis.
2.7.3. **Trichloroacetic acid precipitation**

Dilute protein solutions were concentrated by trichloroacetic acid precipitation. 0.1 (v/v) of 0.15% sodium deoxycholate was added to the protein sample and incubated at room temperature for 5 minutes. Following this, 0.1 (v/v) of 72% trichloroacetic acid was added, the sample mixed, and incubated for a 10 minutes at room temperature. Protein was recovered by centrifugation at 14,000rpm for 8 minutes (Hettich EBA 12). The supernatant was carefully removed and discarded, and the pellet resuspended in 15µl resuspension buffer (4% SDS, 0.2M Tris pH7.4, 0.15M NaOH).

2.7.4. **Concentration of protein**

The elutions were concentrated using Amicon Ultra-15 centrifugal devices (Millipore) and spun at 4000G (Hettich rotanta 46R) for between 10 and 25 minutes until the sample is 10% of the starting volume, and the protein recovered.

2.7.5. **Protein quantitation**

Protein content of samples was measured using RC DC protein assay (Bio-Rad). The assay is based upon the Lowry method for determining protein concentration (Lowry *et al.*, 1951), and was carried out according to the manufacturer’s protocol. Samples were diluted where necessary, and made up to 20µl with dH₂O. A range of BSA (1-20µg) was used to produce a standard curve.
2.7.6. Preparation of protein samples for SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

All membrane preparations were treated with 25U benzonase before addition of Laemmli sample buffer (LSB; 100mM Tris-HCl pH 6.8, 20% glycerol, 10% 2-mercaptoethanol, 4% SDS, 0.1% bromophenol blue). Protein samples were then prepared for SDS-PAGE as described by Laemmli (Laemmli, 1970). Protein was added to an equal volume of LSB and incubated at 37°C for 10 minutes prior to loading the gel.

For whole cell lysates of transiently transfected HEK293T cells, 200µl Buffer 4 (2% SDS in PBS with 40µM leupeptin, 2mM benzamidine and 1µM pepstatinA) with 25U benzonase was used and cells incubated for 30 minutes at room temperature.

2.7.7. SDS-PAGE

Proteins were separated by electrophoresis through vertical 0.75-1.5mm thick polyacrylamide gels (0.1% SDS, 10% acrylamide), using the procedure described by Laemmli (Laemmli, 1970). Briefly, the required volumes of separating and stacking gel, together with SDS-PAGE running buffer, were prepared as follows:
10X SDS-PAGE running buffer:
- 250mM Tris base
- 1.92M glycine
- 1% SDS

Per gel (0.75mm thickness):

Separating gel:
- 750µl 1.5M Tris HCl (pH 8.8)
- 1.15ml dH₂O
- 1ml 30% Acrylamide/Bis-acrylamide mix (37.5:1 ratio; Sigma-Aldrich)
- 30µl 10% SDS
- 23µl 20% Ammonium persulphate
- 5.5µl TEMED (Sigma-Aldrich)

Stacking gel:
- 250µl 1.0M Tris HCl (pH 6.8)
- 1.4ml dH₂O
- 330µl 30% Acrylamide/Bis-acrylamide
- 20µl 10% SDS
- 20µl 20% Ammonium persulphate
- 2µl TEMED (Sigma-Aldrich)

Polymerisation of the separating gel was initiated by the addition of N,N,N’,N’-tetramethylethane-1,2-diamine (TEMED). Once the separating gel had polymerised, TEMED was added to the stacking gel to initiate polymerisation, and it was overlaid onto the separating gel. The wells were rinsed with running buffer prior to loading the samples. A lane containing full range rainbow molecular weight markers (RPN800; GE Healthcare,
Amersham, UK) was loaded onto each gel to enable the subsequent determination of protein molecular weights. The gels were run at 15-40 V/cm for 1-2 hours.

2.7.8. Detection of protein by colloidal blue staining

Total protein visualisation following separation by SDS-PAGE was carried out using the Colloidal Blue Staining Kit (Invitrogen), according to the manufacturer’s protocol. Briefly, the SDS-PAGE gel is placed into colloidal blue solution (per gel: 20ml Stainer A, 5ml Stainer B, 20ml 100% methanol, 55ml dH2O), and incubated for 3-12 hours at room temperature on a rocking platform. The gel is then given one 5-minute, two 1 hour, and one 10-hour washes in dH2O at room temperature, with rocking, in order to de-stain and allow visualisation of protein bands.

2.7.9. Detection of proteins by western blotting

Proteins were separated by SDS-PAGE and then electroblotted onto a polyvinylidene fluoride (PVDF) membrane (Millipore, Billerica, MA) for western blotting analysis. The PVDF membrane was first prepared by soaking in 100% methanol for 1 minute, followed by a 1 minute wash in dH2O. Finally, both the PVDF and 3MM filter paper (Whatman, Maidstone, UK) were soaked in western transfer buffer (25mM Tris HCl, pH8.3, 192mM glycine, 20% methanol). Electroblotting was carried out at 3-4V/cm2 for 45-60 minutes using a B34 Fastblot semi-dry apparatus (Biometra, Goettingen, Germany), according to the manufacturer’s instructions. Following electroblotting, the PVDF membrane was washed with PBST (0.1% Tween20 in PBS) briefly before being incubated in blocking buffer (5%
skimmed milk powder in PBST) for 1 hour at room temperature on a rocking platform. The membrane was then incubated with 1ml primary antibody (diluted 1:500 in blocking buffer, unless otherwise stated) inside a heat-sealed bag overnight at 4°C with gentle agitation. The membrane was given three 15 minute washes with PBST before incubation with secondary goat-anti mouse antibodies, conjugated to horseradish peroxidase (HRP; Dako). The secondary antibody was diluted 1:1000 in 1ml blocking buffer and the membrane was incubated in a heat-sealed bag for 1 hour at room temperature. The membrane was given four 15 minute washes in PBST before visualisation of the protein-bound HRP was carried out using enhanced chemiluminescence (ECL) reagents, as directed by the manufacturer (GE Healthcare), together with X-OMAT LS film (Kodak, Hemel Hempstead, UK).

2.7.10. Solid-Phase Ligand binding assay

1µg of purified protein diluted in 100µl of Buffer 4 (Buffer 2 but 1% OG was replaced with 0.5% OG) was added to each well of a Ni-NTA HisSorb™ white plate (Qiagen) and was incubated at 4°C overnight with gentle agitation. BODIPY Ac-LDL was diluted using ice-cold Buffer 5 (1mM MgCl₂, 1mM CaCl₂, 1% fatty-acid free BSA and PBS) to the following concentrations 0, 0.5, 2, 5, 10, 25, 40, 50 and 60µg/ml each at a final volume of 150µl per well. After the unbound protein was removed, the wells were washed in 150µl of ice-cold Buffer 5 and the ligand was added. The plate was incubated for 2 hours at room temperature with gentle agitation. Once the ligand was removed, each well was washed three times in ice-cold Buffer 6 (Buffer 5 but 1% fatty acid free BSA was replaced with 0.5% fatty acid free BSA) and 100µl of PBS was added to each well.

To measure ligand binding to CD36 expressed on the surface of HEK293T cells, 48 hours following transfection (see section 2.5.2) the wells were blocked with 200µl Buffer 7 (Buffer
5% but 1% fatty-acid free BSA was replaced with 0.2% fatty-acid free BSA) at room temperature with gentle agitation. BODIPY Ac-LDL was diluted to the above concentrations in Buffer 5 and added to each well once the block was removed. The plate was incubated with gentle agitation at 4°C for 2 hours. The cells were then washed twice with 200µl of Buffer 5 and 100µl of PBS was added to each well. The fluorescence was measured using SpectraMax® Gemini Em Dual-Scanning Microplate Spectrofluorometer and Softmax® Pro software (Molecular Devices).

2.7.11. Glycosidase digestion

Proteins were deglycosylated using Endoglycosidase H (Endo H) and Peptide-N-Glycosidase F (PNGaseF) (New England Biolabs) as directed. Briefly, 5µl of lysed cells (see section 2.7.6) or 10pmol of purified protein were incubated for 10 minutes at 100°C in the presence of denaturing buffer (0.5% SDS, 1% β-mecaptoethanol) and incubated on ice for 2 minutes. 500U of Endo H or PNGaseF was added with G5 buffer (50mM Sodium Citrate) or G7 (50mM Sodium Phosphate) supplemented with 1% NP40, respectively and incubated at 37°C for 2 hours.

For deglycosylation of native purified protein, 100ng protein was incubated with increasing amounts of PNGaseF (1-250Units) in the presence of G5 buffer, and incubated at 37°C for 48 hours.
2.8. Mass Spectrometry

Following separation of purified wildtype protein with and without PNGaseF treatment on SDS-PAGE and staining with colloidal blue (see section 2.7.8), the protein bands were excised and digested with tymsin.

2.8.1. Quadrupole Time of Flight MS/MS

In-Gel tryptic digest was carried out by MassPREP Station (Waters) and the peptide was extracted using 0.1% formic acid. The tryptic peptide mixture was analysed by automated liquid chromatography/tandem mass spectrometry (LC/MS/MS) (CapLC, LC Packing, Q-ToF II, Waters) as described (Gavin et al., 2002).

2.8.2. Fourier-transform ion cyclotron resonance (FT-ICR) MS

In-Gel tryptic digest was carried out by BioRobot 3000 (Qiagen) and extracted using 0.1% formic acid, and the tryptic mixture was analysed using LTQ-FT hybrid linear trap/7-T FT-ICR mass spectrometry (Thermo Electron Bremen Germany) as described (Peterman et al., 2005).

The data from LC/MS/MS and FT-ICR was analysed used in conjunction with the MSDB database using the software tool Mascot (Matrix Services) and Sequest database using Bioworks software (Thermo Scientific) respectively.
3. Expression, purification and ligand binding of CD36

3.1. Introduction

The study of membrane proteins encounters many difficulties, one of the most important being that membrane proteins are generally expressed at low levels in their native environment. This chapter explains how this problem was overcome by expressing and purifying recombinant human CD36.

There are four types of expression systems commonly used, bacteria, yeast, insect cells and mammalian cells, each one having its own advantages and disadvantages. Using bacteria and yeast is cheap and they are easy to manipulate with quick growth and high protein expression. However, bacteria lacks biochemical pathways involved in post-translational modifications which can lead to misfolding and aggregation of eukaryotic proteins in inclusion bodies. Yeast also lacks cholesterol in the membrane which may affect the localisation of membrane proteins. Insect and mammalian cells have a longer generation time of approximately 24 hours and are more expensive to grow. Insect cells in particular are more difficult to manipulate and involve the use of baculovirus. Both cell types are able to add complex post-translational modifications to the proteins, and mammalian cells often have the closest native environment for the protein of interest. Despite this, the protein yield in mammalian cells is often poor compared to insect cells.

Once membrane proteins are expressed, they need to be solubilized from the membrane to enable purification. This is routinely achieved by use of detergents.
3.1.1. Detergents

Detergents are amphipathic molecules that consist of a polar head and a hydrophobic tail. In aqueous solutions and at a given concentration, known as the critical micelle concentration (cmc), detergents spontaneously form (generally) micelle structures. These are spherical and can frequently solubilize membrane proteins from a plasma membrane by mimicking the natural lipid bilayer environment. The release of proteins from the membrane occurs in three stages, initially the detergent binds to the membrane and initiates lysis. Secondly, the membrane becomes solubilized and mixed detergent-lipid-protein micelles are formed. Lastly, as the detergent reaches the cmc, the membrane is fully solubilized forming detergent-lipid and detergent-protein complexes (Figure 3.1). There are many types of detergents, and each detergent has its own critical micelle concentration (cmc), however this can vary with changes in temperature, pH and ionic strength (Table 3.1).

Structurally, detergents can be classified into four categories:

1. Ionic detergents – for example SDS. These contain a head group with a net charge that is either cationic or anionic, and are very effective at solubilizing membrane proteins but are almost always denaturing.

2. Bile Acids Salt – are also ionic detergents, but their backbone have steroidal groups, so there is a polar and apolar face instead of a head group. These do not form spherical micelles but small kidney-shaped aggregates. They are relatively mild and therefore are less denaturing than the classic ionic detergents.
3. Non-Ionic detergents for example OG. These contain an uncharged head group and are considered mild and non-denaturing as they break lipid-lipid and lipid-protein interactions but not protein-protein interactions.

4. Zwitterionic detergents such as FC9®, contain properties of ionic and non-ionic detergents but tend to be more denaturing than non-ionic detergents.

Figure 3.1 - Detergent solubilization of a membrane protein.

The detergent binds to the membrane, and as the detergent concentration increases mixed micelles are formed. As the cmc is reached detergent-protein micelles and detergent-lipid micelles are formed (Ahmed, 2005)
Each membrane protein behaves uniquely, so different detergents need to be tested to establish which is suitable for your protein of interest.

<table>
<thead>
<tr>
<th>DETERGENT</th>
<th>CMCmM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Octylglucoside C8 (OG)</td>
<td>25 (0.6)</td>
</tr>
<tr>
<td>N-Dodecylmatoside C12 (DDM)</td>
<td>0.15 (0.0087)</td>
</tr>
<tr>
<td>N-Decylmatoside (DM)</td>
<td>1.6 (0.087)</td>
</tr>
<tr>
<td>TX-100</td>
<td>0.25 (0.015)</td>
</tr>
<tr>
<td>C_{12}E_8</td>
<td>0.11 (0.005)</td>
</tr>
<tr>
<td>FOS-CHOLINE® 9 (FC9)</td>
<td>39.5 (1.2)</td>
</tr>
<tr>
<td>SDS</td>
<td>7.1 (0.17)</td>
</tr>
</tbody>
</table>

Table 3.1 - A List of detergents with the cmc at room temperature and pH7.0

3.2. Expression of CD36 in *E.coli* strains

Bacteria’s inability to add post-translation modifications can in some circumstances be seen as an advantage as the recombinant proteins are likely to be homogeneous, which is ideal for use in crystallization trials to solve the structure of membrane proteins. In order to achieve this homogeneity, attempts were made to purify sections of the extracellular domain of CD36 in 4 different *E.coli* strains, Rosetta Blue, Origami, C41 and C43. Unfortunately, this resulted in the mammalian protein aggregating forming inclusion bodies (Linton, K.J. unpublished data), therefore alternative expression systems were considered. Recombinant P-glycoprotein
(P-gp), a human ABC transporter membrane protein, had previously been purified from insect cells in the lab and insect cells are widely used as an expression system for eukaryotic membrane protein such as human chemokine receptor CXCR4 (Dukkipati et al., 2006), transient receptor potential channel TRPM8 (Phelps and Gaudet, 2007) and human bradykinin B2 receptor (Reyes-Cruz et al., 2000). Insect cells were therefore tested as a suitable host for CD36.

### 3.3. N-linked glycosylation in insect cells

N-linked glycosylation on the sequence Asn-X-Ser/Thr is recognized and occurs, in principle, on the same sites in both mammalian and insect cells (Hsieh and Robbins, 1984), although under glycosylation is observed in proteins derived from insect cells, as judged by glycan analysis (Salmon et al., 1997). It has now become clear that early stages of N-linked glycosylation in insect cells resemble those in mammalian cells, with the Asn of the consensus sequence receiving the core glycan, which is trimmed immediately in the endoplasmic reticulum and sequentially by the enzymes [Glucosidase I and II and α-mannosidases (see section 1.5.1). However, insect cells lack the enzymes present in mammalian cells necessary for elaboration of these trimmed glycans, (Altmann et al., 1999), and so generate minimally elongated structures called paucimannose, which consists of Man3-GlcNAc2. Despite these differences, insect cells are commonly used to express mammalian glycoproteins as the proteins generated are generally biologically and functionally active (Jarvis, 2003).
3.3.1. The baculovirus expression system

Baculovirus infection is a very effective way to introduce a gene encoding your protein of interest to insect cells, and manipulation of the baculovirus genome has resulted in high yield of recombinant protein. The baculovirus used is derived from the *Autographa californica* nuclear polyhedrosis virus (AcMNPV). The life cycle of the wild type AcMNPV encompasses two functionally and morphologically distinct viral forms: budded virus (BV) and occluded virus (OV) (Miller, 1989). Approximately 12 hours following the infection of host cells, new viral capsids pass through the cytoplasm and bud from the cell surface, coated in a membrane envelope. This BV particle can diffuse through the extracellular milieu and efficiently infect neighbouring cells by absorptive endocytosis. Infectious BV particle synthesis then continues, resulting in a logarithmic increase in BV release for a further 10 hours. Formation of OV occurs later, at around 20 hours post-infection, during which time the synthesis of BV is dramatically reduced. Large polyhedral occlusion bodies are formed which contain several enveloped virus particles within a matrix of polyhedrin protein. The synthesis of OV continues for approximately 50 hours, during which the level of polyhedrin protein within the cell can reach 50% of the total protein content (Wickham *et al.*, 1992). This protein envelope allows the virions to be protected in the wild from environmental stresses, however during cell culture, conditions are maintained to suit the virus, therefore OV formation and the expression of polyhedrin are no longer a prerequisite for survival of the virus.

The BAC-n-Blue (Invitrogen) and *flash*BAC baculovirus system (Oxford Expression Technologies) exploits the extremely strong polyhedrin promoter, and the lack of need for polyhedrin protein. The system replaces, via recombination, the non-essential polyhedrin gene with the cDNA of the protein of interest. This allows large quantities of BV but not OV
to be produced, and at a very late stage of infection large quantities of recombinant protein, due to the cDNA being under control of the polyhedrin promoter.

In addition, flashBAC baculovirus system lacks the gene for chitinase (chiA) which is non-essential for replication. When insect cells are infected with wild type AcMNPV, chitinase form para-crystalline array in the endoplasmic reticulum compromising the efficacy of the secretory pathway (Saville et al., 2004; Thomas et al., 1998), therefore deletion of the chiA improves the yield of recombinant proteins that are secreted or membrane targeted.

In the Bac-N-Blue expression system, the viral genome has 3 Bsu36I restriction enzyme sites that allow complete linearization and removal of sequences essential for propagation of the virus in insect cells. These essential sequences can be restored via recombination using the transfer vector, which also introduces the gene of interest downstream of the polyhedrin promoter. Since the removal of these essential sequences is dependent on an enzyme, it is possible that the parental viral DNA can propagate due to incomplete digestion. Hence, β-galactosidase activity resulting from lacZ expression is utilized to identify recombinant baculovirus, enabling rapid screening of a large number of viral plaques. There is no need for this step when using flashBAC since the baculovirus genome lacking an essential gene is purified from bacteria resulting in replication incompetent baculovirus without first recombining with the transfer vector.

### 3.3.2. The baculovirus transfer vectors

The coding sequences for wild type CD36-12His, non-glycosylated CD36-12His (CD36nong) and minimally-glycosylated CD36-12His (CD36N8-10) were excised from the pGEM3 vector (see chapter 2.3.1) by digestion with BstEII and NcoI. CD36nong was generated by site directed mutagenesis of the 10 asn within the N-linked glycosylation
consensus sequence to glutamine (see section 5.2.2.1), and CD36N8-10-12His was generated by replacing the asparagines within glycosylation sites N1-N7 with glutamine (leaving N8, N9 and N10 intact; section 5.2.2.4). These cDNA fragments were then ligated between the BstEII and NcoI sites of pBlueBAC4.5 to generate pBlueBAC-CD36-12His, pBlueBAC-CD36-nong-12His and pBlueBAC-CD36N8-10-12His, respectively. XhoI and EcoRV were then used to excise CD36 from pBlueBAC4.5 and ligate into pBacPAK9 to generate pBacPAK-CD36-12His, pBacPAK-CD36nong-12His and pBacPAK-CD36N8-10-12His. As explained in section 3.3.1, the pBlueBAC and pBacPAK vectors provide a strong polyhedrin promoter upstream of CD36 gene to allow high expression of the protein, and pBlueBAC also contains a 5' fragment of the lacZ gene that allows selectivity for recombinants, as shown in Figure 3.2

### 3.3.3. Generation of CD36-expressing baculoviruses by homologous recombination

Upon successful co-transfection of the insect cell host, homologous recombination takes place between lacZ (pBlueBAC) (or lef2 (pBacPAK9)) and ORF1629 gene fragments on the transfer vector and the AcMNPV genome, resulting in the formation of a recombinant baculovirus genome containing lacZ-CD36-ORF1629 (Figure 3.1) or lef2-CD36-ORF1629 genes. Since baculovirus lacking complete ORF1629 fail to propagate successfully, recombinants are preferentially selected for.
Successful recombination in the co-transfected insect cell host results in the incorporation of the gene of interest, and concomitant repair of lacZ, encoding β-galactosidase and ORF1629, which is essential for virus propagation (Kitts and Possee, 1993). Expression of the gene of interest is controlled by the baculovirus polyhedrin promoter (P_{PH}).
3.4. Results

Following co-transfection, recombinant baculovirus clones derived from the BlueBac vectors were isolated by plaque assay (detailed in section 2.4.3). Each baculovirus was assayed for its ability to express CD36 by western blot analysis following a small scale infection of Sf21 cells (Figure 3.3). CD36-expressing baculovirus clones were selected for further amplification.

![Immunoblot autoradiograph showing expression of CD36 from baculovirus clones.](image)

**Figure 3.3** - Immunoblot autoradiograph showing expression of CD36 from baculovirus clones. Whole cell lysate from $1 \times 10^5$ Sf21 cells were separated by SDS-PAGE and western blotting was carried out as described in chapter 2. Membrane protein samples were probed using anti-His tag antibody. Blot (1) wild type CD36 migrating as an, approximately, 75kDa protein with non-glycosylated immature form at approximately 50kDa. (2) CD36nong. Protein sizes were determined by 1kb protein marker ran alongside the lysates.
3.4.1. Baculovirus amplification and protein expression

Recombinant, CD36-expressing baculovirus was used for a series of Sf21 insect cell infections, as detailed in chapter 2. Each sequential infection was performed at low multiplicity of infection (M.O.I; 1 virus to 10 cells) to generate intermediate and working stocks of the virus at high titre (section 2.4.4). To determine the optimal number of days post-infection that Sf21 should be harvested for optimum protein synthesis, a time course assay for pBlueBAC-CD36 infected Sf21 serum-free was carried out to analyse viability and density of cells (Figure 3.4 A), and each sample was probed using anti-His antibody by immunoblot for protein production (Figure 3.4 B).

Figure 3.4 - A time course assay to monitor infection and expression of CD36 in insect cells at x 5 M.O.I.

In (A), cell density is shown as bars and cell viability show as a line, determined by the number of % of live cells in total counted, and in (B) protein synthesis by probing with anti-His antibody in Sf21 cells infected with pBlueBAC-CD36 wildtype.
This assay established that 72 hours post transfection was the best time point to harvest Sf21 cells since the cell density is high and the cells remain viable, and wild type CD36 is being synthesised. Sf21 cells infected using the \textit{flashback} system were harvested using these same conditions.

\subsection*{3.4.2. Protein purification}

The first step in purification of CD36 from insect cells is to isolate the plasma membrane, and then solubilize the proteins, using detergent, from the crude membrane preparation. Addition of detergent disrupts the membrane, allowing separation of individual proteins, whilst also acting as a substitute for the lipid bilayer and helping to maintain the membrane protein in a stable and functional conformation (Bowie, 2001). Purification of CD36 utilizes the 12-Histidine tag using one-step purification involving metal affinity chromatography. The polyhistidine tag facilitates the binding of the protein to Ni-NTA agarose, separating this protein from the other soluble membrane proteins. CD36-12His can be washed off the resin using imidazole, which competes with the imidazole rings of the polyhistidine tag for binding sites on the Ni-NTA.

\subsubsection*{3.4.2.1. Wild type CD36}

Previous studies on rat Cd36 have identified octylglucoside (OG) as a suitable detergent to solubilize crude membrane preparations (Martin \textit{et al.}, 2007). OG at a concentration of 2\% and a membrane protein concentration of 5mg protein/ml of buffer was used to solubilize human CD36. The soluble proteins were separated from the insoluble fraction by
ultracentrifugation and then incubated with Ni-NTA agarose at 8:1 protein to resin ratio. The resin was washed four times with 20 bed volumes of wash buffer and increasing concentrations of imidazole to remove proteins bound non-specifically. The bound CD36-12His was eluted in 250mM imidazole. Figure 3.5A top panel shows a typical colloidal blue stained SDS-PAGE gel of fractions removed from each stage of purification and the bottom panel shows an immunoblot of CD36 contained in each fraction, probed with anti-His antibody. By measuring the pixels in the immunoblot, and knowing the concentration of the purified protein (see section 3.4.3), it is possible to estimate the amount of CD36 contained in the starting membrane preparation and the efficiency of each purification step. Approximately, 1.9mg of CD36 was contained in the membrane preparation of 50mg total protein (3.8%), and an estimated 88% was solubilized using OG. The main losses of CD36 occurred when the protein failed to bind to the Ni-NTA resin and during the final washes when the imidazole concentration was increased. The concentration of CD36 eluted was about a fifth of the starting material, approximately 368µg, however, the protein was about 83% pure before concentration. Although it is difficult to compare the purification strategy from protein to protein because each protein behaves differently, in a set of 15 human proteins the observed yield after purification was generally higher when the protein is purified using the His tag rather than a GST tag (Scheich et al., 2003). A 200ml culture of CD36-12His-expressing Sf21 serum free cells typically yields 50mg total membrane protein, from which approximately 350µg CD36-12His was purified. Although the yield and purity of the wild type protein was reasonable, the protein was too dilute so the samples needed to be concentrated (see section 3.4.3)
3.4.2.2. CD36nong

CD36nong is expressed via the baculovirus system and is found within the crude membrane preparation pellet. Non-ionic detergents such as OG were unable to solubilize the protein, so the ionic detergent SDS at 0.6 % was used. However, SDS is not compatible with Ni-NTA purification, so following solubilization, the percentage of SDS was diluted to 0.3% before binding to the resin. Figure 3.5 (B) shows a typical colloidal blue stained SDS-PAGE gel of fractions removed from each stage of purification, where very little CD36nong was enriched during elutions.
Figure 3.5 - SDS polyacrylamide gels stained with colloidal blue and an immunoblot probed with anti-His showing protein and wild type CD36 from each purification fraction. Lanes: 1, 10µl (50µg) crude membrane preparation (0.2% of total volume); 2, 10µl solubilized membrane preparation (0.1% of total volume); 3, 10µl of fraction not bound to NiNTA following 90 mins incubation (0.1% of total volume); 4-7, 200µl wash fractions (2% of total volume) with 60mM imidazole wash 4, 80mM imidazole wash 5, 100mM imidazole wash 6 and 120mM imidazole wash 7. 8, 200µl final wash (2% of total volume); 9-12, 50µl elutions (5% of total volume) or 15µl elution on the immunoblot (1.5% of total volume) with 250mM imidazole. Panels (A) wild type CD36, top panel SDS-PAGE stained with colloidal blue, bottom panel immunoblot probed with anti-His antibody, and in (B) CD36nong stained with colloidal blue.
3.4.2.3. **CD36N8-10**

CD36N8-10 was also expressed via the baculovirus system, and since OG was suitable in solubilization of wild type CD36, the same protocol was tested for purification of CD36N8-10. Figure 3.6 shows a typical purification gel, demonstrating that CD36N8-10 was not highly enriched. Immunoblotting demonstrated that OG failed to solubilize membrane bound CD36N8-10 efficiently (Figure 3.7A), therefore, further investigation was necessary to determine whether other detergents could solubilize CD36N8-10 more efficiently.

![Figure 3.6 - SDS polyacrylamide gel stained with colloidal blue, showing total protein CD36N8-10 from each fraction during purification.](image)

**Figures:**

- Lanes: 1, 10µl (50µg) crude membrane preparation (0.2% of total volume); 2, 10µl solubilized membrane preparation (0.1% of total volume); 3, 10µl of fraction not bound to NiNTA following 90 mins incubation (0.1% of total volume); 4-7, 200µl wash fractions (2% of total volume) with 60mM imidazole wash 4, 80mM imidazole wash 5, 100mM imidazole wash 6 and 120mM imidazole wash 7. 8, 200µl final wash (2% of total volume); 9-12, 50µl elutions (5% of total volume) with 250mM imidazole.
3.4.2.3.1. Solubilization of CD36N8-10

To allow purification, it was necessary to establish which detergent was effective at solubilizing CD36N8-10 from the membrane. Membrane preparations from Sf21 cells infected with CD36N8-10 (50µg total protein, at 5mg/ml concentration) were mixed with detergents DDM, DM, TX100, FC-9®, SDS and C₁₂E₈ for 90 minutes at 4°C (all at a concentration of 2%, except TX100, which was used at 1%, in solubilization buffer which includes 20mM Tris-HCl, 150mM NaCl, 1.5mM MgCl₂, 5% glycerol and protease inhibitors) to establish whether the protein could be solubilized from the membrane. Ultracentrifugation was used to pellet the insoluble fraction leaving the soluble proteins in the supernatant (see section 2.7.2.3) Figure 3.7 shows that only FC-9® and SDS managed to solubilize CD36N8-10 efficiently. In each membrane preparation there appears to be at least two forms of the proteins, which are likely to be CD36 with different glycosylation statuses. When wild type CD36 is incubated with OG (Figure 3.7A), only the fully glycosylated protein is soluble (approximately 75kDa), the under glycosylated protein (approximately 50kDa) is probably contained in an insoluble protein aggregate. FOS-CHOLINE® 9 and SDS detergents (Figure 3.7B, lanes 4 and 12 respectively) extracted not only the glycosylated protein of CD36N8-10, but also the under-glycosylated forms. Possibly, the nature of these two detergents (zwitterionic and ionic respectively) allow the solubility of fractions that are otherwise insoluble in non-ionic detergents. Since all of the recombinant CD36 isoforms would still have a 12–His tag, the underglycosylated, possibly misfolded, CD36 may contaminate the mature protein during purification so neither detergents can be used if the protein is to be used in further structural studies. No detergent to date was found that solubilized the membrane bound CD36N8-10 without solubilizing the underglycosylated form of the protein.
Figure 3.7 - Immunoblot of Sf21 membrane preparations, starting and soluble fraction after incubation with detergent.

In (A) starting fraction wild type CD36 (lane 1) and soluble fraction in 2% OG (lane 2), starting fraction CD36N8-10 (lane 3) and soluble fraction in 2% OG (lane 4). In (B), starting fraction of CD36N8-10 (odd numbered lane), soluble fraction in 2% C_{12}E_{8} (lane 2), soluble fraction in 2% FOS-CHOLINE 9 (lane 4), soluble fraction in 1% TX100 (lane 6), soluble fraction in 2% DM (lane 8), soluble fraction in 2% DDM (lane 10) and soluble fraction in 2% SDS (lane 12. Probed with mAb1955.

3.4.3. Concentrating the purified protein

The volume of combined elutions of purified wild type CD36, totalling 4mls, was reduced to approximately 250µl using Amicon Ultra-15 centrifugal devices (Millipore) (see chapter 2.7.4) and the concentration of the protein was estimated using a BSA standard curve. Figure 3.8 shows a typical colloidal blue stained SDS-polyacrylamide gel with increasing concentrations of BSA and purified protein. The same procedure was used to concentrate CD36N8-10, however, the yield was very low.
Figure 3.8 - SDS-PAGE gel stained with colloidal blue showing increasing BSA concentrations and purified wild type CD36-12His.
0.1µg BSA (lane 1), 0.2µg BSA (lane 2), 0.4µg BSA (lane 3), 0.8µg BSA (lane 4), 1.0µg BSA (lane 5) wild type CD36-12His (lane 6-9), 2µl of combined elution fractions from figure 3.4A (0.05% of total volume; lane 6), 1µl concentrated protein (0.4% of total volume; lane 7), 3µl concentrated protein (1.2% of total volume; lane 8) and 5µl concentrated protein (2% of total volume; lane 9).

The density of BSA bands was measured using ImageJ software (NIH) and the standard curve plotted using GraphPad Prism Version 4 (CA, USA) shown in Figure 3.9. The concentration of the protein was measured using linear equation (Equation 1)

\[ y = mx + c \]  

Equation 1

Where y and x are co-ordinates that satisfy the equation, m is the gradient of the straight line and c is the y axis intercept. Typically, the purified protein can be concentrated to 250µg in a total volume of 250µl. Approximately 30% of protein was lost during this concentration step (from 350µg of protein in the starting material), however the protein purity increased from
83% to 93%. The centrifugal devices used in concentrating the protein had a molecular cut off of 50kDa, so it is likely that proteins small than this were not retained in the concentrated fraction, increasing the purity for CD36.

Figure 3.9 - Typical BSA standard curve measured by the density of pixels from SDS-PAGE stained with colloidal blue.

3.4.4. Deglycosylation of native purified CD36

For further structural studies it would be helpful to be able to deglycosylate the native purified protein, therefore CD36 was treated with PNGaseF for 48 hours. Although PNGaseF did remove some glycans from wild type CD36 (Figure 3.10 (A)), the digest was incomplete despite using excess amounts of the enzyme, and therefore the final sample was heterogeneous. Glycosylation may play a protective role for the extracellular protein domain, therefore, it is possible that the glycans prevent the enzyme from accessing the cleavage site effectively, leading to the incomplete digestion observed. There also may be the possibility of protein degradation making the product of this treatment unsuitable for crystallization trials.
CD36N8-10 treated with PNGaseF did remove the glycosylation (Figure 3.10(B)), possibly making it suitable for future studies such as crystallization if sufficient yield could be obtained.

Figure 3.10 – SDS polyacrylamide gel stained with colloidal blue and immunoblot of CD36 treated with PNGaseF.

In (A), wild type CD36-12His and in (B) CD36N8-10 probed with anti-CD36 antibody. In both (A) and (B) untreated (lane 1), with 1 unit PNGase F (lane 2), with 10 units PNGase F (lane 3), with 50 units PNGase F (lane 4), with 100 units PNGase F (lane 5), with 200 units PNGase F (lane 6), with 1000 units PNGase F (lane 7) and denatured protein with 250 units PNGase F (lane 8). Immunoblot was used for CD36N8-10 due to low protein yield during purification.
3.4.5. **Affinity of acetylated low-density lipoprotein for insect cell purified wild type CD36**

Whilst insect cells recognise faithfully the mammalian glycosylation sites (Hsieh and Robbins, 1984) they do not elaborate on the core glycan paucimannose (Jarvis, 2003). It was therefore important to establish that the CD36 produced in the heterologous host was suitable for analysis. To test this, the ligand binding affinity of BODIPY acetylated low-density lipoprotein (BODIPY ac-LDL) for the purified protein immobilized on plates coated with Ni-NTA was determined (see Section 2.7.10). The data are analysed using GraphPad Prism software version 4.0 and the saturation binding curves were best fitted by Langmuir adsorption equation (Equation 2), which describes the binding of ligand to a single class of binding site as defined below:

\[
B = \frac{B_{\text{max}} \times [L]}{K_d + [L]} \quad \text{Equation 2}
\]

Where B is bound ligand (relative fluorescence units), [L] is the concentration of ligand (µg/ml) and Kd is the concentration of ligand giving half maximal binding and a measure of the affinity of ligand-receptor interaction.

The mean Kd of the receptor purified from insect cells for Ac-LDL +/- S.E.M. was 6.4 +/- 1.5 µg/ml (n=3) (Figure 3.11 A). This was compared to the Kd of the receptor expressed transiently on the surface of mammalian HEK293T cells (see section 5.2.2.9). The Kd for Ac-LDL +/- S.E.M. was 8.3 +/- 1.4 µg/ml (n=3) which is not significantly different to the affinity of the protein purified from insect cells. This indicates purified CD36 is folded correctly, and therefore can be used for further structural studies (Figure 3.11 B).
Figure 3.11 - Interaction of wild type CD36 is with BODIPY Ac-LDL.

In (A), wild type CD36-12His purified from Sf21 insect cells, and in (B) wild type CD36-12His expressed on the surface of HEK293T mammalian cells, adjusted for non-specific binding.
3.5. Discussion

As membrane proteins are often present at low levels in native biological membranes, it is often necessary to express proteins recombinantly to obtain sufficient amounts for further studies. Early attempts to express human wild type CD36 in different strains of *E. coli* resulted in aggregation of the protein in inclusion bodies (Linton, K; personal communication). This is most likely caused by the absence of post-translational modifications (as I shall explore in chapter 5). The intracellular retention of CD36nong in transiently transfected HEK293T cells (and probably in infected Sf21 insect cells) suggests that, specifically, N-linked glycosylation is required for the protein to fold and traffic to the cell surface (see section 5.2.2.1) (although this does not exclude the need for other post-translational modifications such as disulphide bridging). Since the *E. coli* expressed CD36 is unlikely to have the correct fold, an alternative expression system was sought that allowed the extensive post-translational modifications of CD36. This study demonstrates that wild type CD36, CD36N8-10 and CD36nong can all be expressed in Sf21 insect cells via the baculovirus systems Bac-N-Blue and flashBAC, and can be detected, with N-linked glycosylation (for wild type CD36 and CD36N8-10), in crude membrane preparations generated from infected Sf21s. The Bac-N-Blue system was used early and flashBAC was used later in the study. The yield of purified wild type CD36 using both expression systems was similar, however flashBAC was the preferable system to use as the titre of the working stock was approximately ten times more than that of Bac-N-Blue.

The mild non-ionic detergent OG was used successfully to solubilize wild type CD36 from the plasma membrane, however, OG failed to solubilize CD36nong. The inability to solubilize CD36nong in any detergent except SDS (Linton, K; personal communication) is indicative that, despite being pelleted with the membrane preparation, the protein is probably
not present within the membrane. CD36nong may possibly be retained intracellularly in large protein aggregates, reminiscent of the inclusion bodies containing underglycosylated CD36 in *E. coli*. Purification attempts on CD36nong extracted using SDS yielded very little protein (SDS is known to inhibit Ni-NTA binding), but since the protein is almost certainly incorrectly folded due to the lack of N-linked glycosylation, no further attempts were made to optimise purification.

OG also failed to solubilize the majority of CD36N8-10 contained in the membrane preparation. A range of non-ionic detergents were used to try to solubilize CD36N8-10, together with the Zwitterion detergent FC-9® and the ionic detergents TX100 and SDS. The only detergents that managed to solubilize the protein were SDS and FC-9®, which are both stronger than non-ionic detergents. Since these two detergents were able to solubilize CD36 isoforms which are otherwise insoluble when non-ionic detergents were used, there is a possibility that CD36N8-10 is not expressed on the cell surface. However, when CD36N8-10 is transiently expressed in mammalian HEK293T cells, occupancy of glycosylation sites N8, N9 and N10 is sufficient for a small amount of the protein to reach the plasma membrane (approximately 12% of wild type expression in mammalian cells) (see section 5.2.2.4). Glycosylation of all or some of these sites also occurs on CD36N8-10 purified from insect cells, as demonstrated by an increase in electrophoretic mobility when the protein is treated with PNGaseF so, like mammalian cells, some of the protein is likely to have reached the cell surface. A small amount of CD36N8-10 was purified from insect cell membrane; however, it is not known whether this purified protein has the correct fold as the yield was too low for further functional studies. Taking this into consideration, the expression of the protein on the cell surface is likely to be very low, so the majority of CD36N8-10 may be contained in intracellular protein aggregates much like CD36nong, which is why it can only be solubilized in strong detergents. It is therefore unlikely CD36N8-10 will be useful in further structural studies. Wild type CD36 was successfully purified from Sf21 insect cells. From a starting
concentration of approximately 1.9mg, a yield 250µg was achieved, and importantly, this was approximately 93% pure. The majority of the 85% loss occurred during the purification procedure, in particular the washes, where about 60% of wild type CD36-12His was lost. Although this yield is reasonable and enabled further functional studies, it may be improved by changing the protein and Ni-NTA resin ratio, and concentration of imidazole in the washes.

Baculovirus expression vectors are widely used to produce recombinant proteins in insect cells, however, there is a fundamental difference in mammalian and insect cell glycosylation processing pathway with insect cells generating less complex glycosylation than a protein generated in mammalian cells. Insect cells express the enzymes involved in N-glycan trimming but no enzymes that elongate N-glycans in mammalian cells so generate trimmed but minimally elongated structures called paucimannose, which consists of Man$_3$-GlcNAc$_2$ (Jarvis, 2003). Measurement of the affinity of the insect cell derived proteins for Ac-LDL was therefore necessary to confirm that the change in glycosylation did not effect the folding of the protein in insect cells. Importantly, the affinity for Ac-LDL of CD36 purified from insect cells and expressed on the surface of mammalian HEK293T cells was not significantly different indicating that the type of glycan is not important for binding of the modified lipoprotein. The insect cell derived protein is also therefore likely to be folded in a similar manner to the native protein indicating that recombinant wild type CD36 protein expressed in and purified from Sf21 cells is suitable for further studies.

The next chapter describes identifying the occupancy of the 10 putative N-linked glycosylation sites using protein derived from insect cells for mass spectrometry and expressed on the surface of mammalian cells for electrophoretic mobility.
4. **Identifying the occupancy of 10 putative glycosylation sites in CD36**

4.1. **Introduction**

N-linked glycosylation occurs when oligosaccharides are attached to the Asn within the consensus sequence Asn-X-Ser/Thr (where X is any amino acid except Pro), however, the presence of this motif only indicates the possibility of glycosylation and not the certainty. This chapter describes the combination of mass spectrometry on purified protein from insect cells, and gel retardation of mutant proteins expressed in mammalian cells to determine which of the ten putative N-linked glycosylation sites in CD36 are modified by glycosylation.

4.1.1. **Mass Spectrometry**

Mass spectrometry determines the mass of a molecule by measuring the mass/charge ratio (m/z) of its ion. The past decade has seen technological improvements to mass spectrometry allowing it to be used as an analytical tool for proteins, peptides, DNA, drugs and other biologically relevant molecules. For example, the technique can be used for:-

- Accurate molecular measurement to establish a sample’s purity, post-translational modifications and verify amino-acid substitutions
- Monitor enzyme reactions and protein digestion
- Amino-acid sequencing to characterize peptides and identify proteins
Although there are many types of mass analysers available, the basic mechanism is the same. Once the sample is introduced into the instrument it is fragmented into ions in the ionization chamber and accelerated into the mass analyser. The mass analyser deflects and separates the samples according to the m/z ratio, which generates ion currents. The currents are then amplified by the detector before being transferred to a data system (Figure 4.1). The ionisation source, analyser and detector are under high vacuum to ensure the sample is not hindered by air molecules when travelling from one end of the instrument to the other.

![Figure 4.1 - General components of mass spectrometer.](http://www.chemguide.co.uk/analysis/masspec/howitworks.html#top)

A vapourised sample is introduced and ionized in the ionization chamber. The fragments are accelerated by a strong electromagnetic field to a detector, generating ion currents. Deflection is dependant on the mass/charge ratio allowing identification of the fragments.

(http://www.chemguide.co.uk/analysis/masspec/howitworks.html#top)
4.1.2. Liquid Chromatography Tandem mass spectrometry using Quadrupole Time-of-Flight (LC MS/MS Q-ToF)

LC MS/MS Q-ToF (using Electrospray Ionisation method) combines the physical separation of a solution with the mass analysis of a mass spectrometer. The sample is introduced into the mass spectrometry instrument via a capillary column following liquid chromatography, and is expelled from a needle in the presence of a strong electrical field resulting in a spray of charged droplets. During solvent evaporation ions with single or multiple charges are ejected from the droplets and are directed to the mass analyser.

The tandem mass spectrometry analysis combines high sensitivity and accuracy by using Quadrupole and Time-of-Flight analysis (Figure 4.2). Quadrupole is a mass filter with 4 parallel metal rods, and uses an oscillating electrical field to stabilize or destabilize ions which pass through a static radio frequency. This enables ions of a specific m/z value range to pass through the quadrupoles and enter the collision chamber, and in this chamber the selected ions are fragmented by an inert gas such as argon. The Time of Flight analysis accelerates the ions to a detector via a reflectron under vacuum. All the ions are given the same amount of energy, so the time taken for the ion to reach the detector is dependent on the mass, and the reflectron increases the path length leading to a higher resolution than a single time of flight analysis. This robust and powerful mass spectrometry technique thus exploits the ability of quadrupole to select a particular range of ions and the ability of ToF to achieve simultaneous and accurate measurements of ions across the mass range.
After separation by liquid chromatography, electrospray ionization introduces the sample to the quadropole which filters the ions according to the m/z ratio by an oscillating electrical field. The selected ions are fragmented in the collision chamber and the time taken for these fragments to reach the detector is measured allowing identification of the fragments.

(adapted from http://masspec.scripps.edu/mshistory/whatisms_details.php#massanalyzers)

4.1.3. Fourier-transform ion cyclotron resonance (FT-ICR) mass spectrometry introduction

FT-ICR is commonly considered as the most complex and most sensitive method of mass analysis and is able to detect a very small mass difference with very high precision. The technique is based on monitoring the orbiting motion of a charged particle in a magnetic field. The ions are generated in the ionization source and enter the ion trap under ultra high
vacuum conditions where the temperature is close to absolute zero (Figure 4.3). The ions pass through a magnetic field and bend in a circular motion that is perpendicular to the magnetic field. At this stage the radius of the ions’ orbit is small and cannot be detected, so an excitation pulse is applied to excite the electrons to a higher orbit and enlarge the radius, inducing an alternating current between the two detector plates. The frequency of this current is equal to the cyclotron frequency of the ions and the intensity is proportional to the number of ions. As the radio frequency goes off resonance for a particular m/z value the ions drop to their natural orbit and a different m/z range is excited.

Figure 4.3 - Fourier transform ion cyclotron analyser.

Ions are generated and enter the ion trap under ultra high vacuum. The ions bend in a circular motion when they enter the high magnetic field and this orbit is increased upon induction of an excitation pulse. This induces a current between the two detector plates which can be measured and the m/z of the fragment identified.

(adapted from http://massspec.scripps.edu/mshistory/whatisms_details.php#massanalyzers)
4.2. Enzymatic deglycosylation of wild type CD36

In preparation for mass spectrometry, approximately 10pmol of purified wild type CD36 was deglycosylated using PNGase F. PNGase F cleaves all asparagine-linked complex, hybrid or high mannose oligosaccharides except in the unusual circumstances of alpha (1-3) core fucosylation, found as a plant glycan. PNGase F will also fail to cleave if the asparagine to which the oligosaccharide is linked is the initial or terminal amino acid (Trimble and Tarentino, 1991). During PNGase F digestion, the asparagine is converted to aspartic acid resulting in a mass shift of +1 for each deamidation (Figure 4.4(A) compared to the predicted molecular mass of the protein backbone containing a non-utilized asparagine. Figure 4.4(B) shows a typical SDS gel stained with colloidal blue showing the products of PNGase F-treated wild type CD36. The mass difference of 20kDa suggests that several of the putative N-linked glycosylation sites are occupied. Purified CD36nong (see section 3.4.2.2) was used as a control to ensure wild type CD36 treated with PNGaseF was fully deglycosylated. Wild type CD36 and deglycosylated wild type CD36 bands were excised from the gel and prepared for Q-ToF MS/MS and FT-ICR.
Figure 4.4 - PNGaseF enzyme reaction and deglycosylation of CD36.

(A) PNGaseF cleaves the oligosaccharide from the protein backbone, deamidating Asn to Asp.

(B) SDS gel stained with colloidal blue showing wild type CD36 (lane 1), wild type CD36 treated with PNGaseF (lane 2) and CD36nong (lane 3)
4.3. Preparation of protein for mass spectrometry

Protein preparation and Q-ToF mass spectrometry was carried out by MRC CSC Proteomics facility. Preparation and FT-ICR mass spectrometry was carried out by Functional Genomics and Proteomics unit at University of Birmingham.

Briefly, following excision from the SDS gel, the protein was digested using trypsin, yielding peptide fragments of known sizes, seven of which contain the putative N-linked glycosylation sites (Table 4.1). The digested protein was extracted from the gel using 0.1% formic acid and analysed by mass spectrometry.

<table>
<thead>
<tr>
<th>mass</th>
<th>Amino acid number</th>
<th>peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>6950.4468</td>
<td>101-164</td>
<td>ENVTQDAEDNTVSFLQPNGA FEPSLSVGTEADNFTVNLAVAAASHIYQNQFYQMILNSLINK</td>
</tr>
<tr>
<td>3876.8827</td>
<td>335-368</td>
<td>EGRPYYISLPFYASPDVSEPIDGLNPNEEEHR</td>
</tr>
<tr>
<td>3351.6823</td>
<td>184-213</td>
<td>DPFLSLVPYPYTTTVGLFYPY\text{N}NTADGVYK</td>
</tr>
<tr>
<td>3318.7558</td>
<td>438-467</td>
<td>INLGLIEMILLSVGVMFVAFMISYACR</td>
</tr>
<tr>
<td>3022.7195</td>
<td>6-36</td>
<td>NCGLIAGAVIGAVLAVFGGLMPVGDLLIQK</td>
</tr>
<tr>
<td>2958.3073</td>
<td>235-260</td>
<td>NL\text{S}YWESHD\text{M}INGTD\text{A}SFPPFVEK</td>
</tr>
<tr>
<td>2782.3327</td>
<td>64-86</td>
<td>QFWFDVQNPQEVMM\text{N}SSNIQVK</td>
</tr>
<tr>
<td>2175.1332</td>
<td>408-426</td>
<td>NYIVPILWLN\text{E}STG\text{T}GDEK</td>
</tr>
<tr>
<td>2034.8572</td>
<td>299-316</td>
<td>AFASPVENPDNYCFCTEK</td>
</tr>
<tr>
<td>1957.0317</td>
<td>369-385</td>
<td>TYLDIEPITGFTLQFAK</td>
</tr>
<tr>
<td>1878.8153</td>
<td>473-487</td>
<td>GHHHHHHHTGHHHHHHH</td>
</tr>
<tr>
<td>1529.7417</td>
<td>261-273</td>
<td>SQVLQFFSSDICR</td>
</tr>
<tr>
<td>Score</td>
<td>Start-Stop</td>
<td>Sequence</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
<td>-----------</td>
</tr>
<tr>
<td>1484.763</td>
<td>274-286</td>
<td>SIYAVFESDVNLK</td>
</tr>
<tr>
<td>1367.8256</td>
<td>387-398</td>
<td>LQVNLLVPSEK</td>
</tr>
<tr>
<td>1333.7361</td>
<td>41-52</td>
<td>QVVLEEGTIAFK</td>
</tr>
<tr>
<td>1299.6249</td>
<td>321-332</td>
<td>NCTSYGVLDISK</td>
</tr>
<tr>
<td>936.4937</td>
<td>177-183</td>
<td>ELLWGYR</td>
</tr>
<tr>
<td>922.5244</td>
<td>224-231</td>
<td>VAIIDTYK</td>
</tr>
<tr>
<td>854.4189</td>
<td>167-173</td>
<td>SSMFQVR</td>
</tr>
<tr>
<td>825.4101</td>
<td>57-63</td>
<td>TGTEVYR</td>
</tr>
<tr>
<td>756.3675</td>
<td>89-94</td>
<td>GPYTYR</td>
</tr>
<tr>
<td>704.4090</td>
<td>287-292</td>
<td>GIPVYR</td>
</tr>
<tr>
<td>690.4185</td>
<td>293-298</td>
<td>FVLPSK</td>
</tr>
<tr>
<td>638.3079</td>
<td>427-431</td>
<td>ANMFR</td>
</tr>
<tr>
<td>619.3409</td>
<td>432-437</td>
<td>SQVTGK</td>
</tr>
<tr>
<td>600.4079</td>
<td>399-403</td>
<td>IQVLK</td>
</tr>
<tr>
<td>581.2170</td>
<td>1-5</td>
<td>MGCDR</td>
</tr>
<tr>
<td>576.2987</td>
<td>219-223</td>
<td>DNISK</td>
</tr>
<tr>
<td>564.3140</td>
<td>214-218</td>
<td>VFNGK</td>
</tr>
<tr>
<td>546.3034</td>
<td>53-56</td>
<td>NWVK</td>
</tr>
<tr>
<td>478.3024</td>
<td>97-100</td>
<td>FLAK</td>
</tr>
<tr>
<td>460.3129</td>
<td>317-320</td>
<td>IISK</td>
</tr>
<tr>
<td>389.2507</td>
<td>174-176</td>
<td>TLR</td>
</tr>
<tr>
<td>374.2398</td>
<td>404-406</td>
<td>NLK</td>
</tr>
<tr>
<td>361.2445</td>
<td>37-39</td>
<td>TIK</td>
</tr>
<tr>
<td>361.2445</td>
<td>470-472</td>
<td>TIK</td>
</tr>
<tr>
<td>303.1775</td>
<td>87-88</td>
<td>QR</td>
</tr>
</tbody>
</table>
Table 4.1 – Complete tryptic digestion of histidine-tagged wild type CD36 results in 46 fragments.

If all 10 putative N-linked glycosylation sites are occupied, the 7 fragments that are italicised will increase in mass by 1 for each deamidation of a glycosylated asparagine to aspartic acid following treatment with PNGaseF. The putative glycosylated asparagines are underlined and in bold.

4.4. LC/MS/MS Q-ToF analysis results

In Q-ToF MS/MS the trypsin-treated fragments are ionised to form amongst others different y-series C-terminal and b-series N-terminal ions, and analysis of the mass and charge of these ions determine the amino acid sequence of the peptide (Figure 4.5). Using this technique, 20% coverage of wild type CD36 treated with PNGase F was achieved. 6 of the 10 N-linked glycosylation sites (N1, N5, N7, N8, N9 and N10) were deamidated from asparagine to aspartic acid and therefore were glycosylated. (Figure 4.6A-E). For example, tryptic fragment 235-NLSWESHCDMINGTDAASFPFVEK-260 contains two putative glycosylation sites N7
(N235) and N8 (N247). As N7 is at the amino terminus of the fragment, the b-ions determine whether this asparagine has been deglycosylated and deamidated to aspartic acid (Figure 4.6A). The b-ions with m/z 229.12 (b2) and m/z 316.15 (b3) are shown on the spectrum (top panel) and both are diagnostic for aspartic acid at position 235 indicative of modification of the wild type Asn (N7) by glycosylation. Putative glycosylation site N8 is located in the middle of the tryptic fragment, therefore both b-ions and y-ions can determine whether this site is occupied. The b-ion with m/z 1709.66 (b14) is diagnostic for aspartic acids in positions 235 and 247 and the y-ions with m/z 1593.78 (y15) and 1480.70 (y14) indicate an aspartic acid in position 247 indicative that N8 is glycosylated in wild type CD36 (Figure 4.6A, bottom panel).

Similarly, diagnostic data for sites N1, N5, N9 and N10 (Figures 4.6B, C, D, E respectively) showed that each of these sites are also modified by glycosylation. An example of a full spectrum is shown in Figure 4.6B top panel for fragment containing glycosylation site N1.

Analysis of wild type CD36 in the absence of PNGase F treatment failed to detect any fragments containing the putative N-linked glycosylation sites, although other fragments were detected. This is consistent with glycosylation, as the N-glycans increase the size of the peptides. Also, fragments with asparagines that are not within the glycosylation consensus sequence showed little evidence of a similar mass change, indicating spontaneous deamidation is rare. For example, N408 (Figure 4.6E) is not contained within a putative glycosylation site. The b-ion with m/z 391.20 (b3) is diagnostic for asparagine at this position and not aspartic acid. Unfortunately, the fragment containing N2, N3 and N4 was not detected which is most likely due to a low charge state for the peptide, which has only one internal histidine, precluding an m/z value lower than 2000. The fragment containing N6 was also not detected probably because the peptide is too small for this technique, therefore, a more sensitive technique FT-ICR was employed to detect these peptides.
Figure 4.5 - Following tryptic digestion, protein fragments are ionized into related y and b-series protonated ions.

A selection of b-ions (green) and corresponding y-ions (blue) are shown for the tryptic fragment containing N7 and N8.
### Figure 4.6 - Q-ToF MS spectra showing that N1, N5, N7, N8, N9 and N10 of CD36 are glycosylated.

Following tryptic digestion, protein fragments are ionised into related y and b-series ions, each with a unique mass/charge detectable by Q-ToF MS. In A-E, selected portions of the Q-ToF MS spectra of the PNGase F and trypsin-digested wildtype CD36 are shown. The spectra show the diagnostic protonated ions which demonstrate that N7 and N8 (A), N1 (B) (with full spectra which includes fragments with different charges), N5 (C), N9 (D) and N10 (E) are deglycosylated and deamidated to aspartic acids by treatment with PNGase F, resulting in a mass shift of +1. The sequence of the tryptic digest fragments with the putative glycosylated asparagines underlined, and the complete y and b-series ions generated by the ionisation of the fragment are shown. Those fragments detected by Q-ToF MS are shown in bold. Fragments not detected are shown in italics.
<table>
<thead>
<tr>
<th>#</th>
<th>b-ions (m/z)</th>
<th>sequence</th>
<th>y-ions (m/z)</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>129.07</td>
<td>Q</td>
<td></td>
<td>23</td>
</tr>
<tr>
<td>2</td>
<td>276.13</td>
<td>F</td>
<td>2655.26</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>462.21</td>
<td>W</td>
<td>2508.19</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>575.30</td>
<td>I</td>
<td>2322.11</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>722.37</td>
<td>F</td>
<td>2209.03</td>
<td>19</td>
</tr>
<tr>
<td>6</td>
<td>837.39</td>
<td>D</td>
<td>2061.91</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>936.46</td>
<td>V</td>
<td>1946.93</td>
<td>17</td>
</tr>
<tr>
<td>8</td>
<td>1064.52</td>
<td>Q</td>
<td>1847.86</td>
<td>16</td>
</tr>
<tr>
<td>9</td>
<td>1178.56</td>
<td>N</td>
<td>1719.80</td>
<td>15</td>
</tr>
<tr>
<td>10</td>
<td>1275.62</td>
<td>P</td>
<td>1605.76</td>
<td>14</td>
</tr>
<tr>
<td>11</td>
<td>1403.67</td>
<td>Q</td>
<td>1508.71</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>1532.72</td>
<td>E</td>
<td>1380.65</td>
<td>12</td>
</tr>
<tr>
<td>13</td>
<td>1631.79</td>
<td>V</td>
<td>1251.61</td>
<td>11</td>
</tr>
<tr>
<td>14</td>
<td>1762.83</td>
<td>M</td>
<td>1152.54</td>
<td>10</td>
</tr>
<tr>
<td>15</td>
<td>1893.87</td>
<td>M</td>
<td>1021.50</td>
<td>9</td>
</tr>
<tr>
<td>16</td>
<td>2008.89</td>
<td>D (N)</td>
<td>890.46</td>
<td>8</td>
</tr>
<tr>
<td>17</td>
<td>2095.93</td>
<td>S</td>
<td>775.43</td>
<td>7</td>
</tr>
<tr>
<td>18</td>
<td>2182.96</td>
<td>S</td>
<td>668.40</td>
<td>6</td>
</tr>
<tr>
<td>19</td>
<td>2297.00</td>
<td>N</td>
<td>601.37</td>
<td>5</td>
</tr>
<tr>
<td>20</td>
<td>2410.08</td>
<td>I</td>
<td>487.38</td>
<td>4</td>
</tr>
<tr>
<td>21</td>
<td>2538.14</td>
<td>Q</td>
<td>347.38</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>2637.21</td>
<td>V</td>
<td>246.18</td>
<td>2</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>K</td>
<td>147.11</td>
<td>1</td>
</tr>
</tbody>
</table>
C

184-DPFLSLVPYPVTTVGLFYPYNYTADGVYK-213

<table>
<thead>
<tr>
<th>#</th>
<th>b-ions</th>
<th>sequence</th>
<th>y-ions</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>116.03</td>
<td>D</td>
<td>3237.64</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>213.09</td>
<td>P</td>
<td>3140.59</td>
<td>29</td>
</tr>
<tr>
<td>3</td>
<td>360.16</td>
<td>F</td>
<td>2993.52</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>473.24</td>
<td>L</td>
<td>2880.43</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>560.27</td>
<td>S</td>
<td>2793.40</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>673.36</td>
<td>L</td>
<td>2793.40</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>772.42</td>
<td>V</td>
<td>2680.32</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>869.48</td>
<td>P</td>
<td>2581.25</td>
<td>23</td>
</tr>
<tr>
<td>9</td>
<td>1032.54</td>
<td>Y</td>
<td>2484.20</td>
<td>22</td>
</tr>
<tr>
<td>10</td>
<td>1129.59</td>
<td>P</td>
<td>2321.13</td>
<td>21</td>
</tr>
<tr>
<td>11</td>
<td>1228.66</td>
<td>V</td>
<td>2224.08</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td>1329.71</td>
<td>T</td>
<td>2125.01</td>
<td>19</td>
</tr>
<tr>
<td>13</td>
<td>1430.76</td>
<td>T</td>
<td>2023.97</td>
<td>18</td>
</tr>
<tr>
<td>14</td>
<td>1531.80</td>
<td>T</td>
<td>1922.92</td>
<td>17</td>
</tr>
<tr>
<td>15</td>
<td>1680.87</td>
<td>V</td>
<td>1821.87</td>
<td>16</td>
</tr>
<tr>
<td>16</td>
<td>1687.89</td>
<td>G</td>
<td>1722.80</td>
<td>15</td>
</tr>
<tr>
<td>17</td>
<td>1800.98</td>
<td>L</td>
<td>1665.78</td>
<td>14</td>
</tr>
<tr>
<td>18</td>
<td>1948.05</td>
<td>F</td>
<td>1552.70</td>
<td>13</td>
</tr>
<tr>
<td>19</td>
<td>2111.11</td>
<td>Y</td>
<td>1405.63</td>
<td>12</td>
</tr>
<tr>
<td>20</td>
<td>2208.16</td>
<td>P</td>
<td>1242.56</td>
<td>11</td>
</tr>
<tr>
<td>21</td>
<td>2371.23</td>
<td>Y</td>
<td>1145.51</td>
<td>10</td>
</tr>
<tr>
<td>22</td>
<td>2486.25</td>
<td>D(N)</td>
<td>982.45</td>
<td>9</td>
</tr>
<tr>
<td>23</td>
<td>2600.30</td>
<td>N</td>
<td>867.42</td>
<td>8</td>
</tr>
<tr>
<td>24</td>
<td>2701.34</td>
<td>T</td>
<td>753.38</td>
<td>7</td>
</tr>
<tr>
<td>25</td>
<td>2772.38</td>
<td>A</td>
<td>652.33</td>
<td>6</td>
</tr>
<tr>
<td>26</td>
<td>2887.41</td>
<td>D</td>
<td>581.29</td>
<td>5</td>
</tr>
<tr>
<td>27</td>
<td>2944.43</td>
<td>G</td>
<td>466.27</td>
<td>4</td>
</tr>
<tr>
<td>28</td>
<td>3045.30</td>
<td>V</td>
<td>409.25</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>3206.56</td>
<td>Y</td>
<td>310.18</td>
<td>2</td>
</tr>
<tr>
<td>30</td>
<td>147.11</td>
<td>K</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>
### Table 1: b-ions and y-ions for D(N)

<table>
<thead>
<tr>
<th>#</th>
<th>b-ions</th>
<th>sequence</th>
<th>y-ions</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>116.03</td>
<td>D(N)</td>
<td>1242.60</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>276.07</td>
<td>C</td>
<td>1242.60</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>377.11</td>
<td>T</td>
<td>1082.57</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>464.15</td>
<td>S</td>
<td>981.83</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>627.21</td>
<td>Y</td>
<td>894.49</td>
<td>8</td>
</tr>
<tr>
<td>6</td>
<td>684.23</td>
<td>G</td>
<td>731.43</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>783.30</td>
<td>V</td>
<td>674.41</td>
<td>6</td>
</tr>
<tr>
<td>8</td>
<td>896.38</td>
<td>L</td>
<td>575.34</td>
<td>5</td>
</tr>
<tr>
<td>9</td>
<td>1011.41</td>
<td>D</td>
<td>462.26</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>1124.49</td>
<td>I</td>
<td>347.23</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>1211.53</td>
<td>S</td>
<td>234.15</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: b-ions and y-ions for E

<table>
<thead>
<tr>
<th>#</th>
<th>b-ions</th>
<th>sequence</th>
<th>y-ions</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>115.05</td>
<td>N</td>
<td>2062.07</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>278.11</td>
<td>Y</td>
<td>1899.01</td>
<td>18</td>
</tr>
<tr>
<td>3</td>
<td>391.20</td>
<td>I</td>
<td>1785.93</td>
<td>17</td>
</tr>
<tr>
<td>4</td>
<td>490.27</td>
<td>V</td>
<td>1785.93</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>587.32</td>
<td>P</td>
<td>1686.86</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>700.49</td>
<td>I</td>
<td>1589.81</td>
<td>14</td>
</tr>
<tr>
<td>7</td>
<td>813.49</td>
<td>L</td>
<td>1476.72</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>999.57</td>
<td>W</td>
<td>1363.64</td>
<td>12</td>
</tr>
<tr>
<td>9</td>
<td>1112.65</td>
<td>L</td>
<td>1177.56</td>
<td>11</td>
</tr>
<tr>
<td>10</td>
<td>1227.68</td>
<td>D(N)</td>
<td>1064.47</td>
<td>10</td>
</tr>
<tr>
<td>11</td>
<td>1356.82</td>
<td>E</td>
<td>949.45</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>1457.77</td>
<td>T</td>
<td>820.41</td>
<td>8</td>
</tr>
<tr>
<td>13</td>
<td>1514.79</td>
<td>G</td>
<td>719.36</td>
<td>7</td>
</tr>
<tr>
<td>14</td>
<td>1615.84</td>
<td>T</td>
<td>662.34</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>1728.92</td>
<td>I</td>
<td>561.29</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>1785.94</td>
<td>G</td>
<td>448.20</td>
<td>4</td>
</tr>
<tr>
<td>17</td>
<td>1900.97</td>
<td>D</td>
<td>319.18</td>
<td>3</td>
</tr>
<tr>
<td>18</td>
<td>2030.01</td>
<td>E</td>
<td>276.16</td>
<td>2</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td>147.11</td>
<td>1</td>
</tr>
</tbody>
</table>
4.5. FT-ICR analysis

In FT-ICR, like Q-ToF MS/MS, the trypsin digested fragments are ionised to form amongst others different y-series C-terminal and b-series N-terminal ions (Figure 4.5), and analysis of the mass and charge of these ions determine the amino acid sequence of the peptide. Using this technique 42% coverage of wild type CD36 treated with PNGaseF was achieved, this compares to 20% coverage using Q-ToF confirming FT-ICR was the most sensitive technique, but once again the fragment containing N2, N3 and N4 was not detected. Figure 4.7A demonstrates a partially digested fragment containing putative glycosylation site N6 (N220) and identified that this glycosylation site can be occupied with an glycan. However, not all fragments exhibited this mass shift indicating that N6 modification is unlikely to be comprehensive (Figure 4.7B).

<table>
<thead>
<tr>
<th>#</th>
<th>b-ions</th>
<th>sequence</th>
<th>y-ions</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>100.07</td>
<td>V</td>
<td></td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>247.14</td>
<td>F</td>
<td>1927.02</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>361.19</td>
<td>N</td>
<td>1779.95</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>418.21</td>
<td>G</td>
<td>1665.91</td>
<td>15</td>
</tr>
<tr>
<td>5</td>
<td>546.30</td>
<td>K</td>
<td>1608.88</td>
<td>14</td>
</tr>
<tr>
<td>6</td>
<td>661.33</td>
<td>D</td>
<td>1480.79</td>
<td>13</td>
</tr>
<tr>
<td>7</td>
<td>776.36</td>
<td>D(N)</td>
<td>1365.76</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>889.44</td>
<td>I</td>
<td>1250.74</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>976.47</td>
<td>S</td>
<td>1137.65</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>1104.57</td>
<td>K</td>
<td>1050.62</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>1203.64</td>
<td>V</td>
<td>922.52</td>
<td>8</td>
</tr>
<tr>
<td>12</td>
<td>1274.67</td>
<td>A</td>
<td>823.46</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>1387.76</td>
<td>I</td>
<td>752.42</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>1500.84</td>
<td>I</td>
<td>639.33</td>
<td>5</td>
</tr>
<tr>
<td>15</td>
<td>1615.87</td>
<td>D</td>
<td>526.25</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>1716.92</td>
<td>T</td>
<td>411.22</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>1879.98</td>
<td>Y</td>
<td>310.18</td>
<td>2</td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>K</td>
<td>147.11</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 4.7 – FT-ICR MS spectra showing that N6 of CD36 both occupied and unoccupied by N-glycan.

Following tryptic digestion, protein fragments are ionised into related y and b-series ions, each with a unique mass/charge detectable by FT-ICR MS. In A and B, selected portions of the FT-ICR MS spectra of the PNGase F and trypsin-digested wildtype CD36 are shown. The spectra ion and diagnostic protonated ions which demonstrate that N6 (A) (which includes fragments with different charges) are deglycosylated and deamidated to aspartic acids by treatment with PNGase F, resulting in a mass shift of +1, however, (B) describes a fragment in which N6 has not been deamidated to aspartic acid. The sequence of the tryptic digest fragments with the putative glycosylated asparagines underlined, and the complete y and b-series ions generated by the ionisation of the fragment are shown. Those fragments detected by FT-ICR MS are shown in bold. Fragments not detected are shown in italics.
4.6. Detecting glycosylation sites by different electrophoretic mobility of mutant proteins

Since the fragment containing N2, N3 and N4 was not detected by either of the mass spectrometric techniques, a different assay was used to establish the glycosylation status of these putative sites and also to corroborate the incomplete utilization of N6. This involves comparing the electrophoretic mobility of CD36 isoforms with and without the test sites. As mammalian cells, unlike insect cells, are able to synthesize mature glycosylation complexes and the addition or removal of these structures may be more evident by the electrophoretic technique, mammalian HEK293T cell line was used to express the different mutants. MS established that CD36 is heavily glycosylated with at least 7 of the 10 putative sites occupied, so removing a single site of interest from wild type CD36 may result in a change that is hard to visualise by immunoblotting, particularly given that glycosylation may not be complete or the glycan matured in every molecule (which often results in glyco-proteins migrating as a smear). Reasoning that the difference in electrophoretic mobility would be more evident in a less complex minimally-glycosylated template, the sites of interest were individually introduced into CD36N8-10, which contains only the three carboxy terminal sites N8, N9 and N10 (see chapter 5.2.2.4 for description of the generation of CD36N8-10). Five CD36 isoforms based on this template were generated, CD36N2,8-10; CD36N3,8-10; CD36N4,8-10; CD36N6,8-10 and CD36N7,8-10.

The coding sequence for these CD36 isoforms were transfected into HEK293T cells using the plasmid vector pCIneo, and the cells were harvested 48 hours post transfection (see section 2.5.2). The cell pellets were lysed using SDS (see section 2.7.6) and the whole cell lysates were analysed by SDS-PAGE and immunoblotting. If the site of interest introduced to CD36N8-10 is glycosylated, the electrophoretic mobility of this protein would be slower than
that of CD36N8-10. As a positive control, CD36N7-10 was used since N7 was identified as occupied by Q-ToF MS/MS in the wild type protein (see section 4.4). Analysis of electrophoretic mobility (Figure 4.8, top panel) found CD36N7,8-10 (lanes 2 and 8), CD36N2,8-10 (lane 9), CD36N3,8-10 (lane 4) and CD36N6,8-10 (lane 6) migrate slower than CD36N8-10 (lanes 1, 3 and 7) suggesting that N2, N3 and N6 are all glycosylated, however, there is no evidence that N4 is occupied by glycosylation (lane 5). Analysis of the electrophoretic mobility following treatment with PNGase F (Figure 4.8, bottom panel) found all the CD36 isoforms migrate similarly indicating the glycans on N2, N3 and N6 are responsible for the change in mobility. The inclusion of N6 in CD36N8-10 results in partial glycosylation as judged by the multiple bands (lane 6, top panel), which is resolved following treatment with PNGase F (lane 6, bottom panel).

![Figure 4.8 - Electrophoretic mobility shift analysis to test glycosylation status of CD36 probed with mAb1955.](image)

Using CD36N8-10 as a template in lanes 1, 3 and 7; the glycosylation sites of interest were introduced, generating CD36N7,8-10 (lane 2 and 8); CD36N3,8-10 (lane 4); CD36N4,8-10 (lane 5), CD36N6,8-10 (lane 6) and CD36N2,8-10 (lane 9). The bottom panels are samples as per top panel following deglycosylation by PNGase F.
To establish whether the occupancy of N4 was dependent upon the presence of glycans on other glycosylation sites, the electrophoretic mobility was compared between wild type CD36 and CD36N1-3,5-10, where N4 is disrupted by mutagenesis (Figure 4.9). Although there appeared to be no change in migration between wild type CD36 and CD36N1-3,5-10 (lanes 1 and 4), consistent with N4 not being glycosylated, there was also little change when either N2 or N8 was removed from wild type CD36 (lanes 2 and 3). Since this study previously confirmed occupancy of these two glycosylation sites, the extensive glycosylation of wild type CD36 does not allow a change in migration to be visualized by immunoblotting when one glycan has been removed, as previously hypothesised.

Figure 4.9- Electrophoretic mobility shift analysis to test occupancy of N4 in wild type CD36 probed with mAb1955.

Using wild type CD36 as a template in lane 1; the glycosylation sites of interest were removed, generating CD36N1-7,9,10 (lane 2); CD36N1,3-10 (lane 3); and CD36N1-3,5-10 (lane 4).
4.7. Discussion

Biophysical analysis was used to assess which of the 10 putative N-linked glycosylation sites (characterized by the motif Asn-X-Ser/Thr) were occupied with oligosaccharides. The combination of mass spectrometric techniques on protein purified from insect cells (Q-ToF MS/MS and FT-ICR) and analysis of electrophoretic mobility of various CD36 mutant proteins, in HEK293T whole cell lysates confirmed 9 of the 10 putative sites were occupied (N1, N2, N3, N5, N6, N7, N8, N9 and N10). No evidence was found for glycosylation of site N4, which is perhaps unsurprising since this is the only site that is not conserved in other CD36 orthologues (Table 4.2).

By separating trypsin treated fragments of purified bovine Cd36 and assaying for the presence for amino sugars, Bergland et al identified all 8 putative N-linked glycosylation sites were occupied in this species (Berglund et al., 1996). Of these sites, 7 are conserved and occupied in human CD36 (corresponding to human sites N1, N2, N5, N7, N8, N9 and N10) (Figure 4.9).

For the related murine scavenger protein SR-BI, Viñals et al generated a collection of mutant SR-BI expression vectors, each with one of the 11 asparagines contained within the glycosylation consensus sequence, mutated to glutamine. These, along with the wild type SR-BI transiently expressed in COS M6 cells, and the electrophoretic mobility of the proteins were compared by SDS-PAGE. All 11 mutants showed a subtle change in mobility compared to the wild type protein (Vinals et al., 2003).
Position of Asn in human CD36

|------|-----|------|------|------|------|------|------|------|------|------|

*Including hamster Cd36 where the site is displaced by 2 amino acids

Table 4.2 - The conservation of human putative N-linked glycosylation sites in members of CD36 family including mouse, rat, hamster and bovine

Four of the SR-BI sites are conserved in human CD36 (using human CD36 nomenclature these correspond to N2, N5, N6 and N8; Figure 4.9). Considering SR-BI amino acid sequence is 33% identical to human CD36 and 54% similar, it is possible the 36% conservation of N-linked glycosylation sites between these two homologues is random and need not suggest the importance of these oligosaccharides for function. However, Viñals et al identified, by antibody binding to an extracellular epitope of SR-BI, and flow cytometry of COS M6 cells transiently expressing mutant forms of SR-BI, that glycosylation sites N108 and N173 were necessary for cell surface expression of the protein (Vinals et al., 2003). Neither of these sites are conserved in human CD36 (although N173 is conserved in bovine Cd36) Conversely, N2, N5 and N8 of human CD36 are conserved in bovine, rat and mouse CD36 (N6 is not conserved in bovine) and mSR-BI. The conservation of these sites throughout the family may be indicative of the importance of these sites for folding and function; however, this will need further analysis (chapter 5).
Figure 4.9 – Primary sequence alignment of CD36 orthologues.

Sequences are aligned from *Homo sapiens* (hCD36), *Bos taurus* (bCd36) and *Rattus norvegicus* (rCd36) and *Mus musculus* SR-BI (mSR-BI) with the putative N-linked glycosylation sites highlighted in yellow. The ten sites in human CD36 are indicated...
Whilst it cannot be ruled out that inclusion of N6 may influence the occupancy of sites N8-10 in the electrophoretic mobility study, the simplest interpretation, and one that is consistent with the FT-ICR data on the wild type protein, is that N6 itself is recognised inefficiently as a glycosylation site. The mass spectrometry and gel retardation data are therefore consistent and suggest that site N6 can be glycosylated, however, modification is unlikely to be comprehensive.

The efficiency of glycosylation can be influenced by surrounding amino acids including whether the third amino acid in the sequon is serine or threonine (Kasturi et al., 1997). By generating variants of the rabies virus glycoprotein, with different amino acids at the Y position (within the glycosylation sequence Asn-Leu-Ser/Thr-Y, where Y is any amino acid), and expressing this protein in a cell free system, Mellquist et al established that the sequence Asn-Leu-Ser-Lys confers 50% core glycosylation efficiency (Mellquist et al., 1998). This sequon is similar to that found at both N6 (Asn-Ile-Ser-Lys) and the unoccupied N4 (Asn-Lys-Ser-Lys), and so the presence of the lysine immediately C-terminal to the motif maybe the reason why N6 in CD36 is only partially glycosylated and N4 is not glycosylated. However, Shakin-Eshleman et al identified, using the same expression system, that the amino acid occupying the X position (between the asparagine and the serine/threonine) may also affect core glycosylation efficiency (Shakin-Eshleman et al., 1996). Data from this study suggest that Leu at X confers to 45% and both Lys and Ile, 85% core glycosylation efficiency. There, however is no comparison in either study that directly corresponds to the N4 or N6 sequon. There are also likely to be other factors affecting glycosylation efficiency such as the accessibility of the sequon and the N-glycans to various glycosidases and glycosyltransferases (Roth, 2002).

Wojczyk et al, has suggested the occupancy of a N-linked glycosylation site maybe dependent on the presence of other N-linked glycans. This may be relevant when considering occupancy of site N4 (when studied in a protein lacking glycosylation sites N1, N2, N3, N5,
N6 and N7). The disruption of occupied glycosylation sites (N2 and N8) from wild type CD36 resulted in a migration change that could not be visualised by immunoblotting, therefore electrophoretic mobility was not suitable to assess the occupancy of N4 in wild type CD36. However, the cell surface expression of CD36N1-3,5-10 (lacking site N4) was very similar to wild type CD36 (see section 5.2.2.6), inconsistent with other CD36 mutants where the disruption of one occupied glycosylation site resulted in a dramatic reduction in cell surface expression. This suggests that either N4 is not involved in the folding or trafficking of the protein, or the simplest interpretation, given the non-occupancy N4 when the site was restored in CD36N8-10, that N4 is not utilized.

Having established that glycosylation status of human CD36, the next chapter describes the role the N-linked glycosylation sites and N-glycans have in the folding and trafficking of human CD36.
5. The role of glycans in trafficking and folding of CD36

5.1. Introduction

N-linked glycosylation is the most common post-translational modification and has been studied extensively in many membrane proteins. Many different functions have been ascribed to this type of modification such as protection from proteases and enhanced solubility of the protein, regulation of protein turnover, mediation of interactions with ligands and facilitation of folding and trafficking of a protein. This chapter explores the role of N-linked glycosylation specifically in relation to the folding and trafficking of CD36. In brief, the cDNA for CD36 was mutated to generate multiple mutant forms with different glycosylation statuses. These were expressed transiently in HEK293T cells and the cell surface expression was analysed using flow cytometry. For selected mutants the binding of ac-LDL was analysed using solid-phase ligand binding assay.

5.1.1. The mammalian expression vector, pCI-neo

Since a large number of mutant CD36 proteins are to be characterised, a high throughput transient transfection system was chosen to express the proteins. The coding sequence for wild type CD36 was previously subcloned into the bi-functional prokaryotic and mammalian vector pCI-neo (Promega), generating the plasmid pCD36-12His (see section 2.3.1 and appendix). This can be used to produce functional CD36, with a twelve histidine tag at the carboxy terminus, in mammalian cells and is also used as a template for mutagenesis of the
CD36 gene and expression of mutant CD36. In bacterial cells, the vector is maintained at a high copy number and, as it contains the β-lactamase coding sequence, confers ampicillin resistance. In a mammalian host, the cloned cDNA is under the control of the human cytomegalovirus (CMV) immediate-early enhancer/promoter which ensures strong and constitutive expression. In addition, a chimeric intron is located downstream of the CMV enhancer/promoter. Studies have shown that the presence of such intronic sequences flanking a target gene can frequently increase expression levels (Buchman and Berg, 1988). The pCI-neo vector also contains an SV40 late polyadenylation signal downstream of the multiple cloning region (MCR). This enhances RNA stability and increases levels of translation (Bernstein and Ross, 1989; Carswell and Alwine, 1989). An additional feature is the neomycin resistance gene, which encodes neomycin phosphotransferase. This can be used to select stably-transfected cell lines, but is not used in this study. All CD36 proteins encoded by the plasmid contain a twelve histidine tag.

5.1.2. Expression of CD36 using a mammalian cell system

All CD36 protein described throughout this chapter are encoded by the human CD36 cDNA, expressed from the pCI-neo expression vector. In order to maintain the glycosylation and post-translational modifications as close as possible to normal physiology, a human cell line (DuBridge et al., 1987) was chosen in which to express the protein. The human embryonic kidney (HEK293T) cells do not endogenously express CD36, as analysed by western blot and flow cytometry (Figure 5.5), and they possess the SV40 large T antigen, allowing maximal expression from the pCI-neo vector.
5.1.3. Polyethylenimine transfection

The branched polycation PEI was first identified as a transfection agent by Boussif and co-workers (Boussif et al., 1995). It generates a very high efficiency of transfection, thought to arise from the ability of DNA/PEI complexes within endosomes to avoid the degradative lysosomal pathway (Akinc et al., 2005). It has been shown previously to transfect HEK293T achieving close to 100% transfection efficiency (Dixon et al., 2000). As a quick and inexpensive transfection reagent, PEI is suitable for use with the large number of mutant CD36 cDNAs generated.

5.1.4. Post-transfection use of butyric acid

Butyric acid is a short chain fatty acid which can act to enhance gene expression in cultured cells (Gorman et al., 1983). It is thought to inhibit histone deacetylase enzymes which results in histone hyperacetylation, a state associated with transcriptional activation (Candido et al., 1978; Davie, 2003). The addition of butyric acid to cells post-transfection has also been shown to increase the level of transgene expression (Goldstein et al., 1989; Gorman et al., 1983; Palermo et al., 1991).

5.1.5. Flow cytometric analysis

Flow cytometry is a sensitive and versatile technology that has been utilised in the study of many aspects of cell biology, including hematopathology, cell cycle analysis, analysis of intracellular pH, apoptosis and glucose uptake (Cherlet et al., 1999; Kaleem, 2006; Plasier et
al., 1999; Zou et al., 2005). A key advantage of flow cytometric analysis is the ability to simultaneously measure one or more fluorophores associated with individual cells, often segregated into different subpopulations on the basis of metabolic or physiological states. A schematic diagram of a flow cytometer is show in Figure 5.1.

Figure 5.1 - Schematic diagram of a FACScan flow cytometer (BD Biosciences).

The system consists of five main components: a laser for fluorophore excitation, the flow manifold or flow cell through which cells pass as they are analysed, optical filters to enable detection of light at specific wavelengths, photomultiplier tubes (PMTs) and photodiodes for signal detection (FL-1, FL-2, FL-3 FSC and SSC), and a data processing unit to control instrument setting and record data. Reproduced from (Purvis and Giorgio, 1994)

Briefly, a cell suspension is taken up into the sample stream of the flow cytometer under negative pressure. This sample stream is then mixed with the sheath stream and injected into the flow manifold. Within the flow manifold, the cells are subjected to the laser beam which
excites any fluorophores associated with the cell. Light emitted by the fluorophores pass through a series of filters onto photomultipliers, enabling the measurement of emissions at specific wavelengths. In addition, the forward and sideways-scattering of the excitation light can be used to measure the size, granularity and morphology of each cell that passes through the manifold. All these individual measurements are recorded by the data processing unit, and multiple scattering or fluorescence signals can be used to divide, or ‘gate’ the cells into subpopulations for further analysis or to examine cellular physiology.

The ability to measure fluorescence associated with an individual cell makes flow cytometry a perfect technique to assess cell surface expression of mutant and wild type CD36.
5.2. Results

5.2.1. Cell surface expression of CD36

Transient transfection of cells with wild type or mutant CD36 results in a heterogeneous population of cells with varying levels of surface CD36 expression. For this study it was important to measure the changes in CD36 expression in the presence and absence of N-linked glycosylation. Cells were labelled for surface CD36 expression using monoclonal antibodies against an external epitope of the protein. Flow cytometry was used to measure the level of surface expression gating on 10,000 cells for normal size and granularity to eliminate dead cells and debris from the subsequent analysis (Figure 5.2A). As the cells are transiently transfected, the overall transfection efficiency is determined by the proportion of cells expressing CD36 (Figure 5.3B). Using the transfection reagent PEI, CD36 expression was typically observed in 98% of cells.

The wild type CD36 trace in Figure 5.2B exhibits a biphasic staining pattern, which is typical when using non-saturating amounts of antibody. If the amount of antibody availability is limited, the high expressing population of cells would be saturated, but the low expressing population of cells would have a proportion of cell surface CD36 that is not labelled. Despite using saturating amounts of antibody (see section 5.2.1.1) the biphasic pattern was still observed for HEK293T cells transfected with CD36. It is likely that the population of cells expressing high levels of CD36 would be about to divide at the point of transfection. In this case the nuclear membrane of these cells would disassemble earlier than cells that divide later, allowing the ectopic DNA to enter the nucleus early. The heights of these peaks (indicating cell number) differs from transfection to transfection but the fluorescence intensity of each peak relative to the positive and negative controls (which reflects the level of CD36 expression) remained consistent. It was necessary to use these controls in each
experiment, the positive control confirming the successful transfection and binding of antibody, and the negative control indicating background non-specific fluorescence. To enable comparison between wild type and mutant CD36 expression, the peak with the highest fluorescence intensity was gated and the median calculated in each experiment (green dotted gate, Figure 5.2B).

Figure 5.2 - Flow cytometric analysis of cells.

Untransfected cells and cells transiently expressing CD36 were harvested and labelled with mAb1258 (see below). In (A), cells were gated (pink outline) for normal size (forward scatter; FSC) and granularity (side scatter; SSC). (B) shows cell surface expression of CD36. Transiently transfected cells (red trace) appear as two populations expressing different levels of CD36, as determined by saturating concentrations of mAb1258 antibody (FL-2 fluorescence). For comparison, a single population of untransfected cells is also shown (blue trace). Using data analysis software, the CD36 expressing cells can be gated (as shown by the black bar), indicating the percentage of cells transfected. The green dotted bar indicates the population used for comparison of cell-surface expression between different isoforms.
5.2.1.1. Titration of primary antibodies

Two primary antibodies were used in this study both of which recognise unknown extracellular epitopes of CD36, the mouse monoclonal 1258 and the rat monoclonal 1955. These antibodies were titrated to find the saturating concentration for labelling CD36-expressing HEK293T cells. Cells were transfected with pCD36-12His, as described in section 2.5.2, and split into 7 fractions, each containing $5 \times 10^5$ cells in 50µl total volume. These fractions were then incubated with 0, 0.25, 0.5, 1, 2.25, 3.5 and 5µl of mAb1955 (0.5µg/ml) or mAb1258 (1µg/ml). Untransfected HEK293T cells were used as a negative control. These cells were washed and incubated with 5µl secondary antibody, a polyclonal fragment-antigen binding (F(ab’)2) rabbit anti-rat immunoglobulin (1mg/ml) for mAb1955 or a polyclonal fragment-antigen binding (F(ab’)2) goat anti-mouse immunoglobulin (1mg/ml) for mAb1258 both conjugated to fluorophore R-phycoerythrin (RPE). The fluorescence associated with the cells was then measured using flow cytometry. Fluorescence from antibody non-specifically bound was determined using the untransfected control population, and this median value was subtracted from the expressing cells to give a value for the specific binding of mAb1955 (Figure 5.3A) and mAb1258 (Figure 5.3B).

It was decided from these data that 2µl of both antibodies (corresponding to 1µg for mAb1955 and 2µg for mAb1258) was the optimum saturating concentration to use, as any increase above this level resulted in a minimal increase in fluorescence.
Figure 5.3 - Titration of mAb1955 and mAb1258 binding to HEK293T cells expressing wild type CD36.

$5 \times 10^5$ cells transfected with pCD36-12His were labelled with mAb1955 (0-2.5µg) or mAb1258 (0-5µg) for 30 minutes at 4°C and binding was determined using flow cytometry. In A, the binding curve of mAb1955 to CD36-12His is shown, and in B the binding curve of mAb1258 to CD36-12His is shown. Specific binding to CD36 was calculated by subtracting the median FL-2 fluorescence values from untransfected cells labelled in the same manner.
5.2.1.2. Titration of secondary antibody

The secondary antibody was titrated in a similar fashion to the primary antibody. Cells were transiently transfected with pCD36-12His, before being split into 7 fractions, each with $5 \times 10^5$ cells in a total volume of 50µl. Each fraction was labelled with saturating concentrations of mAb1955 or mAb1258 (2µl of each) as described in section 5.2.1.1 and then incubated with 0, 0.25, 0.5, 1, 2.25, 3.5, 5µl of secondary antibody (rabbit anti-rat for mAb1955 or goat anti-mouse for mAb1258). Median fluorescence from antibody bound non-specifically was determined using the untransfected control population of cells and this value was subtracted from the median fluorescence of CD36-expressing cells to give a value for the specific binding of the secondary antibody, shown in Figure 5.4A and B. These data show the optimum saturation concentration for rabbit anti-rat secondary antibody is 2µl (2µg) and for goat anti-mouse 4µl (4µg) per $5 \times 10^5$ cells in 50µl volume.
Figure 5.4 - Titration of secondary antibody binding to HEK293T cells expressing wild type CD36.

5x10^5 cells transfected with pCD36-12His were labelled with mAb1955 (1µg) or mAb1258 (2µg) for 30 minutes at 4°C. Following washing, cells were incubated with 0-5µl of rabbit anti-rat (for cells labelled with mAb1955) or goat anti-mouse (for cells labelled with mAb1258) secondary antibodies and binding was determined using flow cytometry. (A) shows the binding of rabbit anti-rat to mAb1955 and (B) shows goat anti-mouse binding to mAb1258. Specific binding to CD36 was calculated by subtracting the median FL-2 fluorescence values from untransfected cells labelled in the same manner.
5.2.2. Mutation of N-linked glycosylation sites in CD36

One of the most widely-used approaches for studying the molecular mechanism of proteins and post-translational modification is protein engineering via site-directed mutagenesis, followed by a comparison of the activities of wild type and mutant proteins. The utility of this approach was initially demonstrated in 1982 by Winter et al in a report dissecting the role of specific amino acid residues in catalysis by tyrosyl-tRNA synthetase (Winter et al., 1982). As described in section 2.3.3., site-directed mutagenesis of the ten putative N-linked glycosylation sites (nine of which are occupied) in CD36 was carried out in order to help elucidate the role of the glycans in folding and trafficking of CD36.

5.2.2.1. Generation and cell surface expression of CD36nong

Generation of non-glycosylatable mutant was achieved by site-directed mutagenesis, where the asparagine within each glycosylation consensus sequence was mutated to glutamine within the mammalian expression vector pCI-neo, to generate pCD36nong (see chapter 2). HEK293T cells were transfected transiently with the pCD36 and pCD36nong, and following harvest and labelling with saturating amounts of antibodies against CD36 (see sections 5.2.1.1 and 5.2.1.2), the cell surface expression was measured using flow cytometry. As a negative control, untransfected HEK293T cells were used. The binding of mAb1258 in untransfected cells and cells transfected with cDNA encoding CD36nong was very similar indicating CD36nong does not reach the cell surface (Figure 5.5A). mAb1258 is an antibody that recognises an extracellular loop of CD36 when it is in the correct confirmation, however, it does not recognise denatured protein. Using mAb1258 there is a possibility that CD36nong does get to the cell surface but the protein is not correctly folded so cannot bind the antibody.
therefore a different antibody was also used, mAb1955. mAb1955 also binds to the
extracellular loop of CD36 and recognizes correctly-folded and denatured protein. This is
demonstrated by the binding of antibody to wild type CD36 in flow cytometry (Figure 5.5B,
black trace), and by immunoblotting on denatured CD36 (Figure 5.6C). Figure 5.6B shows
that the binding of mAb1955 is comparable in untransfected cells and cells transfected with
pCD36nong, confirming the result that CD36nong does not reach the cell surface. An
immunoblot on whole cell lysates from the cells used in flow cytometry (Figure 5.5C)
demonstrates that CD36nong transfection was successful and the protein was synthesised but
must be retained intracellularly. Since CD36nong showed very similar antibody binding as
untransfected cells, cells transiently expressing CD36nong were used as a negative control
for all future flow cytometry experiments ensuring these cells had been treated the same as
the cells expressing the protein of interest.
Figure 5.5 - Analysis of HEK293T cells transfected with pCD36nong.

Untransfected cells and cells transiently expressing wild type CD36 and CD36nong were harvested and labelled with mAb1258 and mAb1955 before being analysed by flow cytometry for cell surface expression. In (A), the binding of mAb1258 is shown and in (B) the binding of mAb1955 is shown. Binding to wild type CD36 is shown by the black trace, CD36nong is shown by the blue trace and binding to untransfected cells is shown by the red trace. In (C), the cells were lysed using 2% SDS and the immunoblot on the whole cell lysate, probed with mAb1955 confirmed that the transient transfection of pCD36nong was successful.
Table 5.1 - The cell surface expression of mutant CD36 compared to wild type, measured by flow cytometry.

The percentage of median fluorescence (minus background) as compared to wildtype, of the binding of a saturating concentration of mAb1258. Expression is normalised to 100% for wild type CD36 and 0% for CD36nong. Each mutant was analysed in at least three independent experiments.

<table>
<thead>
<tr>
<th>CD36 mutant</th>
<th>Cell surface expression % (n=3) +/- std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD36</td>
<td>100</td>
</tr>
<tr>
<td>CD36nong</td>
<td>0</td>
</tr>
<tr>
<td>CD36N1-7</td>
<td>2 +/- 0</td>
</tr>
<tr>
<td>CD36N7-10</td>
<td>11 +/- 1.7</td>
</tr>
<tr>
<td>CD36N1-7</td>
<td>2 +/- 0</td>
</tr>
<tr>
<td>CD36N1-6</td>
<td>1 +/- 0</td>
</tr>
<tr>
<td>CD36N1-5</td>
<td>0</td>
</tr>
<tr>
<td>CD36N2-7</td>
<td>0.3 +/- 0.2</td>
</tr>
<tr>
<td>CD36N3-7</td>
<td>0</td>
</tr>
<tr>
<td>CD36N7-10</td>
<td>11 +/- 1.7</td>
</tr>
<tr>
<td>CD36N8-10</td>
<td>12 +/- 2.0</td>
</tr>
<tr>
<td>CD36N7,9,10</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>CD36N7,8,10</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>CD36N7-9</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>CD36N8,9</td>
<td>0</td>
</tr>
<tr>
<td>CD36N8,10</td>
<td>0</td>
</tr>
<tr>
<td>CD36N9,10</td>
<td>0</td>
</tr>
<tr>
<td>CD36N1-7,9,10</td>
<td>63 +/- 14</td>
</tr>
<tr>
<td>CD36N1-8,10</td>
<td>78 +/- 5.5</td>
</tr>
<tr>
<td>CD36N1-9</td>
<td>75 +/- 4.0</td>
</tr>
<tr>
<td>CD36N1-8</td>
<td>34 +/- 2.0</td>
</tr>
<tr>
<td>CD36N1-7,9</td>
<td>30 +/- 1.7</td>
</tr>
<tr>
<td>CD36N1-7,10</td>
<td>24 +/- 2.1</td>
</tr>
<tr>
<td>CD36N1,8-10</td>
<td>31 +/- 4.2</td>
</tr>
<tr>
<td>CD36N2,8-10</td>
<td>59 +/- 5.6</td>
</tr>
<tr>
<td>CD36N3,8-10</td>
<td>31 +/- 3.2</td>
</tr>
<tr>
<td>CD36N4,8-10</td>
<td>13 +/- 1.5</td>
</tr>
<tr>
<td>CD36N5,8-10</td>
<td>60 +/- 5.0</td>
</tr>
<tr>
<td>CD36N6,8-10</td>
<td>31 +/- 3.2</td>
</tr>
</tbody>
</table>
To establish which of the N-linked glycosylation sites are required for the protein to reach the cell surface, the asparagines were reintroduced into pCD36nong. If each occupied site is treated individually, \(9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1\), which is over 350,000 different CD36 mutants would need to be generated for analysis of all the N-linked glycosylation combinations, therefore a more feasible strategy was considered. Initially, the N-linked glycosylation sites were categorized as carboxy-terminal (including N7, N8, N9 and N10) and amino-terminal (N1, N2, N3, N4, N5, N6 and N7), and two mutants were generated, pCD36N7-10 (with glycosylation sites N7, N8, N9 and N10 intact) and pCD36N1-7 (with glycosylation sites N1, N2, N3, N4, N5, N6 and N7 intact). If the proteins encoded by these sequences reached the plasma membrane, the minimal glycosylation required could be narrowed down by mutagenesis, using these templates. Flow cytometry on HEK293T cells transiently expressing CD36N7-10 (Figure 5.6A) and CD36N1-7 (Figure 5.6B) demonstrated that both mutants reached the plasma membrane, but to very different levels. When the cells expressing the highest levels of CD36N1-7, CD36N7-10 and wild type CD36 were gated, CD36N1-7 displayed 2% cell surface expression compared to wild type, and CD36N7-10, 11% (Table 1). The low level of cell surface expression for these mutants suggests that the combination of glycosylation sites from the amino and carboxy terminal is likely to be necessary to achieve high level expression. It is also likely that there are two minimally-glycosylated sets of CD36, the amino-terminal and carboxy-terminal, that support trafficking, unless N7, which is present in both of the mutants, is key for cell surface expression. The amino-terminal and carboxy-terminal glycosylation mutants were therefore characterized separately before systematic analysis of their interaction.
Figure 5.6 - Flow cytometric analysis of HEK293T cells transiently expressing CD36N7-10 and CD36N1-7.

Cells transfected with pCD36, pCD36nong, pCD36N7-10 and pCD36N1-7 were harvested and labelled with mAb1258 before being analysed by flow cytometry for cell surface expression. In (A), cells expressing CD36N7-10 is shown by the green trace and in (B), CD36N1-7 is shown by the red trace. Cells expressing CD36nong are shown by the blue trace, wild type CD36-12His by the black trace.

5.2.2.3. Characterization of the amino-terminal glycosylation set

To assess which of the seven glycosylation sites from CD36N1-7 (of which, N4 is not modified, see section 4.6) were necessary for the protein to reach the plasma membrane, more mutants were generated by systematically removing the asparagines (and replacing them once again with glutamine) from each end of the amino terminal glycosylation set. Using this strategy, it was also possible to determine whether any site is essential for folding and trafficking of the protein, or whether the number of glycosylation sites was the important factor. pCD36N1-6 (with glycosylation sites N1, N2, N3, N4, N5 and N6 intact) and
pCD36N2-7 (with glycosylation sites N2, N3, N4, N5, N6, N7 intact) were transiently transfected into HEK293T cells and the cell surface expression of the proteins analysed by flow cytometry using mAb1258 as previously described. Both CD36N1-6 (Figure 5.7A) and CD36N2-7 (Figure 5.7B) reached the cell surface, however the level of plasma membrane expression was low. The expression level of CD36N1-6 and CD36N1-7 was very similar (Figure 5.7A) implying that occupancy of N7 is not required for cell surface expression of CD36, however the expression of CD36N2-7 is about five fold lower than CD361-7 (Figure 5.7B) (Table 1).

As these newly generated mutants still reached the cell surface, a further asparagine was removed from pCD36N1-6 and pCD36N2-7 to generate pCD36N1-5 (where sites N1, N2, N3, N4 and N5 are intact) and pCD36N3-7 (where sites N3, N4, N5, N6 and N7 are intact). When these were expressed transiently in HEK293T cells and the cell surface expression analysed using flow cytometry, neither of these proteins were found on the plasma membrane (Figure 5.7C and Figure 5.7D). Despite this, both proteins were synthesised and as they migrated more slowly than CD36nong, likely to be glycosylated. This suggests that the proteins are synthesised and enter the trafficking pathway but are retained intracellularly (Figure 5.7 E and F).
Figure 5.7 Analysis of HEK293T cells transiently expressing CD36N1-6, CD36N2-7, CD36N1-5 and CD36N3-7.

Transfected cells were harvested and labelled with mAb1258 before being analysed by flow cytometry for cell surface expression. In (A) and (B), CD36N1-7 expressing cells are the red trace. In (A) cells expressing CD36N1-6 is shown by the green trace, in (B), CD36N2-7; the black trace, in (C), CD36N1-5; the green trace and in (D), CD36N3-7; the green trace. Cells expressing CD36nong are shown by the blue trace. (E) is an immunoblot of whole cell lysates expressing CD36N1-5 (lane 1), and CD36nong (lane 2). (F) is an immunoblot of whole cell lysates expressing CD36N3-7 (lane 1) and CD36nong (lane 2). Both are probed with mAb1955 and confirm successful transfection.
These data could suggest that glycosylation sites N1, N2, N6 and N7 are the important key sites for CD36 to reach the plasma membrane in the amino-terminal glycosylation set.

Generation of the pCD36N1,2,6,7 (where sites N1, N2, N6 and N7 are intact) and analysis of cell surface expression of the protein product in HEK293T indicated that this mutant failed to reach the cell surface (Figure 5.8A). Once again the protein was synthesised and appeared to be glycosylated, but was retained intracellularly (Figure 5.8B). This suggests that these sites can only exert an effect in the context of a more heavily glycosylated protein. Since N4 is not modified by glycosylation (see section 4.6), a correlation is developing between glycosylation site density and trafficking within the amino terminal set. Thus, all mutants with five or more of the utilized amino-terminal glycosylation sites irrespective of combination appear on the cell surface, while mutants with fewer than five utilized sites are unable to support trafficking.

Demonstrating that CD36N1-6 can be expressed on the cell surface and in the knowledge that it shares no common glycosylation sites with CD36N7-10, indicates that none of the glycosylation sites are essential for the protein to traffic to the cell surface. The recognition of cell surface expression for CD36N1-6 and CD36N7-10 also confirms that the glycosylation occupancy does not affect the binding of the antibody mAb1258.
Cells transfected with pCD36, pCD36nong, pCD36N7-10 and pCD36N1-7 were harvested and labelled with mAb1258 before being analysed by flow cytometry for cell surface expression. Cells expressing CD36nongs are shown by the blue trace, wild type CD36 by the black trace. In (A), cells expressing CD36N1,2,6,7 is shown by the red trace and (B) shows an immunoblot of whole cell lysates expressing CD36N1,2,6,7 (lane 1) and CD36nong (lane 2) probed with mAb1955 confirming successful transfection.

5.2.2.4. Characterization of the carboxy-terminal glycosylation set

Using pCD36N7-10 (which encodes a protein that reaches the cell surface, see section 5.2.2.2) as a template, four more mutants were generated to establish which of the carboxy-terminal glycosylation sites were necessary for the protein to fold and traffic to the plasma
membrane (it had already been established that no single site is essential; see section 5.2.2.3). Each mutant was generated by removing one of the four asparagines and replacing it with glutamine. CD36N8-10 (with sites N8, N9 and N10 intact), CD36N7,9,10 (with sites N7, N9 and N10 intact), CD36N7,8,10 (with sites N7, N8 and N10 intact) and CD36N7-9 (with sites N7, N8 and N9 intact) were expressed transiently and the cells surface expression was measured by mAb1258 binding and flow cytometry. CD36N8-10 reached the plasma membrane (Figure 5.9A), but the remaining mutants, CD36N7,9,10, CD36N7,8,10 and CD36N7-9, showed very little or no cell surface expression (Figure 5.9B), although they were expressed intracellularly (Figure 5.9D). To establish if all three sites within CD36N8-10 were needed to ensure trafficking of CD36, three more mutants were generated using pCD36N8-10 as a template and assessed for cell surface expression using the same system. CD36N8,9 (with sites N8 and N9 intact), CD36N8,10 (with sites N8 and N10 intact) and CD36N9,10 (with sites N9 and N10 intact) did not reach the plasma membrane (Figure 5.9C) but were expressed by the transfected HEK293T cells (Figure 5.9E). This suggests sites N8, N9 and N10 (all of which are shown by Q-ToF MS/MS to be modified by glycosylation; see section 4.4) were essential, within the carboxy-terminal glycosylation sites for expression, and that the three sites were the minimal number that would allow cell surface expression. Although the carboxy and amino-terminal glycosylation sets have been characterized, a combination of carboxy and amino-terminal glycosylation sites is likely to be needed for cell surface expression close to wild type levels. The influence of the carboxy-terminal sites on the trafficking of CD36N1-7 was therefore investigated by restoring each, individually to the CD36N1-7 template.
Figure 5.9 - Analysis of CD36 carboxy-terminal glycosylation sites.

Transfected cells were harvested and labelled with mAb1258 before being analysed by flow cytometry for cell surface expression. In (A), CD36N8-10 expressing cells are the green trace. In (B), CD36N7,9,10 expressing cells are the red trace, CD36N7,8,10 expressing cells are the yellow trace and CD36N7-9 expressing cells are the green trace. In (C), CD36N9,10 expressing cells are the red trace, CD36N8,9 expressing cells are the green trace and CD36N8,10 expressing cells are the light blue trace. Cells expressing CD36nong are shown by the blue trace and wild type CD36 expression cells are the black trace. (D) is an immunoblot of whole cell lysates expressing CD36N7,9,10 (lane 1), and CD36N7,8,10 (lane 2) and CD36N7-9 (lane 3). (E) is an immunoblot of whole cell lysates expressing CD36N9,10 (lane 1) and CD36N8,9 (lane 2) and CD36N8,10 (lane 3). Both are probed with mAb1955 and confirm successful transfection.
5.2.2.5. The effect of individual carboxy-terminal glycosylation sites on the trafficking of CD36N1-7

To assess the influence of the individual carboxy-terminal glycosylation sites, each was restored to the pCD36N1-7 template. CD36N1-7,8, CD36N1-7,9 and CD36N1-7,10 were all expressed transiently on the surface of HEK293T cells (Figure 5.10A). The expression of these mutants was 34%, 30% and 24% respectively, compared to wild type CD36. Each of the carboxy-terminal glycosylation sites was therefore able to significantly increase the expression of a protein when added to the amino-terminal N1-7 glycans (CD36N1-7 was otherwise expressed at only 2% of wild type CD36; Table 1). To establish whether these carboxy-terminal sites behave synergistically or additively N8, N9 and N10 were mutated individually to glutamine in an otherwise wild type CD36 background, to generate pCD36N1-7,9,10, pCD36N1-8,10 and pCD36N1-9. The proteins were expressed in HEK293T cells and reached the cell surface (Figure 5.10 B, C and D), on average at 63%, 78% and 75% of the level of wild type CD36 (Table 1), representing only a 1.3 to 1.6 fold difference in expression which confirms that neither N8, nor N9 nor N10 is individually essential for trafficking. In comparison the low level of CD36N1-7 expression, represents a fold reduction compared to wild type of 62.5. The synergy factor, which is calculated from the fold reduction of CD36N1-7 divided by the product of the fold reduction for CD36N1-7,9,10, CD36N1-8,10 and CD36N1-9, is 22.3, which suggests that N8, N9 and N10 behave synergistically (a synergy factor of 1, would indicate that N8, N9 and N10 behave additively).
Figure 5.10 - Flow cytometric analysis of HEK293T cells describing the influence of the carboxy-terminal sites N8, N9 and N10 on the trafficking of CD36N1-7.

Transfected cells were harvested and labelled with mAb1258 before being analysed by flow cytometry for cell surface expression. In (A), CD36N1-7-expressing cells are the red trace, CD36N1-8-expressing cells are the purple trace, CD36N1-7,10-expressing cells are the blue trace and CD36N1-7,9-expressing cells are the green trace. In (B), CD36N1-7,9,10-expressing cells are the red trace, wild type CD36-expressing cells are the black trace and CD36nong-expressing cells are the blue trace. In (C), CD36N1-8,10-expressing cells are the green trace, wild type CD36-expressing cells are the black trace and CD36nong-expressing cells are the blue trace. In (D), CD36N1-9-expressing cells are the purple trace. Wild type CD36-expressing cells are the black trace and CD36nong-expressing cells are the blue trace.
5.2.2.6. The effect of amino-terminal glycosylation sites on the trafficking of CD36N8-10

To further assess the contribution of amino-terminal glycosylation sites, N1 through to N6 were restored individually to pCD36N8-10 to generate the following mutants; CD36N1,8-10, CD36N2,8-10, CD36N3,8-10, CD36N4,8-10, CD36N5,8-10 and CD36N6,8-10 (CD36N7,8-10 had already been generated, see section 5.2.2.2) and the cell surface expression analysed by mAb1258 binding and flow cytometry. The addition of N7 (Figure 5.11A) and N4 (Figure 5.11B) and to CD36N8-10 made very little difference to the cell surface expression compared to CD36N8-10, consistent with the earlier observation that N7 has no effect on trafficking when part of the amino-terminal set, and that N4 is not utilized as a glycosylation site. Each of the other amino-terminal sites significantly improved cell surface expression of CD36N8-10, with N2 and N5 having the greatest effect. CD36N2,8-10 and CD36N5,8-10 achieve approximately 60% of cell surface expression level compared to wild type protein (Table 1), a five fold increase compared to CD36N8-10 (Figure 5.11C, as an illustration). CD36N1,8-10, CD36N3,8-10 and CD36N6,8-10 all achieve approximately 31% of cell surface expression compared to the wild type protein (Table 1), representing roughly a 2.5 fold increase compared to CD36N8-10 (Figure 5.11C, as an illustration). These data suggest that within the amino-terminal set, N2 and N5 are most important for interaction with the folding and trafficking pathway.
Figure 5.11 - Flow cytometric analysis of HEK293T cells expressing CD36min-g into which amino-terminal glycosylation sites have been reintroduced.

Transfected cells were harvested and labelled with mAb1258 before being analysed by flow cytometry for cell surface expression. In (A), CD36N7-10-expressing cells are shown in black. In (B), CD36N4,8-10-expressing cells are shown in blue. In (C), CD36N1,8-10-expressing cells are shown in blue, CD36N2,8-10-expressing cells are shown in red and wild type CD36-expressing cells are shown in black. For comparison, CD36N8-10-expressing cells are shown in green.
In an effort to establish occupancy of N4 in wild type CD36, the cell surface expression of CD36N1-3,5-10 (with N4 disrupted by site-mutagenesis) was compared to wild type CD36. As a control, CD36N1,3-10 (with N2 disrupted, which was previously identified as occupied) was also compared to wild type CD36 (Figure 5.12). The removal of N2 dramatically reduced cell surface expression, achieving 47% compared to wild type. The reduction is consistent with the other CD36 mutants where one occupied glycosylation site has been disrupted. CD36N1-7,9,10, CD36N1-8,10 and CD36N1-9 (with site N8, N9 or N10 mutated, respectively) achieved 63%, 78% and 75% cell surface expression on the plasma membrane when compared to wild type CD36. Conversely, CD36N1-3,5-10, with N4 disrupted showed very little change in cell surface expression, 92% compared to the wild type protein, which is consistent with non-occupancy of the glycosylation site.

Since the chosen method of analysis of the influence of glycosylation was to mutate the asparagine within the glycosylation consensus sequence, it was important to rule out the possibility that the restoration of the cell surface expression of CD36nong was due to the reintroduction asparagine rather than the subsequent post-translational modification. This possibility is addressed in the next section.
Figure 5.12 - Flow cytometric analysis of HEK293T cells expressing wild type CD36 and proteins with sites N2 and N4 removed.

Transfected cells were harvested and labelled with mAb1258 before being analysed by flow cytometry for cell surface expression. Wild type CD36 expressing cells are shown in black, CD36nong expressing cells are shown in blue, CD36N1-3,5-10 expressing cells are shown in red (with N4 disrupted) and CD36N1,3-10 expressing cells are shown in green (with N2 disrupted)

5.2.2.7. Glycosylation and not re-introduction of asparagines is the important determinant of trafficking efficiency

In the previous sections it has been shown that CD36nong fails to traffic to the plasma membrane, and that this phenotype can be rescued by restoration of the relevant asparagines. To assess whether glycosylation or the re-introduction of the asparagines allowed the restoration of cell surface expression, tunicamycin was used to inhibit in vivo glycosylation of CD36 mutants that traffic efficiently (see section 2.5.3). N-linked glycosylation is accomplished by the transfer of a lipid linked oligosaccharide-P-P-dolichol to the asparagine residue within the appropriate consensus sequence and this oligosaccharide is nearly always
comprised of two N-acetylglucosamines, nine mannoses and three glucoses. The first stage glycosylation is assembly of the oligosaccharide onto the dolichol-P-P, which involves the enzyme GlcNAc phosphotransferase GPT (Hubbard and Ivatt, 1981) which catalyses the following reaction:

\[
\text{GPT} \\
\text{UDP-GlcNAc + dolichol phosphate} \rightarrow \text{GlcNAc-P-P-dolichol + UMP}
\]

Tunicamycin is an antibiotic initially isolated from *Streptomyces lysosuperficus*, and inhibits the enzyme GPT (Keller et al., 1979), thereby preventing N-linked glycosylation. It is structurally similar to UDP-GlcNAc and competitively binds to GPT preventing this initial reaction from taking place and thus N-linked glycosylation (Keller et al., 1979). As tunicamycin prevents N-linked glycosylation globally it is cytotoxic, so it was important to use different concentrations to establish how much was needed to prevent glycosylation but still allow cell viability. Briefly, 5 hours post transfection, fresh medium containing 1ng/ml, 100ng/ml or 500ng/ml of tunicamycin was added to the HEK293T cells. The cells were harvested, as usual, 48 hours post-transfection. At all three concentrations tested, the toxicity of the antibiotic was tolerable, judged visually before harvesting the cells. At 500ng/ml the majority of CD36N8-10 did not display glycosylation, so this was the concentration used in future experiments.

It is likely that some of the chaperone proteins involved in trafficking and folding are glycoproteins themselves, therefore treatment with tunicamycin may affect the folding and transport of CD36 indirectly by preventing glycosylation and function of these chaperones. To ensure this possible indirect effect was controlled for, cells expressing the multi-drug resistance membrane P-glycoprotein (Pgp) were also treated with tunicamycin. Pgp is a member of the superfamily of ATP binding cassette (ABC) transporters and has ten putative N-linked glycosylation sites. Of these, only three are extracellular and occupied, N91, N94
and N99. Schinkel et al, removed all three glycosylation sites by replacing the asparagine with glutamine and transiently-expressing the mutants into the drug sensitive BRO human melanoma cell line. Vincristin, a cytotoxic drug and a Pgp substrate, was then used to select for resistant colonies. The resistance to vincristine is indicative of membrane bound and functional Pgp on BRO cells. The removal of the glycosylation sites did not appear to reduce the number of resistant colonies in this assay, indicating that Pgp does not require glycosylation to reach the plasma membrane or retain functionality (Schinkel et al., 1993). Consequently, Pgp is ideal for use as a control in this tunicamycin assay, since the phenotype due to lack of direct glycosylation can be distinguished from the phenotype due to a non-functional chaperone or other protein involved the folding or trafficking pathway.

HEK293T cells were transfected with CD36 containing either the amino-terminal or carboxy-terminal glycosylation sets (pCD36N8-10 and pCD36N1-7), or Pgp and grown in the presence of absence of 500ng/ml of tunicamycin. Following harvest, CD36 transfected HEK293T cells were labelled with mAb1258 as described previously, and Pgp was labelled with saturating concentration of the mouse mAb 4E3 (Zolnerciks et al., 2007), which recognizes an extracellular epitope of Pgp (Muller et al., 2000). Flow cytometry was used to assess cell surface expression and CD36nong used as a negative control. The expression of CD36N8-10 was almost abrogated following treatment with tunicamycin (Figure 5.13A), as was CD36N1-7 (Figure 5.13B), however there was only a small decrease in cell surface expression of Pgp compared to untreated control (Figure 5.13C). Immunoblots on all proteins treated with tunicamycin confirmed that the antibiotic prevented nearly all glycosylation (Figure 5.13D-F). This was especially important to show for Pgp (Figure 5.13F) because being a drug transporter there is the possibility that tunicamycin is a substrate for the protein. Since immunoblot indicated that Pgp was unglycosylated, this scenario is unlikely.

The simplest interpretation of these data is the oligosaccharides themselves allow cell surface expression of CD36 and not the reintroduced asparagines. However, western analysis has also
shown that even proteins that did not reach the cell surface such as CD36N1-5 and CD36N9,10, they do enter the trafficking pathway and become glycosylated. So, while some glycosylation sites seem to have more influence on trafficking efficiency (e.g. N2, N5, N8, N9 and N10) the density of the sites also appears to be important. Another factor that has yet to be considered is the nature of the added glycan i.e. is the glycosylation matured on route through the pathway and is this important for trafficking. This is considered in the following sections.
Figure 5.13 - Analysis of HEK293T cells expressing CD36 isoforms and Pgp in the presence and absence of tunicamycin treatment.

Transfected cells treated with tunicamycin, harvested, and labelled with mAb1258 (CD36) or 4E3 (Pgp) before being analysed by flow cytometry for cell surface expression. In (A), untreated CD36N8-10-expressing cells are shown in red and CD36N8-10-expressing cells treated with tunicamycin are shown in green. In (B), untreated CD36N1-7-expressing cells are shown in red and CD36N1-7-expressing cells treated with tunicamycin are shown in green. In (C), untreated Pgp-expressing cells are shown in red and Pgp-expressing cells treated with tunicamycin are shown in green. CD36nong-expressing cells are shown in blue. (D-F) are immunoblots of whole cell lysates. Untreated cells (lane 1), treated with tunicamycin (lane 2), treated with PNGaseF (lane 3) and treated with tunicamycin and PNGaseF (lane 4). D shows CD36N8-10 probed with mAb1955, (E) shows CD36N1-7 probed with mAb1955 and (F) shows Pgp probed with mAb C219.
5.2.2.8. Analysis of the nature of the glycosylation in transiently expressed CD36 isoforms

To differentiate between core and complex glycosylation, whole cell lysates prepared from HEK293T cells transiently transfected with the CD36 mutants, were denatured and treated with Endoglycosidase H (EndoH). EndoH is a glycosidase that was originally purified from *Streptomyces griseus* (Robbins *et al.*, 1984). Analysis of EndoH specificity by incubating the glycosidase with various glycopeptides and using paper electrophoresis to separate the products determined that the enzyme will only cleave within the chitobiose core of high mannose and some hybrid oligosaccharides from N-linked glycoproteins (Figure 5.14) (Arakawa and Muramatsu, 1974).

![Glycopeptide structures sensitive to EndoH](image)

**Figure 5.14 – Glycopeptide structures that are sensitive to EndoH.**

Cleavage occurs at the red dotted lines for high mannose or hybrid oligosaccharides. If the glycan structure is high mannose, n represents 2 to 150 mannoses and X represents 1 or 2 mannoses. If the structure is a hybrid that can still be cleaved by EndoH, n represents 2 mannoses and X and/or y represents the sugars AcNeu-Gal-GlcNAc.

High mannose oligosaccharides are generated in the endoplasmic reticulum, and these are sensitive to EndoH. Correctly folded and core glycosylated glycoproteins are generally transported to the Golgi where the glycans are trimmed further and elaborated, making the
oligosaccharides resistant to EndoH (see section 1.5 for further details). The protein is then trafficked to the plasma membrane.

As a control, the whole cell lysates prepared from CD36-expressing cells were also incubated separately with PNGaseF following denaturation. PNGaseF cleaves almost all oligosaccharides regardless of their composition (see section 4.2). This is used to confirm the size of the protein after complete deglycosylation for comparison with the EndoH sensitive fraction.

Whole cell lysates of HEK293T cells transiently transfected with pCD36N7,9,10, were treated with EndoH following denaturation (see section 2.7.11). Wild type CD36 which efficiently traffics to the cell surface and is likely to have mature glycosylation and CD36nong, which has no glycosylation sites, were treated similarly for comparison. Western analysis of the products showed that the mobility of CD36N7,9,10 is increased following treatment with EndoH and PNGaseF (Figure 5.15; lanes 5 and 6 compared to lane 4). Since the glycans are sensitive to EndoH, this CD36 isoform has immature high mannose oligosaccharides and so is likely to be stuck in the early trafficking pathway. A small fraction of wild type CD36 also appears to be sensitive to EndoH judged by the slight increase in mobility of the main protein band (lane 2 compared to lane 1), however, whether the high mannose glycans occupy the same glycosylation site(s) in each molecule or whether the slight EndoH sensitivity reflects incomplete maturation of glycans over the population as a whole, remains to be established.
Figure 5.15 - Electrophoretic mobility of wild type CD36 and CD36N7,9,10 before and after treatment with 1U EndoH or 1U PNGaseF.

Untreated wild type CD36 (lane 1), wild type CD36 treated with EndoH (lane 2), wild type CD36 treated with PNGaseF (lane 3), untreated CD36N7,9,10 (lane 4), CD36N7,9,10 treated with EndoH (lane 5), CD36N7,9,10 treated with PNGaseF (lane 6) and CD36nong (lane 7). Immunoblot was probed with mAb1955.

To test whether the EndoH sensitivity, and so the presence of immature glycans, was common to other CD36 isoforms that reached the cell surface, whole cell lysates expressing CD36N1-7 and CD36N8-10 were treated with EndoH and PNGaseF. Surprisingly, these protein isoforms with no sites in common, were highly sensitive to EndoH (illustrated in Figure 5.16A). However, the cell surface expression of these two mutants is quite low (2% and 12% of wild type levels respectively; Table 1), therefore, there is a possibility that the protein that is expressed at the plasma membrane does have complex glycosylation (and is resistant to EndoH), but the fraction is so small in comparison to a large intracellular pool it cannot be visualized easily by immunoblotting.

The CD36 mutants with a significantly higher cell surface expression than CD36N8-10 (Table 1) were then treated with EndoH and PNGaseF to assess the nature of their glycans. Most of these mutants with the exception of three, were largely sensitive to EndoH indicating the presence of high mannose glycosylation (CD36N1-7,9 is shown in Figure 5.16B and CD36N2,8-10 is shown in figure 5.16C to illustrate). However, three mutants, CD36N1-9,
CD36N1-8,10 and CD36N1-7,9,10 (the latter is shown in Figure 5.16D to illustrate), each with one glycosylation site disrupted were mainly resistant to EndoH indicating the presence of complex glycosylation.

CD36N2,8-10 and CD36N1-7,9,10 both had similar levels of cell surface expression (60% of wild type levels), however, CD36N1-7,9,10 is substantially more resistant to EndoH (Figure 5.15D) than CD36N2,8-10 (Figure 5.16C). The number of glycosylation sites present in these two isoforms is very different; nine in CD36N1-7,9,10 (eight of which are likely to be modified) and four in CD2,8-10 despite similarly high expression levels. These results strongly suggest that it is not necessary for CD36 to have complex glycosylation to reach the plasma membrane.
Figure 5.16 - Electrophoretic mobility of various cell surface expressing CD36 isoforms before and after treatment with 1U EndoH or 1U PNGaseF.

In A, untreated CD36N8-10-12His (lane 1), CD36N8-10-12His treated with EndoH (lane 2) and CD36N8-10-12His treated with PNGaseF (lane 3). In B, untreated CD36N1-7,9-12His (lane 1), CD36N1-7,9-12His treated with EndoH (lane 2) and CD36N1-7,9-12His treated with PNGaseF (lane 3). In C, untreated CD36N2,8-10-12His (lane 1), CD36N2,8-10-12His treated with EndoH (lane 2) and CD36N2,8-10-12His treated with PNGaseF (lane 3). In D, untreated CD36N1-7,9,10-12His (lane 1), CD36N1-7,9,10-12His treated with EndoH (lane 2) and CD36N1-7,9,10-12His treated with PNGaseF (lane 3). Immunoblots were probed with mAb1955.
5.2.2.9. **Affinity CD36 glycosylation mutants for acetylated low-density lipoprotein**

Despite the cell surface expression of many mutants generated, there is so far no evidence that the protein is folded correctly, which can be determined by a ligand binding assay. To establish whether the mutant CD36 proteins generated which reach the cell surface are folded correctly and whether glycosylation is required for the binding of ligand, a ligand binding assay was developed using BODIPY Ac-LDL to mimic the native ligand oxLDL. The assay is described in section 2.7.10 but briefly HEK293T cells transiently expressing CD36 were incubated in increasing concentrations of BODIPY Ac-LDL for several hours at 4°C. Following washing, the fluorescence associated with the cells in each well was measured, subtracting the non-specific binding to HEK293T cells expressing the intracellular CD36nong. A ligand binding curve was generated to determine the affinity of CD36 expressed on the surface of cells for BODIPY Ac-LDL. The data were analysed using GraphPad Prism software version 4.0 and the saturation binding curves were best fitted by Langmuir adsorption equation (Equation 3)

\[
B = \frac{B_{\text{max}} \cdot [L]}{K_d + [L]} 
\]

Where B is bound ligand (relative fluorescence units), [L] is the concentration of ligand (µg/ml) and Kd is the concentration of ligand giving half maximal binding and a measure of the affinity of ligand-receptor interaction.

The Kd of wild type CD36 expressed on the surface of HEK293T cells for Ac-LDL +/- S.E.M. was 8.3 +/- 1.4µg/ml (n=3)(Figure 5.17).
Figure 5.17 - Interaction of wild type CD36-12His expressed on the surface of HEK293T mammalian cells with Ac-LDL.

The affinity of Ac-LDL for CD36N8-10 expressed on the surface of HEK293T cells was measured as Kd +/- S.E.M 5.7 +/- 1.5µg/ml. (Figure 5.18A), which is not significantly different to wild type CD36. Unfortunately, specific binding of Ac-LDL to CD36N1-7 on the surface of mammalian HEK293T cells could not be detected, therefore, to establish whether N8, N9 and N10 was essential for proper folding and Ac-LDL binding, the affinity of CD36N1-7,9,10 (Figure 5.18B), CD36N1-8,10 (Figure 5.18C) and CD36N1-9 (Figure 5.18D) each missing one of the three carboxy-terminal sites was measured. The Kd +/- S.E.M. for each is 6.2 +/- 0.03µg/ml, 4.1 +/- 0.32µg/ml and 4.2 +/- 0.03µg/ml respectively, none of which are significantly different to wild type CD36-12His. This data suggests that no single glycosylation site is essential for the binding of Ac-LDL to CD36.
Figure 5.18 - Interaction of CD36 mutants expressed on the surface of HEK293T mammalian cells with Ac-LDL.

In A, CD36N8-10-12His; in B, CD36N1-7,9,10-12His; in C, CD36N1-8,10-12His and in D, CD36N1-9-12His.
5.3. Discussion

The extracellular domain of the class B scavenger receptor CD36 is heavily modified by glycosylation. Non-glycosylatable CD36, where either the asparagines within the ten N-linked glycosylation site consensus sequences had been mutated to glutamine, or HEK293T cells expressing wild type CD36 had been treated with tunicamycin, fails to reach the cell surface. This, together with the failure to express correctly folded CD36 in *E.coli* (which are unable to post-translationally glycosylate proteins) is consistent with the suggestion that glycosylation is required for correct folding and trafficking of CD36 to the plasma membrane. Despite CD36nong failing to reach the cell surface, it is expressed in HEK293T cells and this combined with the insolubility of CD36nong derived from insect cells in all by the strongest detergent (SDS), is consistent with the protein being shunted from the protein folding and trafficking and pathway into protein aggregates.

The need for glycosylation for proper folding is not unusual, it is also observed in other glycoproteins such as in the human organic anion transporter hOAT4. hOAT4 has four N-linked glycosylation sites and when glycosylation was prevented by mutating the asparagines in the glycosylation consensus sequences to glutamine or treating CHO cells expressing wild type hOAT4 with tunicamycin, the protein did not reach the cell membrane as determined by immunofluorescence (Zhou *et al.*, 2005). However, not all glycoproteins are equally dependent on their glycans for folding and secretion. The importance of glycosylation appears to be highly variable, the general rule being a protein with a large number of glycans are more dependent upon them during folding (Helenius and Aebi, 2004). P-glycoprotein (Pgp) has only three N-linked glycosylation sites and none are required for the protein to reach the cell membrane (Schinkel *et al.*, 1993 and data herein section 5.2.2.7).
The CD36 trafficking defect can be rescued by the reintroduction of carboxy-terminal glycosylation sites CD36N8-10 (approximately 200 fold increase in cell surface expression) and amino-terminal sites CD36N1-7 (approximately 20 fold increase in cell surface expression), however, neither protein is expressed to the same level as wild type (12% and 2% respectively). Since neither of these mutants have glycosylation sites in common, there is a level of redundancy in the glycan pattern for cell surface expression. This is similar to the observation in the human calcium receptor, hCaR. In hCaR, there are eleven putative N-linked glycosylation sites, eight of which are usually occupied. Mutagenesis of three or more occupied glycosylation sites from asparagine to glutamine caused progressive reduction in cell surface expression of the receptor, determined by biotin labelling and immunofluorescence. Like CD36 no individual glycosylation site was essential for folding and trafficking of the protein, however the degree of impairment differed according to the site of disruption. In hCaR, mutations of the amino-terminal sites show more impairment than mutation of carboxy-terminal sites (Ray et al., 1998), however this is not a general rule as the carboxy-terminal glycans appear to be more important in CD36.

In contrast to CD36, analysis of the related class B scavenger receptor, SR-BI from mouse demonstrates that two glycosylation sites (of the eleven in total) are required for cell surface expression (see section 4.7 for further details)(Vinals et al., 2003). The difference in the role of glycosylation between SR-BI and CD36 maybe due to the folding pathway each protein follows which is discussed below.

While glycosylation of CD36 appears to be important for trafficking to the plasma membrane, no individual site seems to be necessary for ligand binding. CD36N8-10 retained its protein fold being able to bind ligand with the same affinity as wild type CD36-12His. However, none of these carboxy-terminal glycans were necessary for ligand binding because mutant proteins each devoid of one site (CD36N1-7,9,10, CD36N1-8,10 and CD36N1-9) retain affinity for ligand. Unfortunately, demonstration of ligand binding to CD36N1-7 could
not be obtained. This is most likely due to the low level of expression (2% of wild type CD36), so it maybe impossible to distinguish specific binding of Ac-LDL to CD36N1-7 in HEK293T cells.

The expression and ligand binding data for CD36 mutants are consistent with a role for glycosylation in the folding and trafficking of CD36. Glycans are able to affect the folding of a protein in two ways, directly and/or indirectly. The direct effect is where the glycans promote folding of the protein biophysically; the addition of large polar carbohydrates affects folding of the polypeptide chain. Such local conformation changes often results in a β-turn (Bosques et al., 2004) and some glycans may stabilize local structures. It is hypothesised that the folding of Erythrina coralloidendron lectin (EcorL) is achieved in this manner. If glycosylated EcorL is denatured, approximately 80-90% of the protein can be refolded correctly in the absence of chaperones. This percentage is dramatically reduced to 10% to 20% if the protein is deglycosylated, indicating that the glycans have a role in folding of this protein (Herrmann et al., 1999). The X-ray structure of EcorL visualizes the glycan, a heptasaccharide on one, Asn17 (of the two) glycosylation sites (Shaanan et al., 1991).

Analysis of this glycan demonstrates that six of the seven sugar residues contact the protein directly or via water-mediated hydrogen bonds. It is hypothesised that this glycan is essential in proper folding as it may guide the folding pathway and stabilise the protein fold (Herrmann et al., 1999). Since CD36N1-7 and CD36N8-10 both reach the cell surface despite not sharing a glycosylation site in common, and none of the N8, N9 and N10 glycosylation sites are essential because single site mutants (CD36N1-7,9,10, CD36N1-8,10 and CD36N1-9) retained ligand binding affinity for Ac-LDL, it is unlikely that CD36 polypeptide folds around a specific glycan.

The requirement for glycosylation to fold and traffic CD36 but the redundancy observed with regard to number and position of these sites, is more consistent with the involvement of lectin-like molecular chaperones such as calnexin and calreticulin. In these folding pathways
glycosylation has an indirect effect on folding with the glycans serving as sorting signals (see section 1.5.1) (Helenius and Aebi, 2004). The association of the glycoproteins with these chaperones promote correct folding and assembly as well as supporting quality control (Rajagopalan and Brenner, 1994). Hebert et al., demonstrated by mutagenesis that the influenza virus glycoprotein hemagglutinin (HA) binds to calnexin and calreticulin as part of the folding process, and the number and location of the glycans was important. The study showed that although calnexin and calreticulin share the same carbohydrate specificity, they display distinct binding properties, and the different glycans on the target protein influence the rate of folding and the formation of disulphide bonds (Hebert et al., 1997). This may explain why the removal of glycosylation sites in CD36 reduces the cell surface expression compared to wild type CD36.

Although the minimal glycosylation sets of CD36 have low expression compared to wild type CD36, the addition of single amino-terminal glycosylation sites to the CD36N8-10 and the addition of single carboxy-terminal glycosylation sites to the CD36N1-7 dramatically increases the cell surface expression. In particular, the restoration of N2 or N5 to CD36N8-10 increased the expression of protein on the plasma membrane to 60% of wild type CD36. This is an important observation as either CD36N2,8-10 or CD36N5,8-10 maybe a suitable candidate for expression and purification from Sf21 insect cells. The five fold increase in cell surface expression compared to CD36N8-10 may enable a suitable yield of purified protein that, importantly may still be sensitive to PNGaseF or EndoH in its native form. If so, these mutants may be used in future structural studies.

One major surprise was the finding that glycans on the partially glycosylated CD36 proteins remained immature and sensitive to EndoH. This is true of proteins that share no sites in common (CD36N1-7 and CD36N8-10) which together include all nine utilized glycosylation sites in the protein, some or all of which must be mature-glycosylated in the wild type protein. It is also true for the proteins that are highly expressed (e.g. CD36N2,8-10). When
hOAT4 was expressed in CHO Lec1 cells, which have no detectable N-acetylglucosaminyl-transferase activity so all the proteins expressed in this cell type carry high-mannose oligosaccharides, hOAT4 reached the cell surface demonstrated by biotinylation and immunofluorescence analysis (Zhou et al., 2005). This suggests that CD36 and hOAT4 are dependent on the oligosaccharide modification in the ER but not dependent upon the processing of more complex oligosaccharides. This is in agreement with wild type CD36-12His expressed in Sf21 cells, which traffics to the plasma membrane and is folded correctly (see section 3.2.5) despite the inability of insect cells to elaborate on core glycosylation. However, it is not clear why the mutant CD36 proteins fail to acquire further oligosaccharides in the Golgi, although there does appear to be a correlation between site density on the protein and maturation of glycans. For example, the glycans in the single site mutant proteins; CD36N1-7,9,10-12His, CD36N1-8,10-12His and CD36N1-9-12His, each containing eight occupied glycosylation sites become mature, yet those on the three double mutants CD36N1-8-12His, CD36N1-7,9-12His and CD36N1-7,10-12His, each with seven occupied glycosylation sites, and CD36N2,8-10-12His (with four sites intact) remain immature.
6. General Discussion

The aim of this project was to investigate the role of N-linked glycosylation in the folding, trafficking and binding of ox-LDL to human CD36.

In order to assess the contribution each individual glycan makes to CD36, the occupancy of the ten putative N-linked glycosylation sites was first determined by comparing the electrophoretic mobility of CD36 glycosylation mutants in lysates prepared from transiently transfected HEK293T cells, and by mass spectrometry of fragments from wild type CD36 purified from Sf21 insect cells. To perform the mass spectrometry analysis, the protein first needed to be isolated from the membrane of the insect cells, which is described in chapter 3. Nine of the consensus sites were found to be modified post-translationally by glycosylation (chapter 4) and the impact of the sites on the trafficking and function of the protein was assessed and described in chapter 5.

In the following sections, the findings of this study are summarized and discussed in relation to the data in the literature, and further work to achieve a fuller understanding of the structure and function of human CD36 is considered.

6.1. Purification of wild type CD36 and mass spectrometry

Endogenous CD36 have previously been purified from human platelets (Tandon et al., 1989), bovine heart endothelial cells (Greenwalt et al., 1990) rat cardiomyocytes (Brinkmann et al., 2006) and rat adipocytes (Jochen and Hays, 1993) using the detergent TX114 to solubilize the protein from the membrane and a combination of different chromatographic techniques; size-exclusion, lectin-affinity and ion-exchange. However, in the current project, human
CD36 was purified for analysis by mass spectrometry, therefore it was important to maximise expression of the protein to obtain a suitable yield. This was achieved by expression in insect cells using a baculovirus expression vector. The protein was also engineered with a carboxy-terminal poly-histidine affinity tag, which allowed the detergent (OG) solubilized CD36 to be isolated using Ni-NTA chromatography. This approach was successful in isolating wild type CD36, however, it was not possible to purify large amounts of non-glycosylated CD36 because this form of the protein was largely insoluble unless treated with strong ionic detergent. This, in combination with the intracellular retention of non-glycosylated CD36 expressed in mammalian cells suggested that N-linked glycosylation is required for the protein to reach the cell surface.

As the yield of purified wild type CD36 was high, there is the possibility that this protein could be used for further structural studies, including crystallization trials. That the purified CD36 also appeared to be folded correctly, determined by the binding affinity of ac-LDL was also encouraging. Prior attempts (Martin and Linton, personal communication) at crystallography using rat Cd36 purified from insect cells, reproducibly generated small crystals but these diffracted only to around 20-25 Å. It is well known that membrane proteins are difficult to crystallize. This is likely due to several factors; the amphipathic nature of the protein, with the hydrophilic extracellular loop and hydrophobic transmembrane domains; and extensive, heterogenous and flexible N-linked. In an attempt to generate a protein that is more suitable for crystallization, PNGaseF was tested on native human CD36 to remove the N-linked glycans without denaturing the protein. Unfortunately, this resulted in incomplete deglycosylation even in the presence of a large excess of the enzyme, making the product unsuitable for crystallization.

Purified wild type CD36 was used to establish the occupancy of the ten putative N-linked glycosylation sites by mass spectrometry. The protein was denatured and deglycosylated using PNGaseF, and following digest of the protein with trypsin, Q-ToF MS identified that six
of the ten sites were occupied (N79 (N1), N205 (N5), N235 (N7), N247 (N8), N321 (N9) and N417 (N10)) The other four sites were in fragments that were not detected using this technique. FT-ICR identified the complete utilization of another site (N220 (N6)). However, N102 (N2), N134 (N3) and N163 (N4) were contained in a fragment that, again, was not detected. To identify whether the remaining three glycosylation sites were utilized, the electrophoretic mobility of different N-linked glycosylation CD36 mutants expressed in HEK293T cells was compared. This required the generation of a minimally-glycosylated mutant of CD36 that retained expression at the cell surface that could be used as a template for electrophoretic mobility studies.

6.2. Comparison of electrophoretic mobility of CD36 mutants

CD36N8-10 (see section 6.3 for a description of how the mutant was generated) can traffic to the cell surface despite having much less complex glycosylation than the wild type protein. It was considered a good template in which to test the utilization of N2, N3 and N4 by electrophoretic mobility. The addition of sites N2, N3 and N4 individually to CD36N8-10 identified that N2 and N3 were both occupied with glycans, but no evidence was found for glycosylation of N4. This, in addition to the mass spectrometry results, suggests that nine of the ten putative sites become glycosylated in CD36.

One limitation of the current study is that the absence of one glycosylation site may affect the occupancy of another. For example, rabies glycoprotein (RGP) has three N-linked glycosylation sites N37, N247 and N319. Generation of glycosylation mutants established that occupancy of N247 allowed glycan processing at N319 (Wojczyk et al., 2005). Thus, the removal of N1, 2, 3, 5, 6 and 7 in CD36N4,8-10 by mutagenesis could have prevented glycosylation of N4. When site N4 was removed from an otherwise wild type CD36 protein,
no change in electrophoretic mobility could be observed, but a similar result was obtained following the removal of site N8 that was shown to be occupied by mass spectrometry. This is probably due to the extensive N-linked glycosylation that occurs in the wild type protein and its inherent heterogeneity; thus if only one site is removed the resulting size difference is too insignificant to visualise using this method. Further evidence suggesting that N4 is not occupied in wild type CD36 comes from analysis of cell surface expression of the N4Q mutant. Data on other CD36 mutants (with sites N2, N8, N9 or N10 individually removed from an otherwise wild type protein) indicate that disruption of just one, normally occupied glycosylation site, can reduce the cell surface expression by up to 2.1 fold when compared to wild type CD36. Preliminary data on CD36 with N4 disrupted, shows very little decrease in cell surface expression, which suggests that the site is either not occupied, or is unimportant for folding and trafficking.

6.3. Establishing the minimal level of glycosylation to allow cell surface expression and function of CD36

This section of the project involved assessment of the role of each individual N-linked glycosylation site in addition to determining the minimum glycosylation required for a functional protein. This was achieved by site-directed mutagenesis and transient expression of the mutant protein in HEK293T cells. Expression was analysed by flow cytometry (to determine protein expressed on the cell surface) and immunoblotting (to primarily ensure that proteins not present on the cell surface were synthesised by the cells). A selection of mutant CD36 proteins that did reach the cell surface were then used in a ligand binding assay to assess the affinity for ac-LDL and to determine whether glycosylation was important in ligand recognition and to demonstrate that the proteins were folded correctly.
Initially, all of the N-linked glycosylation sites were mutated in CD36 and this established that N-linked glycans were required for the protein to be expressed at the plasma membrane. To determine how many and which glycans were important, the glycosylation sites were systematically reintroduced to the protein. Initially, this was most conveniently achieved by restoration of the carboxy-terminal sites (N7-N10) and the amino-terminal sites (N1-N7). Both these sets allowed trafficking to the plasma membrane at levels below wild type. The amino-terminal and carboxy-terminal sets were then characterised individually by systematically removing the glycans from the groups. From this analysis, it was possible to conclude that site number was important within the amino-terminal set, with five sites necessary for trafficking although none of these appear to be individually essential. In the carboxy-terminus the minimal number of glycosylation sites required was three, specifically N8, N9 and N10 (CD36N8-10).

Although both of these mutant CD36 proteins reached the cell surface at levels well above background, the expression compared to wild type was low, CD36N1-7 at 2% and CD36N8-10 at 12%. These CD36 isoforms have no single N-linked glycosylation site in common, so no individual glycan is essential for the folding and trafficking of CD36. This is suggestive of a role for glycosylation in the folding and trafficking of CD36 by mediating the interaction with lectin chaperones, calnexin and calreticulin. CD36N8-10 bound ac-LDL with the same affinity as wild type CD36 indicating that glycosylation sites N1 to N7 were not required for binding of this ligand. However, it was not possible to show binding of ac-LDL to CD36N1-7. This is likely to be due to the low cell surface expression of this mutant with ligand binding below the threshold of detection. It was then demonstrated that none of N8, N9 and N10 are necessary for ac-LDL binding, as single site glycosylation mutants each missing N8, N9 or N10 retain affinity for ac-LDL. The simplest interpretation from these data, taken together, is that no individual N-linked glycosylation site is required for the binding of ac-LDL.
Although, nine of the ten glycosylation sites are occupied in wild type CD36 none of them appear to be essential. To determine whether each site makes an equal contribution to cell surface expression of CD36 many different mutants were generated. Initially, an amino-terminal glycosylation mutant was used as a template (CD36N1-7), and the carboxy-terminal glycosylation sites were added individually to assess the contribution on cell surface expression. The addition of site N8 resulted in the highest increase in the cell surface expression, from 2% to 34%. This correlates well with the removal of N8 from wild type CD36 which resulted in a reduction in the cell surface expression compared to wild type CD36 (and more significant than the removal of either of the other two carboxy-terminal glycosylation sites N9 and N10). Interestingly, a genetic study in a french diabetic population in 1994 identified a naturally occurring single nucleotide polymorphism (SNP) that resulted in the amino acid change N247S (i.e. site N8). Although the protein, when expressed recombinantly, reached the cell surface and appeared to bind ac-LDL (Lepretre et al., 2004), neither the level of expression nor the affinity for ligand were measured. The binding of other ligands such as LCFA, which may be more relevant for type 2 diabetes was also not explored. Additionally, this study was carried out on a single family therefore the link between CD36 and type II diabetes is not conclusive. Identification of SNPs in a larger population with type 2 diabetes would be needed for a clearer understanding of the possible contribution CD36 polymorphisms in this disease.

In the same study another mutation was identified in CD36 associated with insulin resistance and familial type 2 diabetes in one family. This is a nonsense mutation in nucleotide T1079G resulting in a premature stop codon in exon 10 and a truncated protein missing 112 amino acids from the carboxy-terminal. The carboxy-terminal transmembrane domain and N10 glycosylation site are from this isoform, however all extracellular cysteines are still present. When this protein was expressed recombinantly, immunoblot showed that the protein was
synthesised and determined its size as approximately 55-60kDa consistent with significant truncation and loss of glycosylation. It could reach the cell surface of HEK293T cells although the expression was low, but it did not bind ac-LDL. In the current study, mutant CD36 with glycosylation site N10 disrupted was highly expressed at the cell surface, and bound ac-LDL with similar affinity to the wild type protein, so the phenotype of the T1079G mutant is due to the extensive truncation possibly resulting in the misfolding of the extracellular domain.

To assess the contribution of the amino-terminal glycosylation sites on the cell surface expression of CD36, each of N1-N7 were individually added to CD36N8-10. CD36N4,8-10 did not result in an increase in cell surface expression compared to CD36N8-10, which is unsurprising since no evidence was found that N4 was glycosylated. The addition of N7 CD36N8-10 also made very little difference to cell surface expression indicating that this glycosylation site, although occupied by a glycan is unlikely to contribute to the folding and trafficking of the protein. The addition of N1, N3 or N6 to CD36N8-10 induced a subtle upregulation in cell surface expression, causing a 2.6 fold increase compared to CD36N8-10 expression (to 31% of wild type CD36 levels). The two most promising mutants in terms of future studies were CD36N2,8-10 and CD36N5,8-10, of which both displayed a five fold increase in cell surface expression compared to CD36N8-10, achieving 60% of wild type protein levels at the plasma membrane.

6.4. **Purification of mutant CD36 isoforms**

Membrane proteins remain a challenging target for structural biologists. Crystallisation of a protein, in the presence or absence of ligands, allows the protein fold to be determined, can identify ligand binding sites and can provide evidence of conformational changes important
in the function of the protein. This structural information is essential for understanding the function of a protein and for identification of possible targets for intervention in disease. Difficulties in crystallising membrane proteins have been attributed to the partial hydrophobicity, the lack of stability, and the flexibility and heterogeneity of the N-linked glycosylation. Since native wild type CD36 cannot be deglycosylated using PNGaseF, it was not deemed suitable for further structural studies. CD36N8-10, with three occupied glycosylation sites is much less complex than wild type protein so attempts were made to purify and deglycosylate this CD36 isoform. This mutant when expressed in HEK293T trafficked to the cell surface, and bound ligand confirming that the protein was folded correctly. It was hoped expression of this protein heterologously in insect cells would allow purification of this form of the protein which may be more amenable to crystallisation. However, purification of the protein proved to be very challenging due to the relative insolubility of CD36N8-10 in all but the strongest detergent (SDS). SDS is not only likely to denature the protein, but also inhibits the binding to Ni-NTA so is far from ideal for this purpose. Low yield of protein recovered from the insect cells combined with the cell surface expression data from mammalian cells, suggests that there is not enough of CD36N8-10 expressed in the plasma membrane to enable purification for structural studies. However, the small amount of CD36N8-10 protein recovered and concentrated proved to be sensitive to PNGaseF without prior denaturation. In this regard, CD36N2, 8-10 and CD36N5,8-10 may be more suitable targets for purification and crystallization trials because the one additional glycosylation site correlates to a five fold increase in protein at the plasma membrane. With only one additional site these mutants may also be sensitive to PNGaseF in the native form. It is also possible that the combination of the two mutants to generate CD36N2,5,8-10 with five intact glycosylation sites may further increase protein expression at the plasma membrane.
A parallel project running in the lab explored the expression of the extracellular domain (ECD) of CD36. This would also be a good candidate for crystallisation because the protein would no longer contain the hydrophobic transmembrane domains or palmitoylation. This protein is expressed recombinantly in insect cells using a honeybee melittin signal sequence that is cleaved post-translocation into the ER to direct secretion of the protein into the medium. My colleague (Martin, C.A. personal communication) has developed a technique to purify this protein and has demonstrated that it retains affinity for ligand (ac-LDL) and so is likely to reflect the native fold. While the ECD is likely to require glycans for folding it should be possible from the data contained herein, to reduce the complexity of the glycans without significantly reducing the expression level of the ECD. Such a construct may provide the best opportunity for structural determination of the ligand binding domain of this scavenger receptor.

6.5. The nature of mutant CD36 glycans

Comparison of the sensitivity of various CD36 isoforms expressed in HEK293T cells to PNGaseF and EndoH allowed the nature of glycans to be determined. As expected, the majority of wild type CD36 is resistant to EndoH indicating that the glycosylation is complex, but mutant CD36 proteins with few glycosylation sites and which did not reach the cell surface were sensitive to EndoH. These data demonstrated that the mutants had entered the trafficking pathway but also suggested that complex glycosylation may be required for cell surface expression. However, when different partially glycosylated CD36 mutants that were surface expressed were analysed, it became clear that many of these proteins retained immature oligosaccharides.
CD36N8-10 and CD36N1-7 were both sensitive to EndoH indicating that despite their expression at the cell surface, complex glycosylation was absent. However, the proteins analysed were in whole cell lysates, which likely included forms that were retained intracellularly. Since the plasma membrane expression of these proteins was low (compared to wild type) there was a possibility that the correctly trafficked fraction may be outside the range of visualisation by immunoblotting, (i.e. the bulk of the protein observed were the ER retained species which will have high-mannose glycosylation). It was therefore necessary to analyse the nature of the glycans attached to higher expressing CD36 mutant proteins.

CD36N2,8-10 and CD36N1-7,9,10 have similar levels of cell surface expression. However, when incubated with EndoH it became clear that the nature of glycans on these proteins are different, with the glycosylation of CD36N1-7,9,10 mainly complex and CD36N2,8-10 mainly high mannose. As these two mutants are surface expressed to roughly the same level (60% of wild type) this strongly suggests that complex glycosylation is not absolutely necessary for CD36 to reach the plasma membrane of HEK293T cells.

Analysis of the carbohydrate composition of cell surface expressed bovine Cd36 offers a partial explanation for the results obtained herein (Berglund et al., 1996). The earlier study suggested that glycosylation sites N5, N8 and N10 (using human CD36 nomenclature) receive high mannose glycosylation in the wild type protein that is not further matured on route to the plasma membrane. If this is also true of human CD36 these glycans may still be sensitive to EndoH (explaining the slight sensitivity exhibited by wild type CD36, although this is also likely to be true of wild type protein in the ER at the time of harvest). In addition, half of the glycosylation sites in CD36N2,8-10 (three out of five in the case of CD36N5,8-10) would be of a high mannose nature, whereas only one quarter of the glycans would be high mannose in CD36N1-7,9,10. This may account for some of the observed differences in EndoH sensitivity, but in both CD36N2,8-10 and CD36N5,8-10 there is little evidence for any mature glycosylation. Further investigation would be needed to establish the exact
glycans occupying the mutant and wild type CD36 isoforms. This could be achieved by purifying the glyco-peptide fragments of CD36, ensuring that each fragment contained only one occupied glycosylation site. The glycans for each peptide can then be released using endoglycosidase and the resulting carbohydrate mixes can be analysed by mass spectrometry.

6.6. Summary and perspective for future research

This study identified that of nine out of ten N-linked glycosylation sites on the surface in human CD36 are modified by glycans. The glycosylation is necessary for folding and trafficking of the protein in HEK293T and Sf21 cells, but no individual site is essential. Glycosylation did not appear to be necessary for the binding of ac-LDL. This study did not explore the role of individual glycosylation sites in the binding of other ligands such as thrombospondin-1 and LCFA, which is an important consideration particularly if CD36 in different cell types exhibits different glycosylation. The effect of glycosylation on CD36 with respect to binding other proteins was also not explored. These could include binding proteins at the membrane but also the lectin chaperones and glycosyl-transferases involved in the trafficking and folding of CD36. Future analysis might identify these interacting proteins and the use of yeast split-ubiquitin two hybrid analysis (Molecular Biologische Technologie) may help in this regard. Briefly, ubiquitin can be separated into two moieties that do not then interact. If CD36 cDNA is fused to the sequence encoding the carboxy-terminal half of ubiquitin together with an artificial transcription factor, this can be used as bait to identify interacting proteins at the plasma membrane. A library of cDNAs fused with the sequence encoding the amino-terminal half of ubiquitin and if one of these fusion proteins interacts with the CD36 the bait ubiquitin is reconstituted. This attracts ubiquitination-specific proteases and results in proteolytic cleavage of the transcription factor attached to CD36 and activation
of selected reporter genes in yeast. Importantly this may identify partners involved in the receptor mediated internalization of ox-LDL and provide more targets for intervention. In addition, this project has generated CD36 isoforms that may be suitable for further structural studies. If CD36N2,8-10 (and CD36N5,8-10) can be purified from Sf21 cells and deglycosylated using PNGaseF in the absence of denaturation, full length and ECD proteins with this glycosylation pattern may be suitable for crystallisation trials. Crystallisation of CD36 would answer many outstanding questions about the nature of the ligand binding pocket(s) and how the protein evolved such a broad substrate specificity.
References


181


Appendix

DNA sequence pCD36-12His

tcaaatggccattagcataattattcatgtgtatatagcatatatcataaatatg
tactattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg
tacatattatatgcattgcataatagcagcagctatttgctgagctgtatgtaaatatg

196
Abbreviations

12His 12 Histidine
Å Angstrom
Ab Antibody
ABC ATP-binding cassette
ATP Adenosine 5’-triphosphate
Bodipy 4,4-difluoro-5-7-dimetyle-4-bora-3a, 4a-diaza-s-indacene-3-propionic acid
BSA Bovine serum albumin
BV Budded Virus
C12E8 Octaethylene glycol monododecy ether
CDG Congenital disorders of glycosylation
cDNA Complementary DNA
cmc Critical micelle concentration
CMV Cytomegalovirus
DDM n-Dodecyl-β-D-maltoside
dH2O Distilled water
DM N-decymaltoside
DMEM Dulbecco’s modified Eagle’s medium
DNA Deoxyribonucleic acid
DPP Dolicol-pyrophosphate
DTT Dithiothreitol
ECD Extracellular domain
ECL Enhanced chemiluminescence
EcoL Erythrina coralloidendron
EDTA Ethylenediamine tetraacetic acid
EndoH Endoglycosidase H
FABP Fatty-acid binding protein
FACS Fluorescence-activated cell sorting
FAF BSA Essentially Fatty acid-free BSA
FC9® Fos-choline 9
FCS Foetal calf serum
FT-ICR Fourier-transform ion cyclotron resonance mass spectrometry
GPT GlcNAc phosphotransferase
HA Influenza virus glycoprotein hemagglutinin
hCaR Human calcium receptor
HEK Human embryonic kidney cells
hOAT4 Human organic anion transporter 4
HRP Horseradish peroxidase
LC/MS/MS Liquid chromatography dual mass spectrometry
LCFA Long chain fatty acid
LMP-Ag Low melting point agarose
mAb Monoclonal antibody
MOI Multiplicity of infection
MOPS 3-(N-Morpholino)propanesulphonic acid
NTA Nickel-Nitrilotriacetic acid
OG n-Octyl-β-D-glucopyranoside
OST Oligosaccharyltransferase
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OV</td>
<td>Occluded virus</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PEI</td>
<td>Polyethyleneimine</td>
</tr>
<tr>
<td>Pgp</td>
<td>P-glycoprotein</td>
</tr>
<tr>
<td>PNGaseF</td>
<td>Peptide: N-glucosidase F</td>
</tr>
<tr>
<td>PVDF</td>
<td>Polyvinylidene fluoride</td>
</tr>
<tr>
<td>Q-ToF</td>
<td>Quadrupole time of flight</td>
</tr>
<tr>
<td>RPE</td>
<td>Recombinant phycoerythrin</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulphate</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard error of the means</td>
</tr>
<tr>
<td>SHR</td>
<td>Spontaneous hypertensive rat</td>
</tr>
<tr>
<td>TEMED</td>
<td>Tetramethylethylenediamine</td>
</tr>
</tbody>
</table>