

Dissecting the liver tumour micro-environment: Cell- and subcellular-specific characterisation of *N*-glycosylated basigin

Sayantani Chatterjee

A thesis submitted in partial fulfilment of the degree of Master of Research

Supervisor: Dr. Morten Thaysen Andersen



Department of Chemistry and Biomolecular Sciences

Faculty of Science

Macquarie University, Sydney, Australia

9th October 2017

Table of Contents

Statement of Originality	V
Acknowledgements.....	VI
Abbreviations and Symbol Nomenclatures	VIII
Abstract	IX
Chapter 1: Introduction and Aims.....	1
1.1. Liver Injury, Cirrhosis and Cancer	1
1.1.1. Cancer, a heterogeneous disease.....	1
1.1.2 Liver injury and cancer	1
1.2. The Immune System – Neutrophils	3
1.2.1. Origin and structure of neutrophils.....	3
1.2.2. Function of neutrophils	4
1.3. Basigin, a Key Glycoprotein in Cancer and Immunity.....	5
1.3.1. History of basigin.....	5
1.3.2. Basigin, a member of the immunoglobulin family	5
1.3.3. Structure of basigin	6
1.3.4. Tissue-specific expression and binding partners of basigin	8
1.3.5. Functions of basigin.....	9
1.3.6. Protein glycosylation, an important tumour-associated PTM	10
1.3.7. N-glycosylation of basigin.....	12
1.4. Aims of Research	14
Chapter 2: Materials and Methods	14
2.1. Materials.....	14
2.1.1. Chemicals and reagents.....	14
2.2. Methods	15
2.2.1. Isolation of neutrophils from human blood	15
2.2.1.1. Drawing of blood	15
2.2.1.2. Neutrophil isolation using Polymorphprep and Lymphoprep.....	16

2.2.1.3. Determination of cell counts, viability and purity	16
2.2.1.4. Neutrophil lysis and protein precipitation and quantification	17
2.2.2. Subcellular fractionation of blood-derived neutrophils	17
2.2.2.1. Disruption of neutrophils by nitrogen cavitation	17
2.2.2.2. Three-/four-layer Percoll density gradient separation of granules.....	18
2.2.2.3. Validation of plasma membrane isolation using latent alkaline phosphatase	18
2.2.3. Cell culture of liver cancer cells	19
2.2.3.1. Culturing, harvesting and monitoring HepG2 cells	19
2.2.3.2. Lysates and microsomes of HepG2 cells and protein extraction	19
2.2.4. Downstream handling of neutrophil and HepG2 protein mixtures	20
2.2.4.1. SDS-PAGE and staining of protein mixtures	20
2.2.4.2. In-gel trypsin digestion, and peptide extraction and clean-up	21
2.2.4.3. Protein detection using dot-blotting and Western-blotting	21
2.2.5. Immunoprecipitation.....	21
2.2.6. Cell surface protein capture and handling	22
2.2.7. N-glycomics	22
2.2.7.1. N-glycan release and clean-up	22
2.2.7.2. PGC LC-MS/MS of N-glycans.....	23
2.2.7.3. Data analysis of released N-glycans	23
2.2.8. Proteomics and N-glycoproteomics	24
2.2.8.1. ZIC-HILIC glycopeptide enrichment	24
2.2.8.2. LC-MS/MS of peptides and N-glycopeptides.....	24
2.2.8.3. Data analysis of N-glycopeptides	25

Chapter 3: Results 25

3.1. Reactivity of Anti-Basigin Monoclonal Antibodies (mAB) to Basigin.....	25
3.2. Isolation and Characterisation of Basigin from Neutrophils	26
3.2.1. Isolation, cell count, viability, purity and morphology of neutrophils	26
3.2.2. Basigin-centric IP of neutrophil lysates	27
3.2.3. Isolation of the basigin-rich plasma membrane fraction	28

3.2.4. Identification of plasma membrane-basigin using (glyco)-proteomics.....	29
3.3. Isolation and Characterisation of Basigin from HepG2 Cells.....	31
3.3.1. Morphology, viability and growth profile of HepG2	31
3.3.2. Investigation of cell surface-resident basigin on HepG2 cells.....	31
3.3.3. Basigin IP of HepG2 lysates	32
3.4. N-Glycome Profiling of Neutrophils and HepG2	35
Chapter 4: Discussion	37
4.1. Antibody-Based Enrichment and Detection of Basigin	38
4.2 Basigin-Focused Cellular and Subcellular Isolation Strategies	39
4.3. Molecular Characterisation of Neutrophil and Liver Cancer Basigin	41
4.4. N-Glycome Signatures of Key Cellular Components of Liver Tumours	46
Chapter 5: Conclusion and Future Directions	48
5.1. Conclusions.....	48
5.2. Future Directions	50
References	51
Appendices	60

Statement of Originality

This research has been conducted between January 2017 and September 2017 for the completion of Master of Research degree in Chemistry and Biomolecular Sciences (CBMS) in Faculty of Science and Engineering (FSE) at Macquarie University, New South Wales, Australia. This research thesis is certified to be an original work by the author, unless otherwise referenced in the literature and/or acknowledged of personal advice and suggestions.

Human Ethics (Reference #5201500409), Biohazard Risk Assessments (Reference #NIP041214BHA) and Biosafety approvals were duly obtained from Human Research Ethics Committee at Macquarie University to use human samples for research purposes only.

This thesis entitled “**Dissecting the liver tumour micro-environment: Cell and subcellular-specific characterisation of *N*-glycosylated basigin**” is formatted according to Master of Research Thesis guidelines prescribed by the Faculty of Science and Engineering and Department of Chemistry & Biomolecular Sciences and has not been submitted for qualification or assessment to any other institution.

Thank You,

Sayantani Chatterjee, 9th Oct 2017

Student ID: 43971245

Acknowledgements

The Masters of Research has been an incredible turning point in my academic life, and it wouldn't be so effortless and enjoyable without the persistent guidance and direction of my supervisor, **Dr. Morten Thaysen Andersen**. It has been an amazing learning experience for me and I would like to take this opportunity to thank you for your invaluable advice and expertise over these few months, for your constant time and patience towards me, for challenging my reasoning, believing in me and building the confidence that I am in the right place. I would also like to extend my thanks to **Dr. Nicolle H. Packer** without whom I wouldn't even be a part of the Glyco@MQ group. I am really grateful to you for placing your faith in my capabilities and giving me a chance to explore this world of research. Thank you for your constant encouragement and positive attitude during the unsettling moments.

The remarkable team of faculty members in the Master of Research committee including **Dr. Louise Brown, A/Prof Bridget Mabbutt, Prof Paul A. Haynes, A/Prof Mark Molloy, Prof Helena Nevalainen** and **Prof Ian Paulsen** is the reason that my mid-way transfer from coursework to research was so smooth and organised, and I thank each one of you for your decision in providing me with the MQ scholarship. I have gained a lot of knowledge from all your units and that understanding has really helped me to apply myself during the research activities.

During my research year, I have come across several people who definitely needs recognition for the successful completion of my thesis. Firstly, I would like to thank **Dr. Ian Loke** for his unbelievable contribution towards me throughout this year. You have been a strict teacher, yet a friendly senior, and imparted me with all the good qualities of a researcher. It is inspiring to brainstorm with you because of your immense knowledge and awareness. I would also like to thank **Dr. Jodie Abrahams** for leaving me with her valuable resources, her unwavering guidance towards my experimental abilities, for answering to all my queries and for the constant morale-boosting towards my thesis.

Thank you, **Christopher**, for constantly asking me "how are you today?" and "all good?" and solving each and every problem if things were not good. Most importantly, thank you for showing that research also has a relaxed and carefree side which works equally well. Thank you **Shazly** for the constant concern, for praising my work all the time and motivating me to do better each day. Thank you, **Edward**, for all the conversations about failed experiments and ideas of doing something new the next day. A heartfelt gratitude to **Sameera, Wei, Dr. Hannes Hinneburg** and **Dr. Rebeca K.**

Sakuma for their constant care and kindness towards me. Thanks a lot **Harry** for all the favours and support in the lab and for the candies during the writing period.

I would particularly also like to thank **Dr. Liisa Kautto**, **Dr. Robyn Peterson**, **Dr. Nima Sayyadi** and **Dr. Lindsay Parker** for their help and assistance whenever required. Thank you **Dr. Ling Lee** and **Dr. Zeynep S. Bayraktar** for your speedy replies in answering all my queries at every point. Thank you, **Dr. Benjamin Schultz**, for your valuable judgement during my poster activity. Thank you, **Dr. Benjamin Parker**, for your efficient help in fusion mass spectrometry at University of Sydney. Thank you **Dr. Regis Dieckmann** and **Professor Anna Karlsson** for your constructive insight into experimental hurdles and providing samples required for my project work.

To the other group members in the department who have also treated me with utmost care and helped and cooperated with me in times of need – **Hannah**, **Crystal**, **Christoph**, **Daniel**, **Albert**, **David**, **Mafruha** and **Wisam**; thank you very much to each one of you.

To the absolute friend-in-need in Sydney, **Deekshit da**, thank you for always being there with food and alcohol. Thanks to all my friends back in India for constantly being in touch and making life overseas much easier.

This journey would not have been possible without your unreserved support and companionship, **Russel**. Thank you so much for taking care of all my tantrums and mood swings throughout this period, holding on to my sanity and cooking delicious food for me during times of stress! I am glad I could complete this journey with you.

Lastly, and most importantly, none of this would be ever possible without my two pillars of strength, my parents, **Mummy** and **Daddy**. Saying a “thank you” can never be enough but this is a moment for me to be grateful to you for teaching me never to give up, for ensuring me in times of dismay that everything happens and falls into place at the right time and most importantly, to always do what makes you happy! Thank you for your unconditional love and endless support for me. Staying away from you hasn’t been easy, but I have learnt so much about myself from my experiences. I promise, I will make you really proud parents, one day.

I would also like to thank my extended family in India for their love and faith in me, especially, my **Dadu**, **Nani** and my uncle, **Ayan**, for whom the main concern has always been if I am getting enough food to eat and when will I be coming back to them! I miss you all.

“When there’s a will, there’s a way”

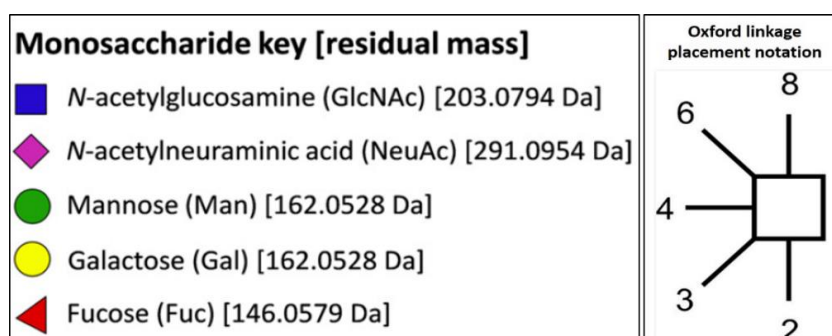
Abbreviations and Symbol Nomenclatures

ACD	Acetose Citrose Dextrose	ETHcD	Electron Transfer Higher-energy Collisional Dissociation
ACN	Acetonitrile		
Arg	Arginine	FBS	Fatal Bovine Serum
Asn	Asparagine	Glu	Glutamic Acid
Asp	Aspartic Acid	Gly	Glycine
BCA	Bicinchoninic Acid Assay	HCC	Hepatocellular Carcinoma
BPC	Base Peak Chromatogram	HCD	Higher-energy Collisional Dissociation
BSA	Bovine Serum Albumin	HEK	Human Embryonic Kidney
CD	Cluster of Differentiation	HLA	Human Leukocyte Antigen
cDNA	Complementary Deoxyribonuclease	HNE	Human Neutrophil Elastase
CID	Collisional Induced Dissociation	HPLC	High Performance Liquid Chromatography
Da	Dalton	HUGO	Human Genome Organisation
DMSO	Dimethyl Sulphoxide	IAA	Iodoacetic Acid
DTT	Dithiothreitol	ICC	Ion Charge Control
EC	Embryonal Carcinoma	IgSF	Immunoglobulin Super Family
EDTA	EthyleneDiamineTetraacetic Acid	IP	Immunoprecipitation
EGF	Epidermal Growth Factors	kDa	Kilo-Dalton
EGTA	Ethylene Glycol-bis- (β -aminoethyl ether) - N, N, N', N'-Tetraacetic Acid	LAP	Latent Alkaline Phosphatase
EIC	Extracted Ion Chromatogram	LC	Liquid Chromatography
ELISA	Enzyme-Linked Immunosorbent Assay	LTQ	Linear Ion Trap Quadrupole
EMMPRIN	Extracellular Matrix Metalloproteinase Inducer	Lys	Lysine
ER	Endoplasmic Reticulum	mAB	Monoclonal Antibody
ESI	Electrospray Ionisation	MCT	Monocarboxylate Transporter
FA	Formic Acid	MMP	Matrix Metalloproteinase
ETD	Electron Transfer Dissociation	MNC	Mononuclear Cell
		MOPS	3 - (N-Morpholino)-PropaneSulfonic acid
		MPO	Myeloperoxidase
		mRNA	Messenger Ribonucleic Acid
		MS	Mass Spectrometry

MS/MS	Tandem Mass Spectrometry	RdCVF	Rod-Derived Cone Viability Factor
NCG	Neutrophil Cathepsin-G		
NETs	Neutrophil Extracellular Traps	RIPA	Radioimmunoprecipitation Assay
OST	Oligosacaryltransferase	ROS	Reactive Oxygen Species
PAGE	Polyacrylamide Gel Electrophoresis	RP	Reversed Phase
PBS	Phosphate Buffered Saline	SDS	Sodium Dodecyl Sulfate
PGC	Porous Graphitised Carbon	SPE	Solid Phase Extraction
PMN	Polymorphonuclear Cell	Ser	Serine
PMSF	Phenyl Methyl Sulfonyl Fluoride	TBS	Tris Buffered Saline
		TBS-T	Tris Buffered Saline Tween
PNGase F	Peptide- <i>N</i> -glycosidase-F	Thr	Threonine
Pro	Proline	TIC	Total Ion Chromatogram
PTM	Post-Translational Modification	TFA	Trifluoroacetic Acid
		WB	Western Blotting
PVDF	Polyvinylidene fluoride	WBC	White Blood Cell
PVP	Polyvinylpyrrolidone	WHO	World Health Organisation
QE	Q-Exactive		
QTOF	Quadrupole Time of Flight	ZIC-HILIC	Zwitterionic Hydrophilic Interaction Liquid Chromatography
RBC	Red Blood Cell		

Glycan Analysis Symbol Nomenclature

The glycan structures presented in this thesis were created using GlycoWorkbench v2.1. Monosaccharide symbols and glycan representation followed the convention outlined by 2nd Essentials of Glycobiology (Varki et al., 2015). The respective glycosidic linkages of the glycan structures are represented according to the Oxford linkage placement notation as described in 2nd Essentials of Glycobiology (Varki et al., 2015).



Abstract

Basigin (CD147), an *N*-glycosylated transmembrane protein occurring in four potential isoforms, is expressed on surfaces of immune and cancer cells forming the heterogeneous tumour micro-environment. Basigin display pro- and anti-tumorigenic functions by interacting with multiple binding partners; interactions that are modulated by glycosylation. However, the molecular features of basigin and the *N*-glycosylation of the cells of the tumour micro-environment remain undescribed. This thesis aimed to characterise the low abundance basigin from primary neutrophils and HepG2 cancer cells as representative immune and cancer components of the liver tumour micro-environment. Multiple enrichment and isolation strategies including immunoprecipitation, cell surface protein capture and plasma membrane separation were performed in conjunction with western blotting and advanced tandem mass spectrometry. For the first time, human basigin isoform-2 was successfully characterised from both neutrophils and HepG2 by extensive mapping of the polypeptide chain. Cell- and subcellular-specific *N*-glycomics revealed paucimannose-rich *N*-glycosylation of neutrophil plasma membranes and dominant high mannose type *N*-glycans of HepG2 microsomes. In conclusion, glycobiology and powerful glyco-analytical technologies were brought together to reveal novel molecular features of a key cancer and immune glycoprotein. These findings contribute to our understanding of the structure and function of basigin and subcellular-specific *N*-glycosylation in the tumour micro-environment. (200 words)

Chapter 1: Introduction and Aims

The glycosylated basigin has roles in both immune and tumour processes in liver cancer [1, 10]. To promote understanding of the pleiotropic functions of basigin in the tumour micro-environment, human innate immune cells (i.e. neutrophils) and liver cancer cells (i.e. HepG2) were investigated in this study with the aims to isolate and characterise the structure of basigin and the cell-specific *N*-glycosylation features of these key cellular components of the liver tumour micro-environment.

1.1. Liver Injury, Cirrhosis and Cancer

1.1.1. Cancer, a heterogeneous disease

Cancer forms a very heterogeneous class of diseases involving uncontrolled growth of cells that can occur in any tissue. The body constantly makes new cells to promote growth, replace worn-out tissues and heal injuries. Normally, cells divide, multiply and die in an orderly and highly regulated way, but these complex processes may be altered upon changes occurring in one or more of the mechanisms controlling cell growth. This may cause abnormal blood or lymph fluid or form tumours in the affected tissue [10]. Thus, cancer a multi-factorial and multi-stage pathogenesis. Tumours can indeed be *benign* in nature, which is a pathophysiological condition that unlike cancer is not considered a life-threatening disease, but is a relatively inert aggregation of cells. In contrast, *malignant* tumours are fast growing cancerous cells that can spread from their primary site of malignant growth to other locations where the cancerous cells can form secondary tumours arising from pre-cancerous conditions e.g. lesions, infections and chronic inflammation [2]. The spreading of cells is referred to as metastasis, a characteristic specific to malignant cells, **Figure 1** [2].

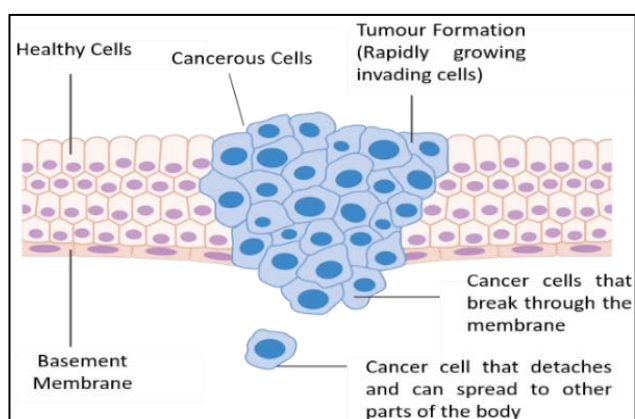


Figure 1: The mechanism of metastasis involving the dissociation of malignant cancer cells from the original tumour location. These cancerous cells are able to spread to another part of the body through the blood stream by adhering and proliferating in a distant tissue [1, 2].

1.1.2 Liver injury and cancer

The liver is the largest internal organ in the body and part of the digestive system. The liver has several important functions such as storing nutrients, metabolising proteins, fats and carbohydrates

to provide energy and building blocks, and removing waste products - functions which are essential to maintain health. The liver is one of the main cancerous sites in the human organism; liver cancer accounted for 788,000 deaths in 2016 according to the World Health Organization [11]. One of the primary and most prevalent types of liver cancer is hepatocellular carcinoma (HCC), which involves uncontrolled growth of hepatocytes, the main cell type of the liver [12, 13].

Upon sustained liver damage, healthy functioning liver tissue gets replaced with scarred tissue that less efficiently performs the normal functions of the liver. The entire organ can transform into a malfunctioning so-called cirrhotic liver. The scar tissue in a cirrhotic liver blocks the flow of blood through the hepatic artery and portal veins and slows the processing of nutrients, drugs and toxins leading to necrosis of liver cells. Cirrhosis can arise from excessive alcohol intake, unhealthy food choices and liver inflammation from other types of injury or due to pathogen infection of the liver [1, 14]. The damage of the liver develops through stages of mildly over moderately to severely injury. Over a prolonged period often spanning years and even decades, the scarred cirrhotic tissues, which are considered pre-cancerous in nature, often develop into liver cancer [12, 13].

HCC is a primary malignancy of the liver that occurs due to the prolonged state of liver injury, chronic liver disease and/or cirrhosis. The tumour spreads with local expansion, intrahepatic spread and distant metastases. Both external and environmental factors may lead to excessive cell proliferation [1]. In addition to the cancerous hepatocytes, the liver tumour micro-environment consists of different cell types i.e. adipocytes, macrophages, fibroblasts, cancer stem cells, non-stem like carcinoma cells, granular leukocytes most significantly neutrophils and B- and T-cells, **Figure 2** [1]. This thesis will focus on studying liver carcinoma cells and neutrophils, two significant cell populations in the HCC tumour micro-environment.

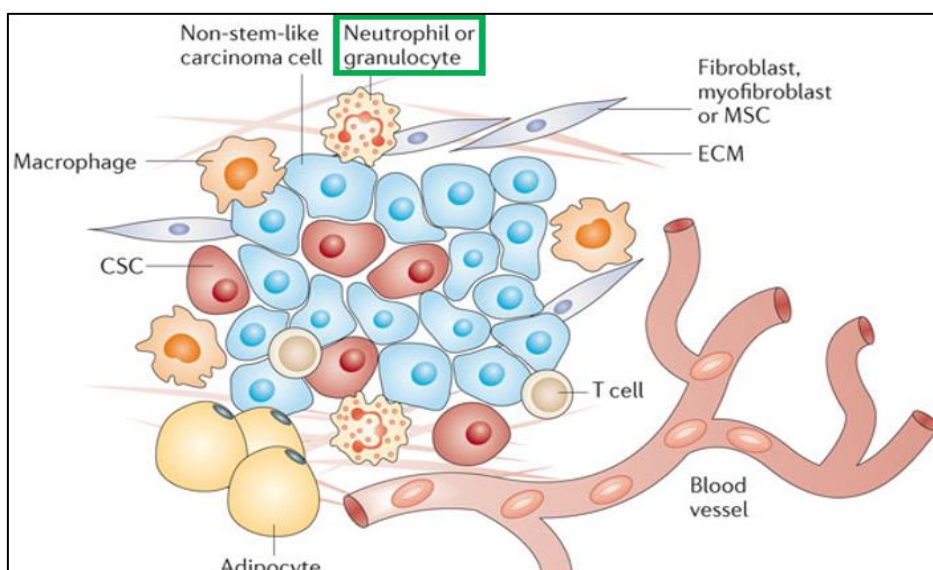


Figure 2: Simplistic representation of the different cell types found in a malignant tumour micro-environment including white blood cells e.g. neutrophils [2].

1.2. The Immune System – Neutrophils

1.2.1. Origin and structure of neutrophils

In the 1980s, immunologists investigated the different cellular components in blood and structurally and functionally classified the blood cells. Crude distinctions were made between the erythrocytes (also known as red blood cells, RBCs) and the leukocytes (known as white blood cells, WBCs) [3].

The bone marrow is the tissue origin of many blood and immune cell types that eventually end up in the peripheral blood or the distal organs and tissues. This plethora of cell types arises from common bone marrow stem cells. Stem cells are defined as precursor cells with an ability to differentiate into various mature cells – The marrow consists of two types of stem cells; hematopoietic stem cells, which produce blood cells, and the stromal stem cells, which produce bones and cartilage. Hematopoietic stem cells can differentiate into both myeloid and lymphoid committed precursor cells, **Figure 3**. The so-called myeloid lineage gives rise to red blood cells, platelets and granular white blood cells of which neutrophils, of focus here, are the most abundant cell type. The lymphoid lineage produces B- and T-cells and natural killer cells. Similar to the myeloid cells, the lymphoid cells also belong to the WBC family that is essential for fighting abnormal growth and infection [7].

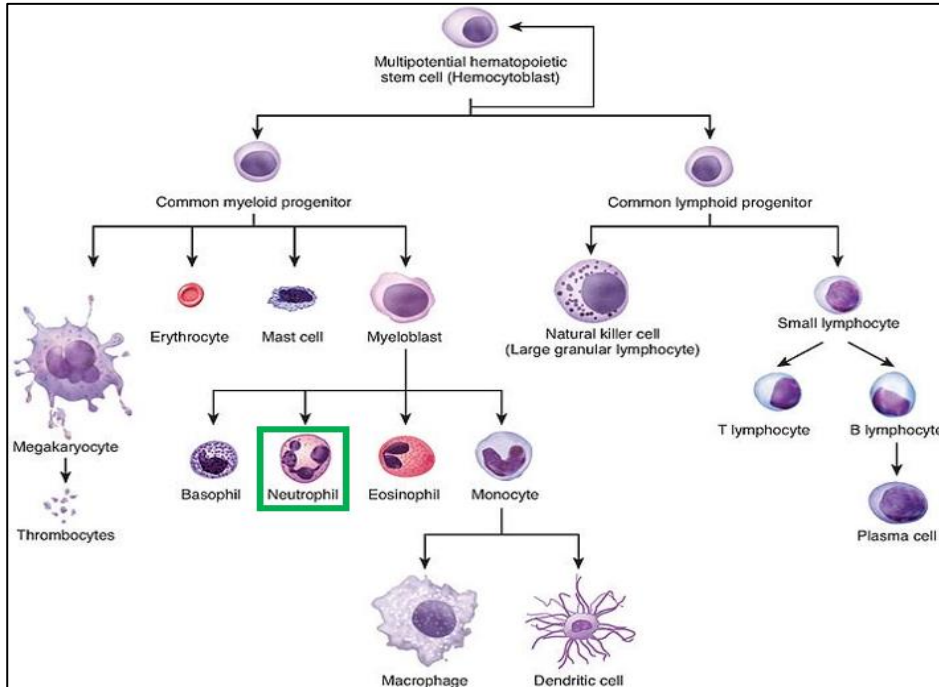


Figure 3: The map of human haematopoiesis. Hematopoietic stem cell gives rise to myeloid and lymphoid progenitors and a range of specialised effector cell types. Neutrophils investigated in this project are the most abundant form of the WBCs [7].

Neutrophils, also referred to as polymorphonuclear (PMN) leukocytes, are the most abundant WBCs found in humans. After maturation in the bone marrow, the mature “resting” neutrophils typically have a half-life of 6-8 hours in the blood circulatory system, after which they may exit circulation through the blood vasculature and migrate to the inflamed tissues where they can stay for 2-3 days

if a pathogenic infection, injury or threat is imminent [3]. Neutrophils have multi-lobed nuclei, and other more common organelles e.g. mitochondria, endoplasmic reticulum, phagosomes. Importantly, neutrophils carry three main types of granules – the *primary* or azurophilic, *secondary* or specific and *tertiary* or gelatinase granules, **Figure 4** [3, 15]. These granules comprise distinct proteomes with glycosylation patterns differ between granules and the individual proteins and their sites [16]; for example, the primary granules comprise important glycoproteins including myeloperoxidase (MPO) [16], neutrophil elastase (HNE) [17] and neutrophil cathepsin G (NCG) that show unusual and different glycosylation patterns [18]. Protein glycosylation is described below.

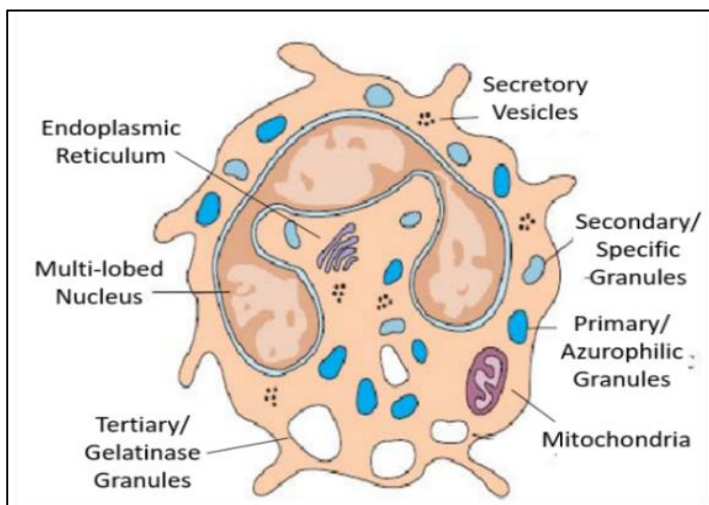


Figure 4: Morphology of a resting human neutrophil displaying the characteristic multi-lobed nucleus and multiple types of granules [3].

1.2.2. Function of neutrophils

Neutrophils primarily act as the first mediators of the rapid innate immune response of the host against bacterial and fungal pathogens before the acquired immune system is initiated to promote a more specific response to the threat [3, 7, 19, 20]. To carry out the protective functions against pathogens, neutrophils are capable of performing many fascinating cellular processes including phagocytosis, adhesion, degranulation, chemotaxis, apoptosis, extracellular trap formations and biochemical processes e.g. ion channel regulation and alterations in phospholipid metabolism [3, 21]. When inflammatory signals arise from injured tissues, circulating neutrophils react to such signals and have specialised mechanisms to exit the blood vasculature and target the site of inflammation (or cancerous tumours). The neutrophils rapidly migrate to the inflamed sites via amoeboid movement where they extend long projections called pseudopodium through which the granules flow; this action is followed by contraction of filaments in the cytoplasm that draws the nucleus and rear of the cell forward. Neutrophils are able to engulf the foreign bodies via active phagocytosis. The engulfment results in the formation of phagosomes containing microbicidal superoxides or reactive oxygen species (ROS) that serve to effectively kill the foreign substances

[18]. Neutrophils are also known to combat infection via a process called degranulation. Upon degranulation the neutrophil granules that harbour compounds with antimicrobial properties are released at the sites of infection leading to death of foreign invaders [22]. A third killing mechanism of activated neutrophils is the formation of a web of chromatin fibres and serine proteases referred to as neutrophil extracellular traps (NETs) that serve to trap and kill the pathogens and act as a physical barrier that prevents the further spreading of the pathogens in the body [21].

1.3. Basigin, a Key Glycoprotein in Cancer and Immunity

1.3.1. History of basigin

Basigin is a single-pass type-I transmembrane glycoprotein. Similar to two other proteins, embigin and neuropilin, basigin belongs to the immunoglobulin superfamily (IgSF). Basigin was first isolated from the plasma membrane of the human LX-1 lung carcinoma cell line in 1982 where its interaction with hyaluronate was studied [23]. Several types of tumours, but not isolated tumour cells in culture, are known to contain high levels of hyaluronate, a cell-surface located anionic non-sulphated glycosaminoglycan. Hyaluronate is also widely distributed throughout non-cancerous epithelial and connective tissues and is associated with cell proliferation and metastases of malignant cells. This landmark study showed the significance of hyaluronate production and the involvement of basigin in these induction processes [23].

Over the years, basigin (UniProtKB: P35613) has been studied from various species by various research groups. Basigin has thereby obtained different designations in the literature. For example, the human protein is known by multiple names including cluster of differentiation-147 (CD-147), extracellular matrix metalloproteinase inducer (EMMPRIN) [24], hBasigin [25], leukocyte activation antigen M6 [26] and HAb18g [19]. Names in other species include OX-47 [27] and CE9 [28] in rats, GP42 [29] and basigin-1 [30] in mice, and HT7 [31], neurothelin [32] and 5A11 [31] in chicken. Basigin will consistently be used in this thesis regardless of the species origin to ease the reading. As per the Human Genome Organisation (HUGO), the gene encoding for human basigin is *BSG*.

1.3.2. Basigin, a member of the immunoglobulin family

Structural studies have found that many cell-surface proteins belong to distinguishable super-families. The IgSF comprises a broad range of molecules such as the founder embigin and the T-cell receptors, major histocompatibility complex (MHC) class I-II, neural cell adhesion molecule (N-CAM), CD4 and CD8 [33]. The Ig domains are generally classified into V, C1 and C2 and I sets; the latter is an intermediate domain of the V and C sets.

Genetic approaches using cDNA expression libraries revealed the presence and important structural characteristics of basigin in mouse embryonal carcinoma (EC) cells [34]. Based on these studies, basigin is known to be homologous to both the Ig V domain as well as the MHC class II antigen β -chain [33, 35]. Southern and Northern blotting revealed the open reading frame (ORF) of *BSG* and indicated that basigin appears as a single or few mRNA copies per mouse genome. Homologous genes were detected in hamsters and humans [33, 35]. The cDNA analyses of EC cells were also used to estimate that the molecular mass of human basigin with the putative signal sequence was approximately 32 kDa. However, basigin from human EC cells migrated as a broad sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) band of 42-60 kDa. The broad molecular mass range was attributed to heavy glycosylation of the polypeptide chain. Upon treatment with peptide:N-glycosidase F (hereafter PNGase F) that removes asparagine (N)-linked sugar moieties from proteins, basigin was reduced to 32 kDa indicating significant N-glycosylation [34].

1.3.3. Structure of basigin

In the 1990s, the structure of basigin was studied extensively to identify its key features and advance the functional understanding of basigin in biological systems. *BSG* is located on chromosome 19 at p13.3 consisting of ten exons that span about 12 kb [36]. The gene has been detected in almost all vertebrates, and in *Drosophila melanogaster* and *Schistosoma*. The N-terminal or V-set domain of basigin is responsible for counter-receptor activity and protein oligomerisation. The C-terminal or C1-set domain (see topology below) is involved in forming associations with different binding partners e.g. caveolin-1, integrins ($\alpha 3\beta 1$ and $\alpha 6\beta 1$), annexin II amongst other proteins [37, 38].

The polypeptide chain of human basigin have four isoforms arising from alternative splicing and differences in transcription initiation sites, **Figure 5** [6, 39]. The common variant of basigin is isoform 1, basigin-2 or simply *BSG*; this common isoform has a predicted protein mass of approximately 42 kDa, spans 385 amino acid residues, and comprises two Ig domains. Isoform 2 or basigin-1 is a less known isoform. Two other isoforms of human basigin known as basigin-3 (isoform 3) and basigin-4 (isoform 4) are weakly expressed in normal and tumour tissues. This work will consistently refer to the four basigin isoforms as isoform 1-4.

Isoform 1 (basigin-2) contains a 21-residue signal peptide sequence, an extracellular domain (ECD) of 185 amino acid residues, a glutamic acid-rich transmembrane helical domain of 20 amino acid residues and a 40-amino acid cytoplasmic domain, **Figure 6** [37-39]. The glutamic acid residues in the transmembrane domain are conserved across species implying that they serve important functions e.g. in the intramembrane association between basigin and interaction partners.

A) Basigin-2 (canonical) - P35613|BASI_HUMAN Isoform-1 - *Homo sapiens*
¹MAAALFVLLGFGALLGTHGASGAAGFVQAPLSQQRWVGGSVELHCEAVGSPVPEIQWVWFEQ
 QGPN²DTCSQLWDGARLDRVHIHATYHQHAASTISIDTLVEEDTGTIECRASNDPDRNHLTRAP
 RVKWWRAQAVVVLVLEPGTVFTTVEDLGSKILLTCSL³NDSATEVTGHRWLKGGVVLKEODALPG
 QKTEFKVDSDDQWGEYSCVFLPEPMGTANIQLHGPPRVKAVKSSEHINEGETAMLVCKSESVP
 PVTDWAWYKITDSEDKALM⁴NGSESRRFFVSSSQGRSELHIENLNMEADPGQYRC⁵NGTSSKGSQ
 QAIITLRVRSHLAALWPFLGIVAEVLVLVTIIFIYKRRKPEDVLDDDDAGSAPLKSSGQHQNDKG
 KNVRQRNSS³⁸⁵
(385 AA, predicted polypeptide mass: 42.2 kDa)

B) Basigin-1 - P35613-2|BASI_HUMAN Isoform 2 - *Homo sapiens*
¹MAAALFVLLGFGALLGTHGASGAAGTVFTTVEDLGSKILLTCSL²NDSATEVTGHRWLKGGVVLK
 EDALPGQKTEFKVDSDDQWGEYSCVFLPEPMGTANIQLHGPPRVKAVKSSEHINEGETAMLVCK
 KSESVPVPTDWAWYKITDSEDKALM³NGSESRRFFVSSSQGRSELHIENLNMEADPGQYRC⁴NGTS
 SKGSQAIITLRVRSHLAALWPFLGIVAEVLVLVTIIFIYKRRKPEDVLDDDDAGSAPLKSSGQHQ
 NDKGKNVRQRNSS²⁶⁹
(269 AA, predicted polypeptide mass: 29.2 kDa)

C) Basigin-3 - P35613-3|BASI_HUMAN Isoform 3 - *Homo sapiens*
¹MGTANIQLHGPPRVKAVKSSEHINEGETAMLVCKSESVPVPTDWAWYKITDSEDKALM²NGSE
 SRFFVSSSQGRSELHIENLNMEADPGQYRC³NGTSSKGSQAIITLRVRSHLAALWPFLGIVAEVL
 VLVTIIFIYKRRKPEDVLDDDDAGSAPLKSSGQHQNDKGKNVRQRNSS¹⁷⁶
(176 AA, predicted polypeptide mass: 19.4 kDa)

D) Basigin-4 - P35613-4|BASI_HUMAN Isoform 4 - *Homo sapiens*
¹MKQSDASPQERVDSDQWGEYSCVFLPEPMGTANIQLHGPPRVKAVKSSEHINEGETAMLV
 CKSESVPVPTDWAWYKITDSEDKALM²NGSESRRFFVSSSQGRSELHIENLNMEADPGQYRC³NGT
 SSKGSQAIITLRVRSHLAALWPFLGIVAEVLVLVTIIFIYKRRKPEDVLDDDDAGSAPLKSSGQHQ
 QNDKGKNVRQRNSS²⁰⁵
(205 AA, predicted polypeptide mass: 22.8 kDa)

Figure 5: Polypeptide sequence and predicted protein mass of the four known isoforms of human basigin i.e. (A) isoform 1 (basigin-2, canonical form), (B) isoform 2 (basigin-1) (24-139 AA missing from isoform-1), (C) isoform 3 (1-209 AA missing from isoform-1) and (D) isoform 4 (12-191 AA missing from isoform-1) [6]. Signal peptides (orange) and potential *N*-glycosylation sites (blue) are marked. Information obtained from UniProtKB.

The transmembrane domain also contains a typical leucine zipper motif comprising three leucine residues and a phenylalanine that appears at every seventh residue facilitating membrane-protein associations. The two Ig domains of basigin were assigned slightly different than the usual classification i.e. D0, I-set, D1, C2-set, and D2 set, I-set [37-39]. Position 138-219 of isoform 1 is similar to Ig-like C2 type whereas the position 221-315 is similar to Ig-like V type. Isoform 2 (basigin-1) is a retina-specific isoform containing a total of three Ig domains. The additional isoform 2 D0 domain is lacking glycosylation of the additional D0-located potential glycosylation site [30, 40].

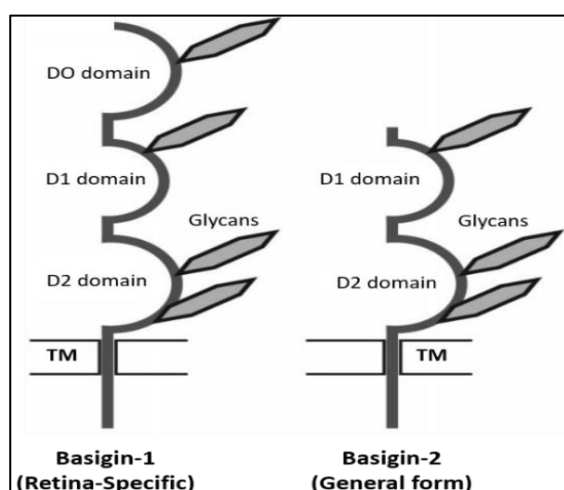


Figure 6: Schematic representation of the topology and domain structure of the two most abundant isoforms of human basigin i.e. isoform 2 (left) and the canonical isoform 1 (right). TM: transmembrane domain. The grey-shaded hexagons indicate potential *N*-linked glycosylation sites i.e. sequon-located Asn residues [5].

Basigin is a highly glycosylated transmembrane protein that is capable of recognising molecules on the surface of the same cells (*cis* recognition) and molecules on other cells or in the extracellular matrix (*trans* recognition). The *cis*- and *trans*- interactions are better understood of isoform 2

compared to isoform 1, **Figure 7** [37, 38]. Isoform 2 is known to form homo-oligomers in a *cis*-dependent manner on the plasma membrane [41, 42]. Isoform 3 reportedly serves as an endogenous inhibitor of isoform 1 caused via hetero-oligomerisation with isoform 2 [37-39]. Isoform 3 and isoform 4 are known to contain a single extracellular Ig domain and consist of so-called “highly” and “lowly” glycosylated forms (described below). However, the exact glycosylation patterns of any of the human basigin isoforms remain unknown.

1.3.4. Tissue-specific expression and binding partners of basigin

Basigin carries out numerous physiological and pathological functions, which reflects its diverse cell and tissue expression. For example, basigin was reported to be broadly expressed on hematopoietic and non-hematopoietic cells e.g. monocytes and granulocytes including neutrophils and epithelial and endothelial cells [43]. As described above, basigin is also highly expressed on different cancerous cell types e.g. liver (HepG2) and kidney (HEK293) cancer cells [44-46]. Basigin is known to serve multiple and diverse cell type-specific functions in the tumour micro-environment. It is likely that the cell-specific functions of basigin are defined or fine-tuned by cell-specific *N*-glycosylation in the tumour micro-environment. Additionally, basigin is a cancer biomarker candidate for the diagnosis and prognosis of a wide range of cancers [44].

In addition to having the propensity of forming homo-dimers, basigin has multiple other endogenous binding partners including rod-derived cone viability factor (RdCVF), GLUT1, CD44, integrins, cyclophilins amongst others, **Figure 7**. RdCVF is a truncated thio-redoxin-like protein that non-enzymatically enhances the survival of cone photo-receptors. The retina-specific isoform 2 may be a receptor for RdCVF [47]. It also binds to GLUT1, a glucose transporter with 12 membrane-spanning regions. Interestingly, knockdown of isoform 2 suppresses the functions associated with RdCVF and GLUT1 thereby supporting the biological importance of interaction to basigin [48-50].

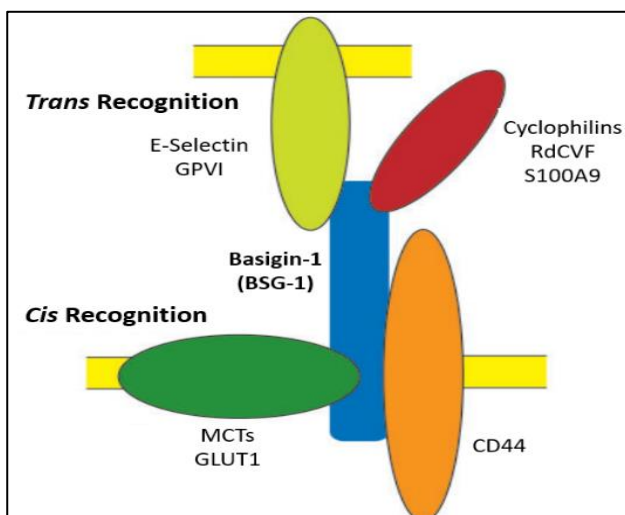


Figure 7: Isoform 2 (basigin-1) (in blue, here denoted as “BSG1”) interacts with various proteins via *cis*- and *trans*- interactions. In the *cis* form, isoform 2 binds to proteins residing in the same cell e.g. MCT/GLUT1 (green) or CD44 receptors (orange). In the *trans* form, isoform 2 binds to soluble proteins like cyclophilins or S100A9 (red) or proteins on adjacent cells e.g. GPVI or E-selectin (pale green) [5].

Integrins are transmembrane receptors that function to facilitate ECM adhesion to cells. Upon binding with ligands in the ECM, integrins can activate signal transduction pathways. Different integrins e.g. CD98, CD43, MCT4 and galectin-3 are known to mediate interactions with basigin. Integrin $\alpha 3\beta 1$ and $\alpha 6\beta 1$ were co-immunoprecipitated with basigin from cell lysates. Their binding was further supported by co-localisation of basigin and integrin $\alpha 3\beta 1$ and $\alpha 6\beta 1$ at the cell surface [47, 51]. The exact molecular interactions between basigin and the integrins still remain unclear, however, the elimination of *BSG* expression did affect the phenotype of the cell thereby adding valuable support for their proposed interactions and functional consequences of their binding.

Cyclophilins form a family of proteins that are known to bind to cyclosporin A. Cyclosporin A is an immunosuppressant, which can be used therapeutically to avoid tissue rejection during organ transplantation. Basigin was found to be transiently linked to the cyclophilin and thereby acting as a temporally-regulated receptor in processes central to cell signalling [52]. Another protein displaying basigin binding affinity is the calcium-binding protein called S100A9 (UniProtKB: P06702) that together forms the heterodimeric calprotectin protein complex [53]. When studying the interaction between basigin and platelet glycoprotein VI (GPVI), it was found that the adhesion and rolling of ADP-stimulated platelets were significantly enhanced on an immobilised basigin base. This indicates that basigin: GPVI interactions are crucial for platelet activation and function [54].

Human basigin is known to be over-expressed in a range of malignant tumours. The interactions between the glycans of basigin to other abundant proteins in the tumour micro-environment including galectin-3, hyaluronan, and integrins show that carbohydrate-protein interactions are, at least in part, responsible for the invasive properties, proliferation as well as survival of malignant cells [37]. The multiple interaction partners and the presence of various basigin isoforms that may receive different glycosylation signatures dependent on cell origin and (patho)physiological conditions may explain potential cell-specific functions of basigin in tumours.

1.3.5. Functions of basigin

Since basigin resides on the surfaces of immune cells, this glycoprotein is likely to play central roles in the inflammatory response e.g. regulate the release of pro- or anti-inflammatory cytokines. Cytokines manipulate the immune system by, for example, changing the activity of lymphocytes and thereby alter the inflammation outcome of affected tissues. Hence, basigin may directly or indirectly be involved in the regulation of immune responses. Recent studies have indicated that human basigin also suppresses T-cell receptor-dependent activation of T-cells developing in the thymus. Basigin also plays an important role in several other non-immune and non-tumourigenic processes

such as spermatogenesis [55, 56], fertilisation [55], embryo transplantation [56] and in neural functions e.g. vision, behaviour, memory and olfaction [34, 57]. Interestingly, human basigin was recently found to be the receptor of an invasive protein RH5 belonging to the reticulocyte-binding protein homologue (RH) in the malarial parasites *P. falciparum* [57, 58].

The most well understood function of basigin in cancer is its ability to induce the secretion of matrix metalloproteinases (MMPs) from neighbouring fibroblasts in the tumour environment. The MMPs form a family of 23 zinc-dependent enzymes that are involved in pathological and physiological processes by proteolytically degrading tissues and fibres in the extracellular matrix [57, 58]. The production and activation of MMPs are tightly regulated by complex mechanisms driven by cytokines, growth factors and hormones. Cancer is associated with high levels of MMP activity, which suggests increased expression of MMP-regulating molecules in the tumour tissue micro-environment that contribute significantly to the tumour initiation, growth and metastasis [48, 49]. However, the exact mechanism underpinning this pro-tumourigenic induction is still not known [57]. It was suggested that the structure of basigin partially resemble that of the MMPs, but the complete primary, secondary, tertiary and quaternary structures of basigin remain undocumented to allow an accurate comparison [59]. The MMP family comprises several key members including MMP-1, MMP-2, MMP-3, MMP-9 and MMP-11. Basigin can stimulate the production of various MMPs e.g. MMP-1 (interstitial collagenase), MMP-2 (gelatinase A) and MMP-3 (stromelysin 1); these MMPs have slightly different functions in the ECM [48-50]. Transcriptional studies showed that basigin is a prerequisite for MMP induction [45], but, the molecular mechanisms of this cancer- and immune-related glycoprotein driving the MMP secretion are not understood.

1.3.6. Protein glycosylation, an important tumour-associated PTM

Post-translational modifications (PTMs) are a diverse class of covalent additions of functional chemical groups to polypeptide chains that significantly increase the structural and functional diversity of the proteome. The addition and removal of PTMs can occur at any time during the protein “life cycle” by chemical or enzymatic processes. Proteolytic cleavage and disulphide bond formation after polypeptide translation are also considered as PTMs. Examples of PTMs of proteins driven by highly specific enzymatic reactions (e.g. kinases, phosphatases, glycosidases) are glycosylation, phosphorylation, nitrosylation, methylation, acetylation, lipidation and ubiquitination. The identification and understanding of the dynamic regulation of PTMs can open pathways to understanding the exact structure><function relationship of proteins and the cellular processes that underpin the complex biology of life [60, 61].

Protein glycosylation is a common type of PTM found in almost all living organisms. This family of complex PTMs involves the transfer of glycans (also known as oligosaccharides, sugars or complex carbohydrates) to a protein backbone (glycosyl acceptor). Activated (nucleotide-containing) sugars are most often the glycosyl donors of the glycosylation process, which can occur on both secreted and membrane proteins facing the extracellular environment or the lumen of organelles [62, 63]. The glycosylation process depends on several factors including the enzyme and substrate availability. Without the necessary enzymes, the glycosylation event cannot occur or will only form incompletely. Specific amino acid residues and their local sequence (primary structure) and conformation/orientation (secondary and tertiary structure) of the polypeptide chain affect the formation of glycosidic bonds and therefore the initiation of the glycosylation process. Finally, the trafficking speed and micro-environments in the biosynthetic route in the glycosylation machinery are important factors for the glycosylation event. After attachment, the glycans will, in turn, affect the secondary, tertiary and quaternary structure of the yet-to-be-fully-folded glycoprotein [62, 63].

Glycosylation frequently targets asparagine residues in restricted sequons known as *N*-glycosylation, NxS/T, $x \neq P$), **Figure 8**. Consensus-free threonine and serine residues known as *O*-linked (mucin type) glycosylation is also an abundant type of glycosylation [62-64]. In the *N*-glycosylation process, which is the focus of this thesis, the so-called *N*-glycans are attached to the amide nitrogen atom on asparagine side-chains of the protein.

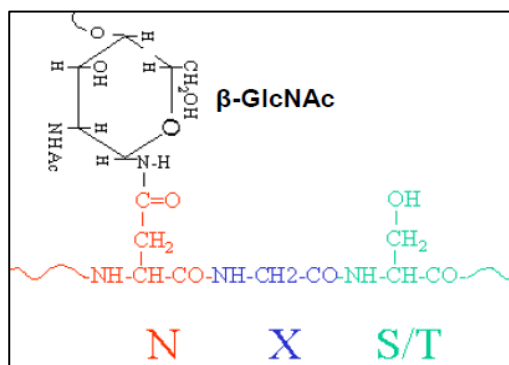


Figure 8: Protein *N*-glycosylation occurs in consensus NxS/T motifs ($x \neq P$) [4].

N-glycans comprise an enormous, but finite, spectrum of structures. The *N*-glycans are built on scaffolds (cores) with limited structural heterogeneity; the structural diversity is usually arising from variations of monosaccharides or larger glycoepitopes in the antenna region i.e. their non-reducing termini. The individual *N*-glycans consist of a limited set of monosaccharide residues that, however, can be linked in many different ways through various glycosidic bonds; the common building

blocks of *N*-glycans include fucose (Fuc), mannose (Man), galactose (Gal), *N*-acetylglucosamine (GlcNAc) and *N*-acetylneuraminic (NeuAc) residues [65-68]. The trimannosylchitobiose core (Man₃GlcNAc₂) common to all *N*-glycans reduces the structural heterogeneity and eases structural determination. The *N*-glycans cover three major types i.e. high mannose, complex and hybrid types, **Figure 9** [62-64]. High mannose glycans comprise Man and GlcNAc residues, whereas complex

glycans additionally may contain Gal, Fuc and NeuAc residues. A combination of high mannose and complex type features of the two glycan antennas give rise to the so-called hybrid type glycans.

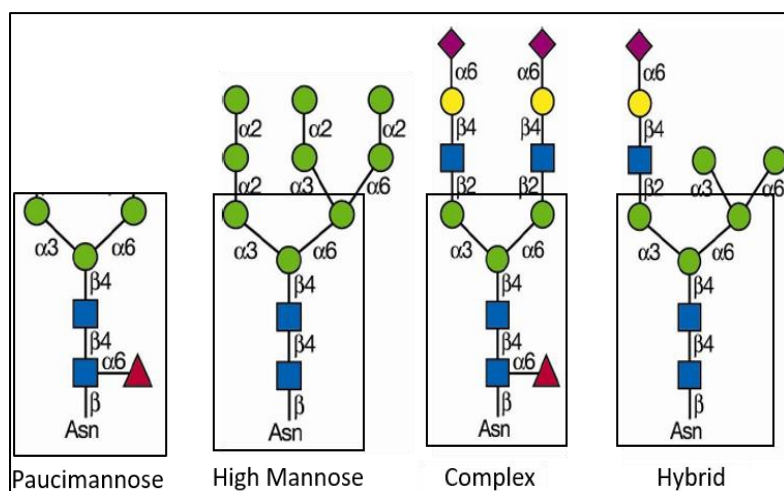


Figure 9: The three types of *N*-glycans i.e. high mannose, complex and hybrid type and the unusual paucimannose type. The common core at the non-reducing end is boxed. The non-reducing end of the glycan has a variety of monosaccharide residues and larger glycoepitopes that cap the glycans [6].

also expressed in humans under specific conditions including in inflammation and cancer as investigated in this thesis [48].

1.3.7. *N*-glycosylation of basigin

Although direct biochemical evidence is lacking, the literature supports the notion that human basigin only carries *N*-glycosylation [38]. Human basigin isoform 1 has three potential *N*-glycosylation sites i.e. Asn160, Asn268 and Asn302, **Figure 10**. The importance of human basigin *N*-glycosylation [44], and structure-function relationships [59, 65] of basigin have been reviewed. Evidently, substantial literature supports that *N*-glycans are directly and/or indirectly involved in important biological functions of basigin e.g. roles in tumour malignant transformation and metastasis, angiogenesis and chemo-resistance. As described above, basigin interacts with a range of inflammation-associated molecules e.g. integrins, cyclophilins, and E-selectin in these processes to mediate adhesion, invasion and metastasis of malignant tumour cells [18, 38, 44, 59, 70]. As mentioned previously, the molecular mass of glycosylated basigin is ~42-60 kDa, whereas unglycosylated basigin is 32 kDa [37, 41, 71]. Basigin obtained from various cells and tissues produced two distinct Coomassie and Western blot bands indicating that the protein exists in two major proteoforms (or “glycoforms”) produced by variations in the glycosylation.

A fourth class of *N*-glycan has recently been identified, also known as paucimannosidic type glycans. This unusual class of *N*-glycans comprises short chains of Man and GlcNAc residues with and without core fucosylation ($\text{Man}_{1-3}\text{GlcNAc}_2\text{Fuc}_{0-1}$), and show affinity towards mannose-binding lectin and other specialised lectin receptors [69]. Paucimannosylation has frequently been reported in invertebrate and plants, but recent reports indicate that paucimannosidic proteins are

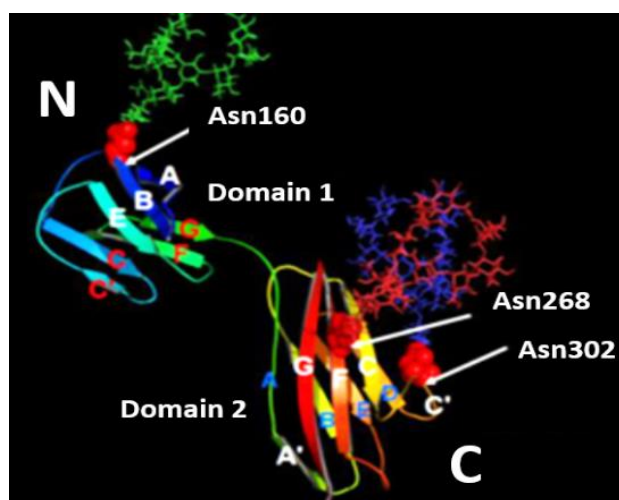


Figure 10: Molecular model (generated using Pymol) of the extracellular region of human basigin isoform 1. The three potential *N*-glycosylation sites at Asn160, Asn268 and Asn302 are indicated. The *N*-terminal domain is a C2 immunoglobulin and the *C*-terminal domain is an I-immunoglobulin, both consisting of β -barrel sheets and a conserved disulphide bond between the strands B and F [8].

Thus, the literature refers to a so-called “high glycosylation” form (HG-CD147) at 40-60 kDa presumably containing biantennary complex-type glycans that is sensitive to PNGase F, and a so-called “low glycosylation” form (LG-CD147) at ~32 kDa containing endoglycosidase H (Endo H)-sensitive high mannose type glycans [18, 38, 44, 59, 70]. Based on the knowledge of the general protein glycosylation process and the structural characteristics of basigin, these glycosylation signature differences can be assumed to relate to differences in the glycosylation initiation (site occupancy) and maturation processes (glycoform heterogeneity) of basigin as it travels through the glycosylation machinery in the ER and Golgi [38, 72]. Some researchers reported the presence of both HG-CD147 and LG-CD147 on the plasma membrane of breast carcinoma (MCF7 cells) [44, 70]. Only a fully complex type containing glycosylated form of basigin (HG-CD147) was found on the plasma membrane of HCC tumour cells [6, 44]. Systematic elimination of the three potential *N*-glycosylation sites (N44Q, N152Q and N186Q) by site-directed mutation in human embryonic kidney (HEK293) cells caused an equal decrease in the molecular mass of both HG-CD147 and LG-CD147 as evaluated by their increased mobility on electrophoretic gels. These observations indicated that all three *N*-sites are occupied on the polypeptide chain of basigin and that *N*-glycans on all sites contribute equally to the mature molecular mass [70, 73].

Given the extensive focus on basigin as a key driver and modulator of the tumour micro-environment and the substantial evidence of the biological involvement of basigin *N*-glycans in tumourigenic and immune functions, it is remarkable that the exact glycan structures attached to the potential *N*-glycan sites of human basigin expressed in cancer cells and neutrophils forming critical components of the tumour environment to date remain unidentified. In order to learn more about basigin glycosylation, our group recently mapped the site-specific *N*-glycan profile of HEK-derived human basigin isoform 1 using LC-MS/MS based glycopeptide analysis [74]. This provided

the first direct evidence of glycosylation of Asn160 and Asn268, but only of this recombinant (artificial) human basigin. Thus, the *N*-glycans on native basigin expressed by cells in the tumour environment are yet to be identified. This gap in our knowledge formed the rationale for the aims targeted in this thesis.

1.4. Aims of Research

This study comprised three major aims. The overall focus was to structurally characterise human basigin and the *N*-glycosylation from liver cancer cells and neutrophils forming key components of the liver cancer micro-environment. Specifically, this thesis aimed to:

1. To develop, optimise and validate analytical procedures for the isolation and characterisation of native human basigin using recombinant human (rh) basigin as a model.
2. To isolate and characterise native human basigin from primary neutrophils and liver cancer cell cultures using the developed methods in aim 1.
3. To perform detailed spatial *N*-glycome profiling of primary neutrophils and liver cancer cell cultures to show cell- and subcellular-specific differences in the *N*-glycosylation signatures expressed by key tumourigenic and immune components in the tumour micro-environment.

Targeting these three aims using a variety of techniques in cell and molecular biology and state-of-the-art methods for molecular characterisation (**Figure 11**) promised to yield insight into the molecular mechanisms underpinning the diverse functions of basigin in the liver tumour micro-environment. These aims would also provide important structural insight of the cell- and subcellular-specific *N*-glycosylation of the functionally different immune and liver cancer cells in the liver tumour micro-environment and thereby advance our knowledge of liver cancer glycobiology.

Chapter 2: Materials and Methods

2.1. Materials

2.1.1. Chemicals and reagents

All chemicals and control proteins (e.g. bovine fetuin and serum albumin) used for the described experiments (see **Figure 12**) were sourced from Sigma Aldrich/Thermo Fisher Scientific (Sydney, Australia) unless otherwise specified. Ultra-high-quality water was from a Millipore Academic Milli-Q system. Peptide-*N*-glycosidase F (PNGase F) and trypsin were from Promega (Sydney, Australia). Protease inhibitor (PI) cocktail tablets (cysteine and serine inhibitors) were from Roche Diagnostics (Sydney, Australia). Recombinant human (rh) basigin (product # CD7-H5222) from ACRO Biosystems

(Sydney, Australia) was stored in -30°C until use. Polymorphprep and Lymphoprep reagents for isolation of neutrophils from whole blood were from Axis Shield (Norway). Percoll (1.129 g/ml) was from GE Healthcare (Denmark). The HepG2 liver cancer cell line from American Type Culture Collection (ATCC® HB-8065™) were kept at -140°C until use. Cross-species antibodies against human basigin i.e. mouse monoclonal anti-CD147 [8D6] (ab194401) were from Abcam (Sydney, Australia).

2.2. Methods

Many methods were applied to achieve the described aims of the thesis, **Figure 11**. This workflow provides an overview of the applied methods and refers to the relevant method sections.

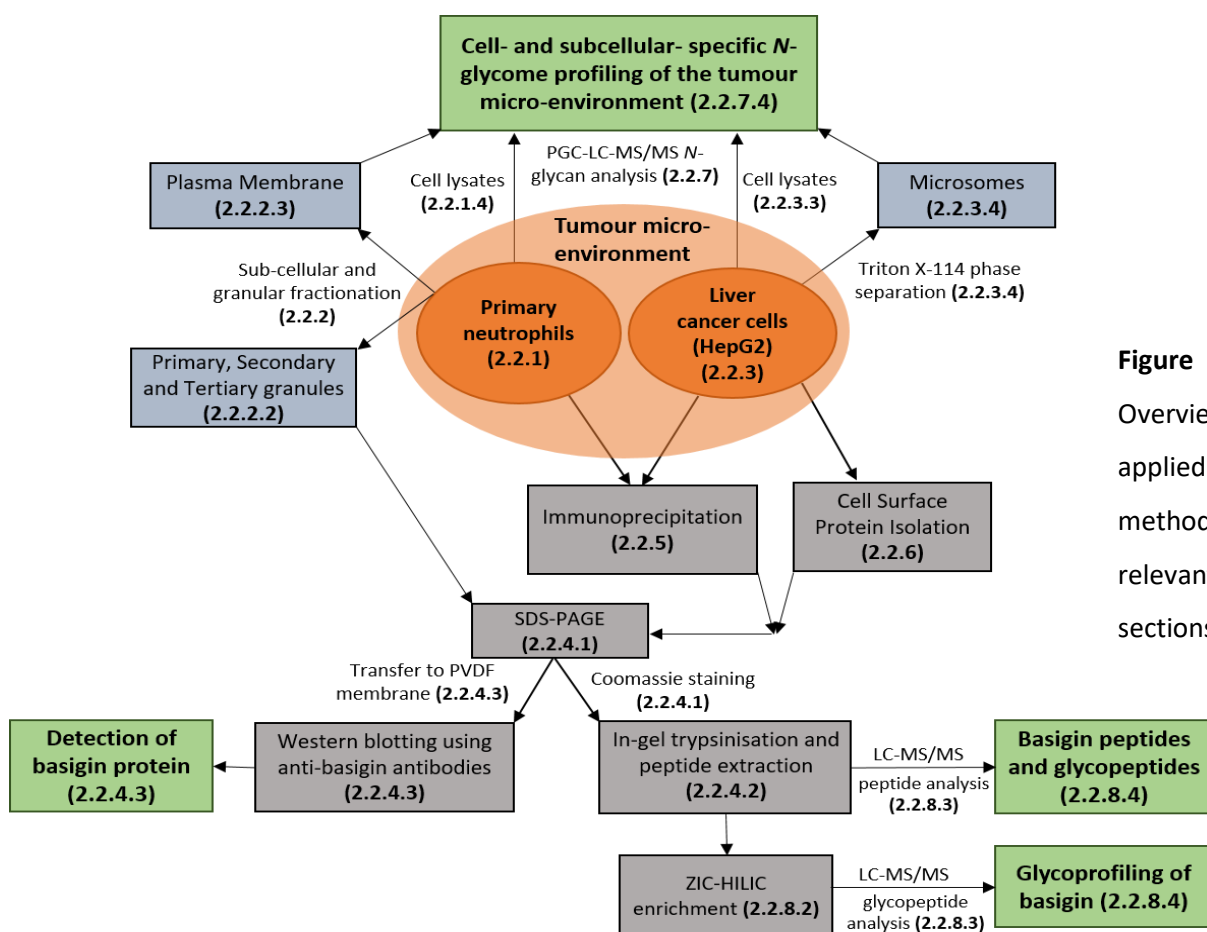


Figure 11:
Overview of
applied
methods and
relevant
sections.

2.2.1. Isolation of neutrophils from human blood

2.2.1.1. Drawing of blood

Peripheral blood (60-200 ml) was collected over multiple donations from the medial cubital vein of a healthy adult male donor by a trained phlebotomist after informed consent was obtained. The collection of blood and the subsequent isolation of resting neutrophils were approved by the Human Research Ethics Committee (Reference #5201500409) and BioHazard Risk Assessment (Non-GMO, Reference # NIP041214BHA) at Macquarie University, Sydney, Australia (**Appendix Figure S1**).

2.2.1.2. Neutrophil isolation using Polymorphprep and Lymphoprep

For comparative purposes, neutrophils were isolated using both Polymorphprep and Lymphoprep. For the Polymorphprep-based isolation, the freshly drawn blood was collected into ethylenediaminetetraacetic acid (EDTA)-sprinkled anti-coagulant tubes to prevent clotting of the blood [75] and layered with 1:1 (v/v) Polymorphprep separation media [13.8% (w/v) sodium diatrizoate and 8% (w/v) dextran] [76] in a 1:1 ratio at room temperature. The tubes were centrifuged on swing bucket at 550 g for 35 min at 20°C (max acceleration, no brakes). The plasma (top layer) and the mononuclear cells (MNCs) were carefully removed using a disposable Pasteur pipette. The polymorphonuclear (PMN) cells located on top of the highly dense red blood cells were mixed with cold sterile filtered aqueous 0.9% (w/v) sodium chloride (NaCl) and MilliQ water in 1:1 (v/v) ratio to restore the isotonicity and centrifuged at 400 g for 6 min at 4°C. The supernatant was removed and the cell pellet was re-suspended twice in cold sterile filtered MilliQ water and vortexed for 30 s to lyse any remaining red blood cells. For the Lymphoprep isolation, freshly drawn blood (60 ml for whole neutrophil analysis and 200 ml for subcellular fractionation) was collected into acetose citrose dextrose (ACD)-sprinkled anti-coagulant tubes and mixed with 2% (w/v) dextran in cold sterile filtered aqueous 0.9% (w/v) sodium chloride in a 1:1 (v/v) ratio by repeated inversion (at least 10 times) with incubation performed at room temperature to ensure efficient sedimentation. Differential sedimentation resulted in two clearly separated layers – a light yellow plasma-rich upper layer that was centrifuged at 200 g for 10 min at 4°C, and a red lower layer of dense RBCs. The pellet that contained MNCs and PMNs was re-suspended in 30 ml cold sterile filtered aqueous 0.9% (w/v) sodium chloride. 15 ml Lymphoprep separation media [9.1% (w/v) sodium diatrizoate and 5.7% (w/v) dextran] [77] was layered on top using a disposable Pasteur pipette. The tubes were centrifuged at 400 g for 30 min at 4°C, resulting in a MNC-rich supernatant, which was discarded and a PMN-rich pellet that was predicted to contain neutrophils [78]. The cell pellet was washed in cold sterile filtered MilliQ water until it appeared white as described above.

2.2.1.3. Determination of cell counts, viability and purity

The neutrophil counts and viabilities were determined at different dilutions using a Muse Cell Analyzer (Merck-Millipore, catalogue #0500-3115), which provides a measure of live and dead cells based on their differential side scattering [79]. Alternatively, a haemocytometer (BlauBrand, Germany) was also used to determine the neutrophil counts, however, the viability of cells was not obtained using this device [80]. The purity of the neutrophil population was estimated based on Wright-Giemsa staining of concentrated cells using cytospin centrifuge (Thermo) with Sorensen's

buffer for 12 min [81]. The slides were mounted in DPX mounting media and observed using a bright field microscope (Olympus Life Science) at 10x, 20x and 40x magnification. The purity of the isolated neutrophils (performed in triplicates) were determined by counting ≥ 100 cells in different fields and assessing their cell types based on morphology in each of the chosen microscopic fields.

2.2.1.4. Neutrophil lysis and protein precipitation and quantification

The pellets containing the granulocytic cells of interest was re-suspended with cold phosphate buffered saline (PBS) buffer, centrifuged at 14,000 g for 5 min at room temperature and the pellet was then taken up with 500 μ l radioimmunoprecipitation assay (RIPA) buffer [1 M Tris-HCl, pH 8.0, 5 M NaCl, 0.01% (v/v) Triton X-100, 0.01% (w/v) sodium deoxycholate, and 10% (v/v) SDS, PI tablets (Roche)] and incubated on ice for 20 min. The cells were lysed using probe sonication (Sonar) on ice at a 10% amplitude (3x) at 3 s intervals and centrifuged at 10,000 g for 10 min at 4°C. The supernatant containing the protein extracts was collected [82].

The total protein extracts of neutrophil lysates were mixed with chloroform, methanol and MilliQ water in volume ratio 8:4:3; and vortexed thoroughly. A white membrane protein-containing disc was formed on the methanol-water interface, with soluble protein in the upper aqueous layer and lipids and other non-polar compounds in lower chloroform layer. [83]. The concentration of the protein extracts was determined using a bicinchoninic acid assay (BCA) on a 96-well-plate according to manufacturer's instructions [84] with bovine serum albumin (BSA) as a protein standard prepared in various concentrations (0-2 mg/ml) [85]. The dried precipitated protein pellet was re-suspended in 8 M urea and diluted down to 2 M urea with MilliQ water in all samples to maintain compatibility with the assay [85]. The concentration of protein samples in the 96-well plate was measured in three technical replicates on a Fluostar Optima spectrophotometer (BMG Labtech) at 562 nm absorbance. Concentrations of protein samples of interest were determined from BSA standard curves.

2.2.2. Subcellular fractionation of blood-derived neutrophils

2.2.2.1. Disruption of neutrophils by nitrogen cavitation

The blood drawing and the isolation and validation of neutrophils were performed as described in **Section 2.2.1.1-3**. The subcellular fractionation, described below, follows a published protocol [86].

PI tablets (5 mM, final concentration) were added to the purified neutrophils and placed on ice for 10 min. The cells were then centrifuged at 200 g for 6 min at 4°C and the supernatant was discarded. Subsequently, 7 ml 1x disruption buffer [100 mM KCl, 3 mM NaCl, 3.5 mM MgCl₂, 1.5 mM ethylene glycol-bis (β -aminoethyl ether)-N,N,N',N'-tetraacetic acid (EGTA), 10 mM piperazine-N,N'-bis [2-

ethanesulfonic acid], 1 mM ATP (Na)₂ and 0.5 mM phenylmethylsulfonyl fluoride (PMSF) (all final concentrations)] was added to the pellet and vortexed. The Parr cavitation bomb (Parr Instruments, USA) was pre-cooled overnight at 4°C and 99.99% pure nitrogen (N₂) gas (at 375 psi) was introduced and left in the Parr bomb for 15 min. Next, the cell suspension was placed into the Parr bomb. The disrupted cells (called the “cavitate”) was collected in a drop-wise manner without introduction of bubbles in a parafilm covered fresh 50 ml falcon tube containing 100 mM EGTA. The cavitate was centrifuged at 400 g for 15 min at 4°C. Only the supernatant was collected for further use. [86, 87].

2.2.2.2. Three-/four-layer Percoll density gradient separation of granules

Percoll (1.129 g/ml) was mixed with a 10x disruption buffer (see above). Four densities of Percoll solution were prepared i.e. 1.03 g/ml, 1.11 g/ml, 1.09 g/ml and 1.12 g/ml to obtain a four-layered density gradient that can separate the different neutrophil granules and the basigin-rich plasma membrane of interest. The primary granules are reportedly the densest compartment (bottom layer after centrifugation) followed by secondary and tertiary granules, secretory vesicles and the plasma membrane (top layer). The gradient was packed in a polycarbonate centrifuge tube (Beckmann Coulter, 50 ml capacity); 9 ml of the 1.03 g/ml Percoll solution was dispensed using a 14G x 3^{1/4} inch needle followed by the fractions of increasing density. The supernatant obtained from the cell lysates was mixed in a 1:1 (v/v) ratio to the 1.11 g/ml Percoll solution and the gradient tubes containing all four Percoll density solutions were centrifuged at 37,000 g for 30 min at 4°C in a fixed angle rotor centrifuge (Sorvall RC-5B). After centrifugation, 1 ml fractions were collected from the bottom of the Percoll gradient into Eppendorf tubes using a peristaltic pump (Amersham) and an automated fraction collector (FRAC-200, Amersham). A total of 35 fractions were collected [86, 87].

A three-layer Percoll density gradient separation of disrupted neutrophil was also performed by our collaborator (Prof Anna Karlsson, Gothenburg Univ, Sweden). The three Percoll densities were 1.03 g/ml, 1.09 g/ml and 1.12 g/ml. The resultant Percoll layers contained the primary, secondary and tertiary granules and a combined secretory vesicles and plasma membrane as the top layer [86, 87].

2.2.2.3. Validation of plasma membrane isolation using latent alkaline phosphatase

Several marker assays measuring the activity of abundant enzymes specific to the individual neutrophil compartments were used to validate the subcellular isolation. This allowed pooling of the correct fractions. A latent alkaline phosphatase (LAP) assay was used to distinguish the basigin containing plasma membrane layer from the secretory vesicles in the neighbouring fractions on the basis of the AP activity in the presence and absence of a detergent i.e. 10% (v/v) Triton X-100. The assay was performed on all the fractions in two 96-well-plates using an alkaline sodium carbonate-

bicarbonate buffer in a 9:1 (v/v) ratio (pH 10.5) mixed with the phosphatase substrate (p-nitrophenyl phosphate disodium hexahydrate). Subsequently, 100 µl Triton X-100 was added to all the fractions in just one of the plates. The plates were incubated in tin foil at 37°C for 40 min, and then the optical density was read on a plate reader at a wavelength of 405 nm. LAP assay readouts (and other assays) informed which fractions were pooled together [86, 88].

In order to remove the LC-MS/MS-incompatible Percoll, the pooled plasma membrane fractions transferred into new tubes were ultra-centrifuged at 100,000 g for 90 min at 4°C [89]. This spin resulted in the formation of a hard thick pellet containing Percoll below a white disc of precipitated proteins that was carefully collected using a glass Pasteur pipette, transferred to fresh tubes, dried and stored in -30°C until further use [86].

2.2.3. Cell culture of liver cancer cells

2.2.3.1. Culturing, harvesting and monitoring HepG2 cells

The human liver cancer cell line HepG2 was thawed and cultured in Corning Cell Culture Flasks with Dulbecco's modified Eagle's medium (DMEM) (catalogue #30-2003) supplemented with 10% (v/v) fetal bovine serum (FBS) to provide nutrients to the cells and 1% (v/v) penicillin and streptomycin (100 U/ml each) to prevent microbial contamination in the media. Stable growth conditions were provided by an incubator i.e. constant 37°C and 5% of carbon dioxide (CO₂) level. When the confluence level reached 90-95% (72-96 h), sub-culturing was initiated by removing DMEM media, washing the monolayer cells with PBS and detaching it from the cell culture flask using 0.25% (w/v) trypsin and incubating for few minutes at 37°C. Fresh media was added in a 1:1 volume ratio to stop the trypsin cleavage and the cell suspension was redistributed to 4 fresh flasks and fresh media was added every 48 h [90]. To harvest the cells, the cell suspension (media and trypsin) was collected into fresh conical tube, spun down at 500 g for 6 min at room temperature and pelleted out. The cells were washed twice with cold PBS and the cell pellets were stored at -30°C. 20 µl cell suspension was mixed with 20 µl 0.4% (w/v) trypan blue dye. A fraction of the sample (25%, 10 µl) was loaded onto counting slides (BioRad, catalogue #145-0011) and cells were counted using an automated cell counter (BioRad, catalogue #AU390653). The cells were visualised under an inverted light microscope at 10-40x magnifications to monitor their morphology [91].

2.2.3.2. Lysates and microsomes of HepG2 cells and protein extraction

The pellet was thawed on ice, washed in 1 ml cold PBS buffer, vortexed and transferred into a Dounce homogeniser tube to gently disrupt the cells using a tight glass pestle into fine particles with minimal damage to cell nuclei [92] to obtain whole cell lysates. The frothy cell suspension was then

centrifuged at 14,000 g at 4°C for 12 min. The supernatant containing the cell lysates was collected and stored at 4°C. The pellet containing intact cells and nuclei was discarded. For enrichment of microsomal fractions, the pellet was re-suspended in a solution of 25 mM Tris-HCl (pH 7.4), 150 mM NaCl, and 1 mM EDTA with PI tablets. The cell suspension was ultra-sonicated using a Sonifier 450 on ice over three cycles of 10 s bursts and centrifuged at 2,000 g for 20 min at 4°C to remove unlysed cells and nuclei. The resultant supernatant was ultra-centrifuged at 120,000 g for 80 min to obtain the total membrane component containing pellet that was re-suspended in a buffer of 25 mM Tris-HCl (pH 7.4), 150 mM NaCl and 1% (v/v) Triton X-114. The suspension was phase partitioned on ice for 10 min, and then at 37°C for 20 min. Centrifugation at 1,000 g for 10 min at 25°C created two distinct layers. Only the lower detergent phase was kept for analysis [93].

For protein precipitation, ice-cold acetone was added to both the HepG2 cell lysates (1:4, v/v) and microsomal (1:9, v/v) fractions. The tubes were vortexed thoroughly and incubated overnight at -20°C in an upright position. The next day, the tubes were centrifuged at 14,000 g for 12 min resulting in an acetone-rich supernatant, which was discarded and a protein-rich pellet, which was left to air-dry for about 12 min and stored in -30°C until further analysis [94]. The protein quantification was performed using a BCA assay as described in **Section 2.2.1.5** [84].

2.2.4. Downstream handling of neutrophil and HepG2 protein mixtures

2.2.4.1. SDS-PAGE and staining of protein mixtures

NuPAGE 4-12% gradient bisacrylamide-Tris (Bis-Tris) mini protein gels (catalogue #NP0336BOX; 1.5 mm, 15-well, 25 µl loading volume/lane) were used for multiple SDS-PAGE experiments. These gels have a wide separation range (15-260 kDa) when used with a 3-(*N*-morpholino)-propanesulfonic acid (MOPS) running buffer at constant 200 V with variable current (50-100 mA) for 50 min. Generally, 30-60 µg total protein treated with 500 mM dithiothreitol DTT (10X) and NuPAGE sample loading buffer (4X) was loaded into each gel lane after heating at 70°C for 10 min. A molecular marker (Chameleon Duo Pre-stained protein ladder, LI-COR, catalogue #928-6000) was used to indicate the molecular mass of the separated proteins [95, 96]. After gel electrophoresis; the gels were washed in MilliQ water, stained with Coomassie stain solution [40 ml dye, 10 ml diluent] and covered in tin foil to be incubated on a shaker overnight at room temperature [97, 98]. Next day, the stain was poured off and the gel was de-stained with 1% (v/v) acetic acid and then with MilliQ water. Gel images were taken on an Odyssey Imaging Systems (LI-COR Biosciences).

2.2.4.2. In-gel trypsin digestion, and peptide extraction and clean-up

Protein bands of interest were excised from SDS-PAGE gels, cut in small pieces and de-stained further, reduced with DTT and alkylated with iodoacetamide (IAA), and washed thoroughly in MilliQ water. Protein digestions were performed with porcine trypsin (1:30 enzyme/substrate w/w ratio) in 25 mM ammonium bicarbonate (ABC) (pH 8.0) at 37°C for 18-20 h. The resulting peptides were extracted using successive rounds of solvent extractions using acetonitrile (ACN) and formic acid (FA) and then dried using a vacuum centrifuge. The dried samples were re-suspended in 1% FA for reversed-phase solid phase extraction (RP-SPE) employing pipette tips containing an immobilised C18 resin (Merck-Millipore) primed with 60% ACN and then washed with 0.1% FA. The retained peptides were eluted using firstly 50 µl 60% ACN/0.1% FA and then 50 µl 90% ACN/0.1% FA. The samples were dried in a vacuum centrifuge and stored at -30°C until LC-MS/MS analysis [99].

2.2.4.3. Protein detection using dot-blotting and Western-blotting

Dot-blotting and Western blotting are similar techniques used for a range of applications including protein identification. Dot-blotting was performed to validate the reactivity of the anti-basigin antibodies used in this project. Two replicates of each 2 µg rh basigin and the control protein bovine fetuin were reduced and alkylated with DTT (70°C, 10 min) and IAA (in dark, 30 min), dot-blotted in discrete spots on a nitrocellulose membrane and dried overnight at room temperature. Non-specific binding sites of the membrane were first blocked with 5% BSA in Tris buffered saline (TBS) at 4°C for 1 h, followed by incubation with primary antibodies (**Section 2.1.2**) [1:1000, in 5% BSA in TBS-Tween 80 (TBS-T) at 4°C for 1 h. The membranes were then treated with secondary anti-mouse antibody (1:14,000, in 5% BSA in TBS-T) and incubated at room temperature for 30 min in the dark [100]. Images of the membrane were taken on an Odyssey Imaging Systems (LI-COR Biosciences). Western blotting is usually performed after gel electrophoresis when specific protein bands of particular interest require further identification and/or quantitation. Proteins separated on washed mini gels were transferred to an ethanol-activated PVDF membrane and handled as described above [100].

2.2.5. Immunoprecipitation

Immunoprecipitation (IP) is a widely used method to isolate and analyse protein targets in complex mixtures using specific antibodies conjugated to Protein A or G agarose beads [101]. An IP kit (Protein G beads) and associated reagents were used in this project (Roche, catalogue #1719386). Cell pellets obtained from isolated neutrophils and HepG2 cells were lysed using lysis buffer (from the kit) in a Dounce homogeniser as described in **Section 2.2.3.2**. To reduce non-specific adsorption of proteins to the beads, a pre-clearing step was performed by adding 50 µl Protein-G agarose beads

to the sample of interest and incubating it on a rocking platform at 4°C overnight. After centrifugation (12,000 g for 20 s) in a microfuge at room temperature, the pellet was stored in -20°C until further use. The supernatant was incubated with primary antibody on a rocking platform at 4°C for 3 h. The antigen-antibody complex was IP'ed by adding 50 µl Protein-G agarose beads to the sample. Incubation was performed overnight on a rocking platform at 4°C. After repeated rounds of washing with buffers and centrifugation, the pellets containing the IP'ed proteins were collected. All supernatant wash fractions were stored for analysis. IP'ed proteins were eluted from the Protein-G beads using NuPAGE sample loading buffer (LDS, 4X) containing 500 mM DTT and separated by SDS-PAGE (**Section 2.2.4.1**). Proteins were identified by Western blotting (**Section 2.2.4.3**) and mass spectrometry (**Section 2.2.4.2**).

2.2.6. Cell surface protein capture and handling

Cell surface proteins displaying accessible lysine residues (primary amines) can be selectively labelled and thereby isolated using biotin-neutravidin affinity chromatography [102-105]. The cell surface capture kit (Thermo, catalogue #89881) was used to capture cell surface protein (including plasma-membrane resident basigin) from HepG2 liver cancer cells. Four flasks (T75 cm²) of 95% confluent HepG2 cells were prepared and incubated with 10 ml Sulfo-NHS-SS-Biotin solution on a rocking platform at 4°C for 1 h. The biotinylation was quenched and the cells were gently scraped off the walls of the flask and pooled together before centrifuging at 500 g at 4°C for 5 min. The resulting pellet in lysis buffer was lysed on ice using probe sonication at 10% amplitude (5x, 1 s bursts) and centrifuged at 10,000 g at 4°C for 5 min and the supernatant was collected. Next, the neutravidin agarose resin was packed in the column and the sample was repeatedly loaded and incubated with end-over-end mixing at 4°C for 2 h for maximum recovery. The enriched proteins of interest were eluted with NuPAGE sample loading buffer (LDS, 4X) containing 500 mM DTT, analysed using SDS-PAGE (**Section 2.2.4.1**) and mass spectrometry (**Section 2.2.4.2**).

2.2.7. N-glycomics

2.2.7.1. N-glycan release and clean-up

The employed *N*-glycan release protocol followed closely the method published by Jensen *et al.*, 2012 [73]. In short, 25 µg protein obtained from biological samples e.g. lysates of neutrophils and HepG2 cells (n = 3 technical replicates), as well as the standard *N*-glycoprotein i.e. bovine fetuin, were reduced and alkylated and immobilised and visualised on ethanol-activated PVDF membranes using 0.1% (w/v) Direct-Blue Stock 71. Protein spots were transferred into a 96-well plate. After blocking for non-specific binding and washing, 2.5 U PNGase F (Promega) was applied to each spot

in 10 µl MilliQ water. *N*-glycans were released overnight by incubation at 37°C [106] and then incubated in 100 mM ammonium acetate (NH₄CH₃CO₂) (pH 5.0) at room temperature for 1 h to remove glycosylamines from the reducing end of the glycans. Samples were then dried.

The dried glycans were reduced with 0.5 M sodium borohydride (NaBH₄) in 50 mM aqueous potassium hydroxide (KOH) at 50°C for 3 h. The reduction reaction was quenched by the addition of 2 µl glacial acetic acid (CH₃COOH). Removal of the high sodium amounts from the samples was performed by AG 50W X8 strong cation exchange resin (200-400 mesh) (30 µl bed volume) (Bio-Rad) packed into C18 stage tips (Merck-Millipore). Glycan samples were loaded onto the columns and the flow-through were collected by passing 50 µl water through the column twice, dried in a vacuum centrifuge, dissolved in 100% (v/v) analytical-grade methanol and re-dried to remove residual borate. For further clean-up, the samples were re-suspended in water and subjected to porous graphitised carbon (PGC) (Bio-Rad) on SPE. The PGC resin was packed onto C18 stage tips and washed with 90% ACN in 0.05% aqueous trifluoroacetic acid (TFA) followed by 40% (v/v) ACN in 0.05% (v/v) aqueous TFA and MilliQ water. The glycans were repeatedly loaded onto PGC columns, washed in MilliQ water and eluted using two rounds of 25 µl 40% (v/v) ACN in 0.05% (v/v) TFA. The eluted glycans were dried and stored at -20°C until LC-MS/MS analysis [73].

2.2.7.2. PGC LC-MS/MS of *N*-glycans

The dried glycans were reconstituted in 12.5 µl MilliQ water, centrifuged at 14,000 g for 10 min at 4°C and the supernatant layer was transferred into high recovery glass vials (Waters) for LC-MS/MS analysis. The *N*-glycans were separated by capillary PGC LC (Hypercarb KAPPA capillary column, 180 µm inner diameter x 100 mm length, 5 µm particle size, 200 Å, Thermo, Australia) using an Agilent 1260 Infinity HPLC instrument. A linear gradient of 0-45% (v/v) ACN and 10 mM (ABC) over 85 min at a constant flow rate of 2 µl/min was applied. Glycans were ionised using ESI and detected using a 3D ion trap mass spectrometer (LC/MSD Trap XCT Plus Series 1100, Agilent Technologies, Australia). The full scan acquisition range was kept constant at *m/z* 200-2,200 in negative ionisation polarity mode. Data-dependent acquisition was enabled; the top three most abundant precursors in each full scan spectrum were selected for MS/MS using resonance activation CID fragmentation. The specific LC-MS/MS acquisition details are described in **Appendix Table S1**.

2.2.7.3. Data analysis of released *N*-glycans

The glycan LC-MS/MS raw data was browsed and structures were solved manually using Bruker Compass Data Analysis v4.0 (Bruker Daltonics, Germany). The characterisation of the individual *N*-glycans was based on accurate matches of the experimental and theoretical monoisotopic precursor

masses and the annotation of glycan fragment ions in the corresponding CID-MS/MS spectra according to the established nomenclature of glycan fragmentation [107]. The Expasy GlycoMod tool [108] and GlycoWorkbench v2.1 [109] were used to aid the structural characterisation. Our in-house knowledge of the PGC-LC retention patterns for reduced *N*-glycans was used as a powerful orthogonal identifier in addition to the MS and MS/MS data [73, 110-113].

2.2.8. Proteomics and *N*-glycoproteomics

2.2.8.1. ZIC-HILIC glycopeptide enrichment

The protein extracts obtained from neutrophil and HepG2 cell lysates were separated by SDS-PAGE, stained with Coomassie Blue and digested with trypsin as described in **Sections 2.2.4.1-2**. The resulting peptide mixtures were analysed by LC-MS/MS with or without glycopeptide enrichment.

Glycopeptide enrichment followed closely our published method [114]. In short, the dried peptide mixtures were reconstituted in 30 µl 80% ACN in 1% aqueous TFA, vortexed and kept on ice until use. Custom-made micro-columns were prepared by packing zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC) stationary phase (Merck Millipore, 5 µm particle size, 200 Å, 2 cm) in a SPE format on top of a C18 disc in gel loading tips. The columns were equilibrated in 50 µl mobile phase containing 80% ACN in 1% aqueous TFA. The peptides were loaded repeatedly onto the column and then eluted using 2 x 50 µl 1% TFA. All resulting fractions i.e. enriched, unenriched, flow-through fractions were dried in a vacuum centrifuge and stored at -20°C until LC-MS/MS [72].

2.2.8.2. LC-MS/MS of peptides and *N*-glycopeptides

Peptides were reconstituted in 10 µl 0.1% (v/v) FA, and centrifuged at 14,000 g for 10 min at 4°C. The supernatants were transferred into high recovery glass vials (Waters) and analysed on ESI-MS/MS using an HCT 3D ion trap (Bruker Daltonics) coupled to an Ultimate 3000 LC (Dionex) in positive polarity mode. The C18 column (Proteocol HQ303, 10 cm length, 3 µm particle size and 300 Å pore size) was equilibrated in 100% solvent B (0.1% FA in 99.9% ACN) for 30 min followed by 100% solvent A (aqueous 0.1% FA) for 1 h. Samples were analysed at 5 µl/min. The gradient was as follows: 0-30% (0.5%/min slope) of solvent B over 60 min and 30-60% (4.2%/min slope) of solvent B over 7 min. Full MS scans (m/z 300-2,200) at a scan speed of 8,100 $m/z/s$ were followed by data-dependent fragmentation of the three most abundant precursor ions using resonance activation CID-MS/MS fragmentation. The mass spectrometer was calibrated using a tune mix (Agilent Technologies). The mass accuracy of the precursor and product ions was typically better than 0.5 Da.

Alternatively, the glycopeptide samples were analysed on an Orbitrap Fusion Tribrid mass spectrometer coupled to a Dionex UPLC system in collaboration with Dr. Benjamin Parker, Univ Sydney. The peptides were separated on a 75 μm x 15 cm custom-made UPLC column packed with 1.7 μm , 150 Å, BEH C18 (Waters) through a similar binary solvent system and linear gradients as described above. The samples were analysed at a flow rate of 300 nl/min. The mass spectrometer was operated in data-dependent mode switching between MS (m/z 300-1,800) and MS/MS of the top 10 precursors with highest charge states. Three fragment scan type were done – CID at 35% NCE, ETD performed with calibrated charge-dependent reaction time and HCD activation at 15% NCE. Ion species were analysed in the Orbitrap at a resolution of 60,000 FWHM (at m/z 200) [115].

2.2.8.3. Data analysis of N-glycopeptides

The (glyco) peptide LC-MS/MS raw data was browsed using Xcalibur v3.0 (Thermo). Some structural aspects of the glycopeptides were solved manually with support from GPMAW v10.0 (Lighthouse, Odense, Denmark) [116]. The protein sequences of all human basigin isoforms were used in databases such as Mascot v2.5 (Matrix Science) [117], Byonic (Protein Metrics) [118] and Proteome Discoverer v2.10 [119] with a 1% false discovery rate (FDR) [120] to identify peptides and glycopeptides. The oxonium ions i.e. m/z 204.08, 366.14, 512.10, 528.12, 657.23 [121] indicated the presence of glycopeptides and specific (sub-)structural features [121-123].

Chapter 3: Results

In this project, primary human neutrophils were purified from whole blood and HepG2 liver cancer cells were cultured in order to facilitate isolation and characterisation of basigin and the *N*-glycome from cancerous and immune cell components of the liver tumour micro-environment. A multitude of approaches were applied to isolate cells and basigin-rich subcellular compartments prior to LC-MS/MS-based molecular characterisation. The individual steps were performed and validated using many molecular and cellular methods and strategies, see **Figure 11** for overview of used workflow.

3.1. Reactivity of Anti-Basigin Monoclonal Antibodies (mAB) to Basigin

To confirm basigin reactivity of our anti-basigin antibodies for downstream WB and IP, dot-blotting experiments were performed using recombinant human (rh) basigin and a mouse monoclonal anti-human basigin primary antibody (mAB). Immobilised rh basigin was incubated with primary and secondary antibodies or just the secondary antibody to control for unspecific reactivity, **Figure 12**.

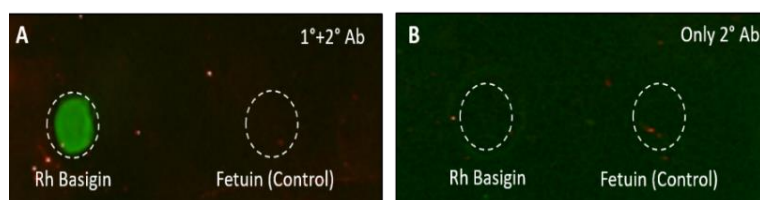


Figure 12: Dot blotting of rh basigin and bovine fetuin (negative control) by application of **(A)** 1° and 2° anti-human basigin mABs and **(B)** the 2° mAB alone.

The model glycoprotein fetuin was included as another negative control. These experiments demonstrated that the mouse anti-human basigin mAB, as expected, reacted with antigens on rh basigin and importantly showed compatibility with, and absence of unspecific reactivity of, the secondary antibody. These antibodies were thus considered suitable for WB and IP. The rabbit polyclonal anti-human CD147 (ab64616) antibody was also tested for rh basigin reactivity. Due to lack of reactivity this antibody was excluded in subsequent WB and IP experiments (data not shown).

3.2. Isolation and Characterisation of Basigin from Neutrophils

3.2.1. Isolation, cell count, viability, purity and morphology of neutrophils

For comparative purposes, resting neutrophils were first isolated from human donor blood using both Polymorphprep and Lymphoprep reagents. The characteristic neutrophil morphology i.e. multi-lobed nuclei were visually confirmed using bright field microscopy of Giemsa-Wright stained neutrophils isolated using both reagents, **Figure 13**.

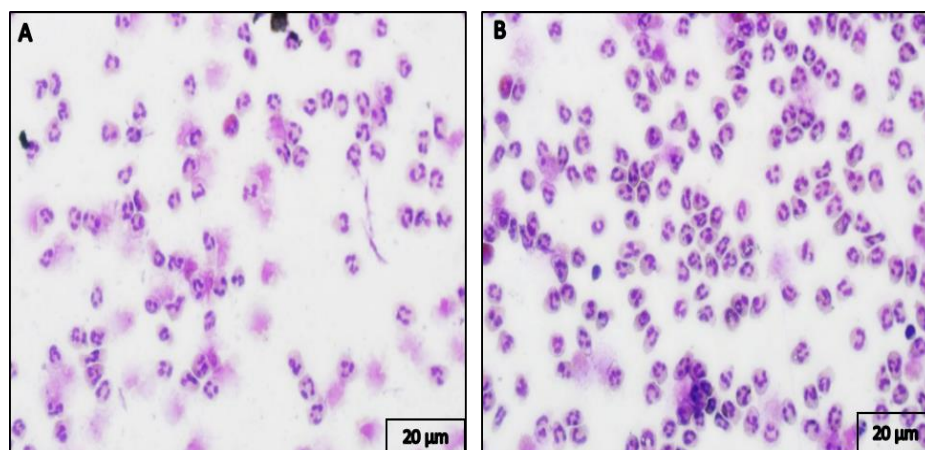


Figure 13: Multi-lobed nuclei morphology of resting neutrophils after isolation. Bright-field microscopy (20x magnification) of Giemsa-Wright stained neutrophils isolated using **(A)** Polymorphprep and **(B)** Lymphoprep reagents.

The neutrophil yields, viabilities, and purities were quantified using an automated cell counter and by microscopy, (**Appendix Figure S2**). Both preparations showed very high neutrophil viability and purity (>90-95%), but a significantly higher neutrophil yield was achieved using Lymphoprep (3.5×10^7 cells/ml blood) compared to Polymorphprep (5.9×10^6 cells/ml blood). Taken together, both of these isolation methods were evidently useful for efficient isolation of resting neutrophils displaying intact morphology and high viability and purity. However, due to the superior yield, Lymphoprep was used for the downstream subcellular isolation and basigin characterisation.

3.2.2. Basigin-centric IP of neutrophil lysates

IP of neutrophil lysates was performed in attempts to isolate neutrophil basigin. The protein-rich chloroform-methanol-water interface (the fraction expected to harbour basigin) of the IP fractions were run on two matching SDS-PAGE gels; one was Coomassie blue stained and protein bands of interest were identified using LC-MS/MS after in-gel trypsin digestion and one was used for basigin-focused WB, **Figure 14**.

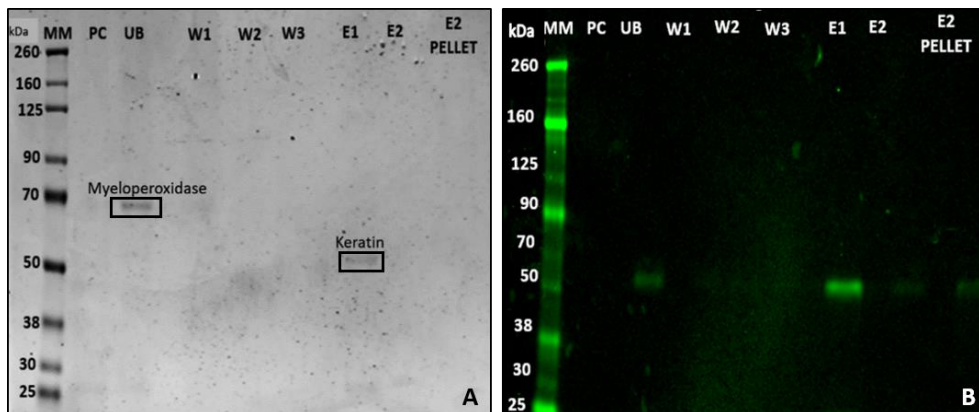


Figure 14: IP fractions from neutrophil lysates on matching gels after **(A)** Coomassie and **(B)** WB. MM, molecular marker, PC, pre-clear fraction, UB, unbound supernatant, W1-3, wash fractions, E1-2, elution fractions and E2 pellet, antibody-bound protein G beads. Bands (boxed) were identified using LC-MS/MS.

Surprisingly few proteins were observed in PC and UB indicating less-than-expected protein extract being applied to the gels e.g. resulting from protein degradation. The characteristic neutrophil proteins, myeloperoxidase (MPO) [86] were identified in UB by ion trap LC-MS/MS and Mascot and Byonic searches, **Table 1**. Although the 55 kDa protein band in E1 showed strong reactivity to the anti-basigin mAB, no basigin was identified in this fraction by LC-MS/MS most likely due to masking/suppression by the more abundant keratins and/or too little protein starting material. Alternatively, the heavy chain of IgG (~55 kDa) used in the IP experiment may have reacted unspecifically with the secondary antibody. The aqueous (hydrophilic) and chloroform (hydrophobic) fractions from the chloroform-methanol-water preparation were also investigated, but also without convincing evidence of basigin (data not shown). Hence, a different approach was attempted to isolate sufficient neutrophilic basigin for molecular characterisation.

Band	UniProtKB Acc. #	Protein name	Score	Coverage (%)	# Proteins	# Peptides	MW (kDa)
UB	P05164	Myeloperoxidase	520.8	10.40	2	7	83.8
E1	P35527	Keratin, Type I Cytoskeletal	1183.4	12.84	2	6	62.0

Table 1: LC-MS/MS-based protein identification from UB and E1 IP fractions (only most abundant protein from each band shown).

3.2.3. Isolation of the basigin-rich plasma membrane fraction

As the transmembrane basigin resides predominantly in the plasma membrane [86], enrichment of the plasma membrane fraction of neutrophils was attempted after gently disrupting the cells via nitrogen cavitation. A disruption efficiency of ~95% was achieved based on the relative MPO activity in the supernatant (disrupted) and pellet (undisrupted cells) after centrifugation (data not shown) [124]. The released, still intact, neutrophil granules, and the resealed plasma membrane were separated by high speed centrifugation in a four layered Percoll density gradient, **Figure 15**.

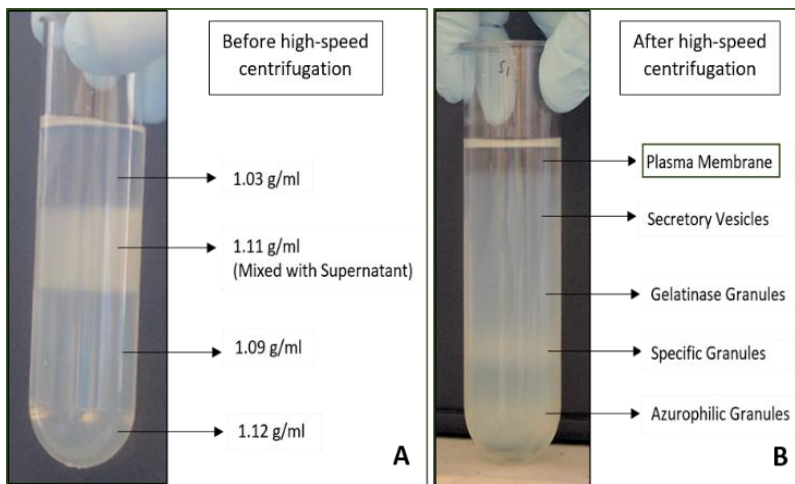


Figure 15: Four-layer Percoll density gradient **(A)** before and **(B)** after high speed centrifugation. The densities are indicated. Disrupted neutrophils were mixed with the 1.11 g/ml fraction before centrifugation. The basigin-rich plasma membrane was isolated as the top layer (see also **Figure 16**).

The granule separation yielded 35 fractions (1 ml/fraction) [86]. An MPO activity assay indicated that fraction 1-6 contained the primary granules (not described further), **Figure 16**. LAP activity assay determined at 405 nm absorbance (detects p-nitro phenol phosphate) [86, 124] showed significance activity in fractions 27-35 that formed the basigin-rich plasma membrane fractions.

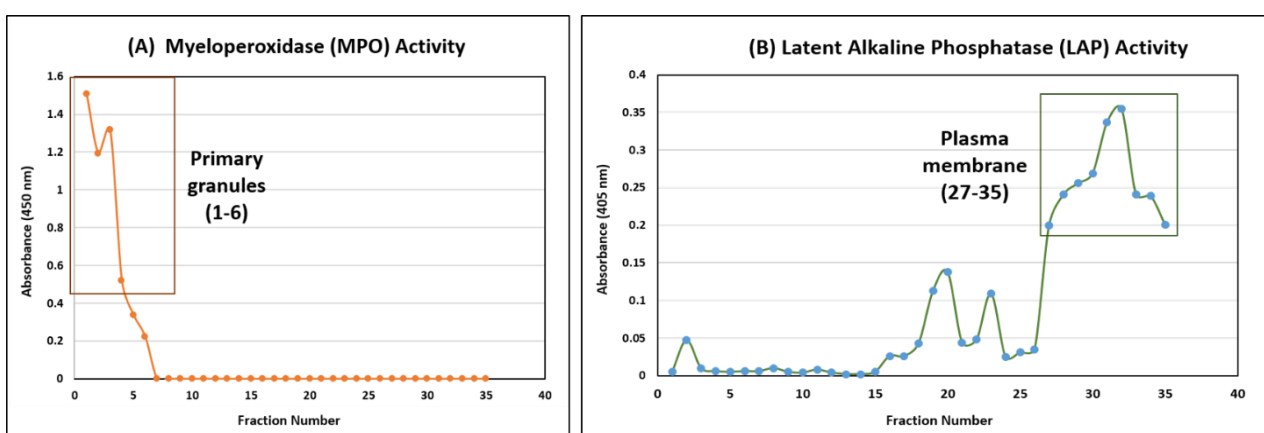


Figure 16: Enzyme activity assays were performed on all 35 subcellular fractions to identify the separated granules and the basigin-rich plasma membrane of neutrophils. **(A)** MPO activity in fraction 1-6 indicated primary granules. **(B)** Significant LAP activity in fraction 27-35 identified the plasma membrane. Average absorbance readings of technical duplicates ($n = 2$) are given for both assays.

3.2.4. Identification of plasma membrane-basigin using (glyco)-proteomics

Before acquiring and analysing the LC-MS/MS proteomics data to identify basigin, theoretical tryptic peptides of the four known isoforms of human basigin were calculated using the GPMW sequence tool, **Table 2**. Many peptides common to all basigin isoforms were found, leaving only few isoform-specific peptides to distinguish between basigin variants in particular the less common isoform 2-4.

Common theoretical tryptic peptide sequences of basigin found in		Theoretical tryptic peptide sequences of basigin found only in each isoform			
Isoforms 1,2,3 and 4	Isoforms 1 and 2	Isoform-1	Isoform-2	Isoform-3	Isoform-4
R.VK.W K.AVK.S K.SSEHINEGETAMLVCK.S K.SESVPPVTDWAWYK.I K.ITDSEDK.A K.ALMNGSES.R (Asn268) R.FFVSSSQGR.S R.SELHIENLNMEADPGQYR.C R.CNGTSSK.G (Asn302) K.GSDQAIITLR.V R.VR.S R.SHLAALWPFGLGIVAEVLVLTIIYIEK.R R.KPEDVLDDDDAGSAPLK.S K.SSGQHQNCK.G K.GK.N K.NVR.Q R.QR.N R.NSS-(COO ⁻)	K.ILLTCSLNDSATEVTGHR.W (Asn160) R.WLK.G K.GGVVLK.E K.DALPGQK.T K.TEFK.V	(NH ₂)-AAGFVQAPLSQQR.W R.WVGGSVELHCEAVGSPVPEIQWW FEGQGPNDTCSQLWDGAR.L R.LDR.V R.VHIHATYHQHAASTISIDTLVEEDTG TYEGR.A R.ASNDPDR.N R.NHLTR.A R.APR.V K.WVR.A R.AQAVVLVLEPGTVFTTVDELGSK.I	(NH ₂)-AAGTVFTTVDELGSK.I 	(NH ₂)-MGTANIQLHGPPR.V 	(NH ₂)-MK.Q K.QSDASPPER.V

Table 2: Theoretical tryptic peptides of the four isoforms of basigin. The common peptides shared between isoform 1-4 and isoform 1-2 (left), and the isoform-specific peptides (right) are indicated.

To remove Percoll and ensure clean peptide samples for LC-MS/MS analyses, enriched plasma membrane proteins (~60 µg) were ultra-centrifuged and run 1 cm into SDS-PAGE gels (**Appendix Figure S3**). The non-separated proteins were excised *en bloc*, washed and trypsin digested.

The resulting desalted peptide mixture was analysed using a traditional data-dependent proteomics strategy on a high resolution/high mass accuracy Orbitrap Fusion Tribrid mass spectrometer with HCD fragmentation. This experiment facilitated the positive identification of 795 non-redundant proteins using conventional proteomics type settings on a Byonic search engine (**Appendix Table S2**). Human basigin-1 (isoform 2) (Byonic score 710.1, 275th ranked protein by confidence) and many other proteins were confidently identified in the enriched plasma membrane fraction by detection of five unique peptides belonging to basigin, **Table 3** (see **Appendix Figure S4** for MS/MS).

UniProtKB Acc. #	Protein Name	Byonic Score	Sequence Coverage (%)	# Identified Non-Redundant Peptides	Observed MW (kDa)	Expected MW (kDa)
P02788	Actin, Cytoplasmic 1	1510.7	84.00	88	69.2	78.1
P12814	Alpha-Actinin-1	1333.7	70.63	139	101.5	103.0
P00738	Haptoglobin	1250.2	61.50	36	45.0	45.2
P14780	Matrix Metalloproteinase 9	1136.0	61.24	122	76.3	78.5
P04040	Catalase	1194.0	61.86	38	59.7	59.7
P02787	Serotransferrin	1106.1	63.75	84	77.6	77.6
P35579	Myosin-9	1091.5	52.96	166	223.6	226.5
O14745	Na ⁺ /K ⁺ Exchange Factors	1079.5	55.95	20	38.4	38.8
P27797	Calreticulin	1046.3	52.11	68	48.2	48.2
P35613	Human Basigin-1 (Isoform 2)	710.1	19.33	5	19.3	29.2

Table 3: Shortlist of most abundant proteins identified in the neutrophil plasma membrane using LC-MS/MS proteomics. Human basigin isoform 2 was confidently identified.

The identification of basigin isoform 2 was very exciting and the data was therefore scrutinised further by manual interrogation for confirmation. The presence of theoretical tryptic peptides from the isoform 2 was explored. Four additional peptides (R.WLK.G, K.TEFK.V, K.ITDSEDK.A, R.FFVSSSQGR.S) could be observed but at very low abundance without MS/MS confirmation. With a total sequence coverage of 41.5%, the nine peptides covered extensively the isoform 2, **Figure 17**.

The peptide mixture was also analysed after glycopeptide enrichment on the same Orbitrap Fusion instrument, but using glycoproteomics tailored acquisition. Alternating HCD, EThcD and CID fragmentation modes were applied [125]. No additional non-modified or PTM-modified basigin peptides including any *N*-glycopeptides were identified through this strategy even after 1) broadening the search criteria on Byonic (one of only few software capable of identifying intact glycopeptides) [74] e.g. including semi-tryptic conditions and inclusion of multiple variable PTMs and 2) manually interrogating the LC-MS/MS data for likely basigin (glyco)peptides (data not shown).

AAGTVFTTVEDLGSKILLTCSLNDSATEVTGHRWLKGGVVLKEDALPGQKTEFKVDSDDQWGEYSCVFLPEPMGTANIQLHGPPRVKAVKSSSEHINEGETAMLVCKSESVPVPTD
WAWYKITDSEDKALMNGSESRRFFVSSSQGRSELHIENLNMEADPGQYRCNGTSSKSGSDQAIITLRVRSHLALWPLGIVAEVLVLTIIIFIYEKRRKPEDVLDDDDAGSAPLKSSGQ
HQNDKGKGNVRQRNSS

Coverage Summary: 41.5% (103 out of 248)

#	Peptide Sequence	Byonic Score	Modification Types	m/z	Charge (z)	Observed [M+H] ⁺ (Da)	Calculated [M+H] ⁺ (Da)	Start-End AA	Cleavage	Scan Time (min)	Comment
1	(NH2)-AAGTVFTTVEDLGSK.I	592.6		748.3	2	1495.8	1495.8	22-36	NRagged	110.4	Isoform-2 Specific
2	R.WLK.G			446.5	1	446.5	446.6	56-58	Specific	69.1	Manually Identified, No MS/MS
3A	K.GGVVLKEDALPGQK.T	650.1		705.9	2	1410.8	1410.7	58-71	Specific	73.5	Common in Isoforms 1 and 2
3B	K.GGVVLKEDALPGQK.T	622.6		470.9	3	1410.8	1410.7	58-71	Specific	73.5	Common in Isoforms 1 and 2
4	K.TEFK.V			524.5	1	524.5	524.5	72-75	Specific	75.6	Manually Identified, No MS/MS
5	K.SSEHINEGETAMLVCK.S	710.9	C[+57],M[+16]	607.6	3	1820.8	1820.8	112-127	Specific	64.2	Common in all Isoforms
6	K.ITDSEDK.A			807.8	1	807.8	807.8	142-149	Specific	94.2	Manually Identified, No MS/MS
7	R.FFVSSSQGR.S			507.5	2	1015.1	1015.1	158-166	Specific	88.4	Manually Identified, No MS/MS
8	R.SELHIENLNMEADPGQYRC	485.8		705.9	3	2115.9	2115.9	167-184	Specific	99.5	Common in all Isoforms
9A	R.RKPEDVLDDDDAGSAPLK.S	693.6		970.9	2	1940.9	1940.9	233-250	Specific	72.3	Common in all Isoforms
9B	R.RKPEDVLDDDDAGSAPLK.S	514.2		485.9	4	1940.9	1940.9	233-250	Specific	72.3	Common in all Isoforms

Figure 17: Top: Position of the nine identified tryptic peptides identified in the polypeptide chain of human basigin isoform 2 (Byonic, underscored green; manual identification, green). The potential *N*-glycosylation sites (NDS/NGS/NGT) are bolded and underlined. Bottom: List of the observed basigin peptides.

No glycosylated peptides for human basigin were observed using both strict and relaxed Byonic search strategies. Upon manual investigation of the LC-MS/MS data of the glycopeptide-enriched sample validated to contain the human basigin isoform 2 peptides, only weak TIC traces and no oxonium ions (m/z 204.08, 366.14, 512.10, 528.12, 657.23) at the HCD-MS/MS level could be detected, which indicated absence or trace levels of any *N*-glycopeptides in the analysed samples.

3.3. Isolation and Characterisation of Basigin from HepG2 Cells

3.3.1. Morphology, viability and growth profile of HepG2

HepG2, human hepatocellular carcinoma cells, were cultured and studied as a biological source of liver cancer basigin and tumour-specific *N*-glycans. The characteristic epithelial morphology and exponential growth of HepG2 were monitored over 27 days, **Figure 18**. Even after the extended culture period, the HepG2 cells remained viable and were thus harvested, lysed and the protein complement extracted for downstream basigin and *N*-glycome isolation and characterisation.

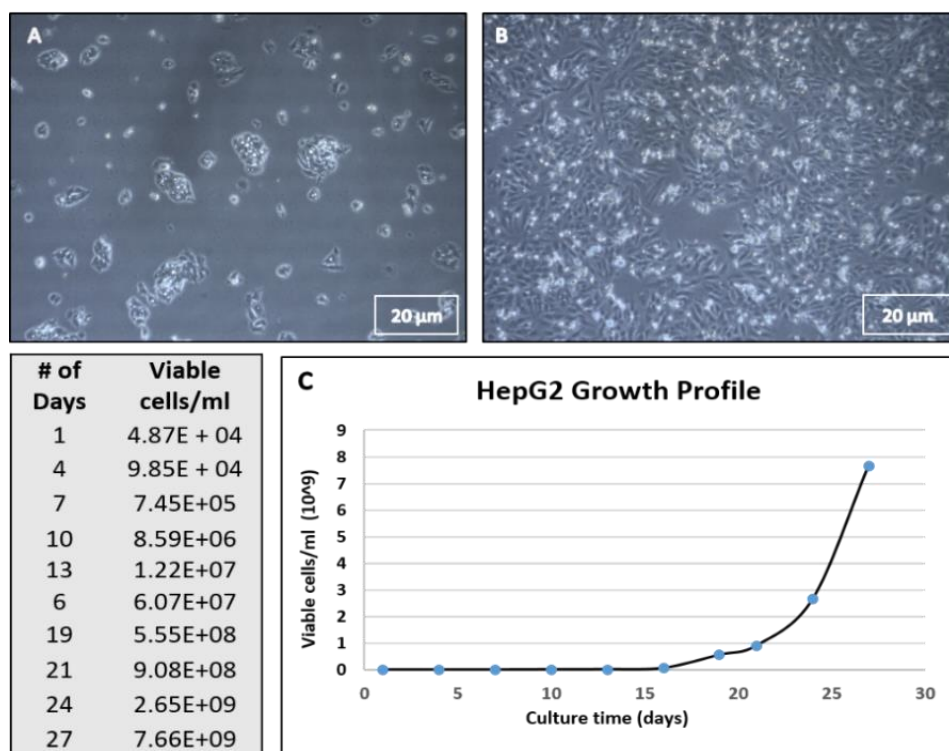


Figure 18: Morphology of HepG2 liver cancer cells at (A) low confluency and (B) high confluency as visualised by inverted light microscopy. Scale bars indicate 20 µm. (C) Exponential growth profile of viable HepG2 liver cancer cells cultured over 27 days.

3.3.2. Investigation of cell surface-resident basigin on HepG2 cells

The neutrophil studies indicated that plasma membrane enrichment strategies were useful to isolate basigin prior to characterisation. Thus, a biotin-based cell surface protein capture method

was applied to enrich the cell surface-resident basigin from HepG2 cultures. The streptavidin bound fractions containing the biotinylated cell surface proteins and the unbound fractions containing cytosolic proteins were separated and visualised on SDS-PAGE, **Figure 19**. Only a minor proportion of the HepG2 proteome was captured reflecting that the cell surface proteome only forms a subset of the total protein complement. Two protein bands (1 and 2) in the predicted basigin mass region (40-60 kDa) was identified as human type I cytoskeletal keratin (UniProtKB: P35527) on ion trap LC-MS/MS. No human basigin was detected by automated (Byonic) and manual interrogation of the LC-MS/MS data possibly reflecting that basigin is a very low abundance protein on HepG2 surfaces.

This biotin labelling and capture method is dependent on available primary amines (lysine residues) on the extracellular side of cell surface resident proteins. Since no 3D structures of basigin are available it cannot be assessed if the extracellular domains of basigin lack available primary amines that could serve to further explain its absence using this approach. Thus, an IP based approach was performed to enrich human basigin for molecular characterisation.

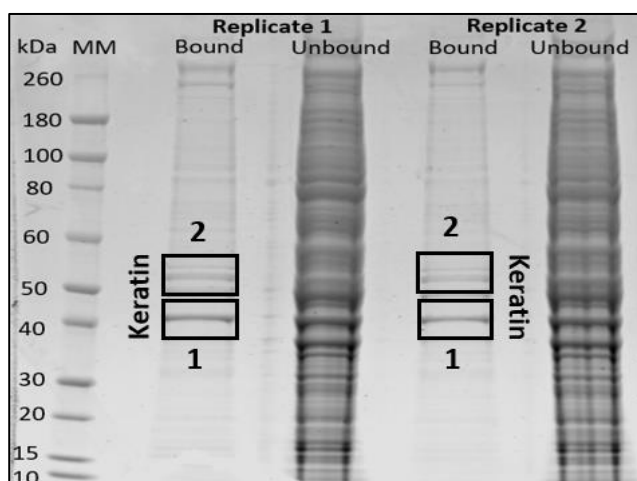


Figure 19: SDS-PAGE of cell-surface captured proteins from HepG2 cells. MM is the molecular marker. The streptavidin bound (cell surface) and unbound (non-cell surface) protein fractions are indicated. Protein bands of interest (boxed) were analysed further using LC-MS/MS.

3.3.3. Basigin IP of HepG2 lysates

Isolation of cancer basigin was attempted using IP of HepG2 cell lysates. Fractions were analysed by Coomassie stained SDS-PAGE and WB, **Figure 20**. Unlike the IP from neutrophil lysates, high HepG2 protein amounts were used as evaluated by the many intense gel bands in the PC and UB fractions. Few significant protein bands in the molecular mass region of 30-55 kDa appeared in the eluted (E) fraction. Following in-gel trypsin digestion, the corresponding proteins were identified using ion trap LC-MS/MS proteomics. Band 2 (42 kDa) was identified as human basigin (Mascot score 41) by the confident detection of the basigin peptide (R.SELHIENLNMEADPGQYR.C) using the Mascot search engine against the *Homo sapiens* proteome, **Table 4**. Since the identified peptide is shared amongst all basigin isoforms (see **Table 2**), the isoform of basigin was however not specified using Mascot.

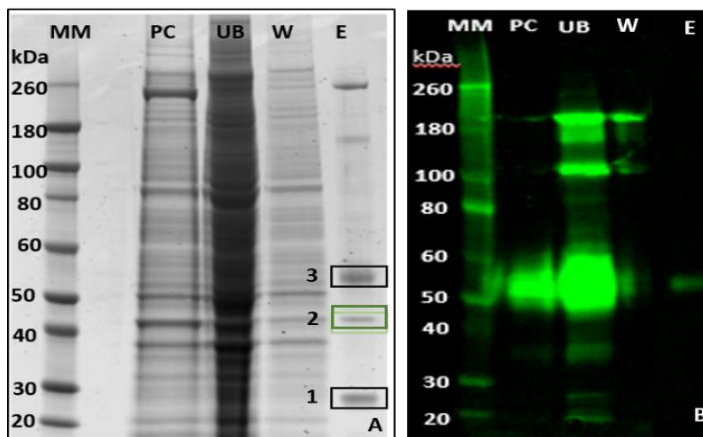


Figure 20: Visualisation of IP fractions of HepG2 lysates by matching SDS-PAGE gels after (A) Coomassie staining and (B) basigin-focused WB. MM is a molecular marker, PC is pre-clear, UB is unbound supernatant, W is wash and E is the eluted fraction. Protein bands of interested (boxed) were identified using LC-MS/MS.

The proteasome activator complex (Q9UL46) and the protein disulfide isomerase (P07237) were also identified in the IP'ed fraction (band 1 and 3) indicating possible interaction with basigin in HepG2 cells. However, as supported by the corresponding WB, the co-IP may arise from unspecific reactivity against the proteasome activator complex and/or the heavy chain of IgG (from the IP) with the secondary basigin antibodies as indicated by strong reactivity to a ~55 kDa band in the WB.

Band	UniProtKB Acc. #	Protein Name	Mascot Score	Sequence Coverage (%)	# Identified Non-Redundant Peptides	Observed MW (kDa)	Expected MW (kDa)
1	Q9UL46	Proteasome Activator Complex Subunit	55.3	8.37	6	27.0	27.4
2	P35613	Human Basigin	41.0	15.84	1	42.0	42.2
3	P07237	Protein Disulfide Isomerase	185.1	13.98	7	55.0	57.1

Table 4: Proteins identified in the IP fractions of HepG2 lysates using LC-MS/MS. Human basigin (green) was confidently identified.

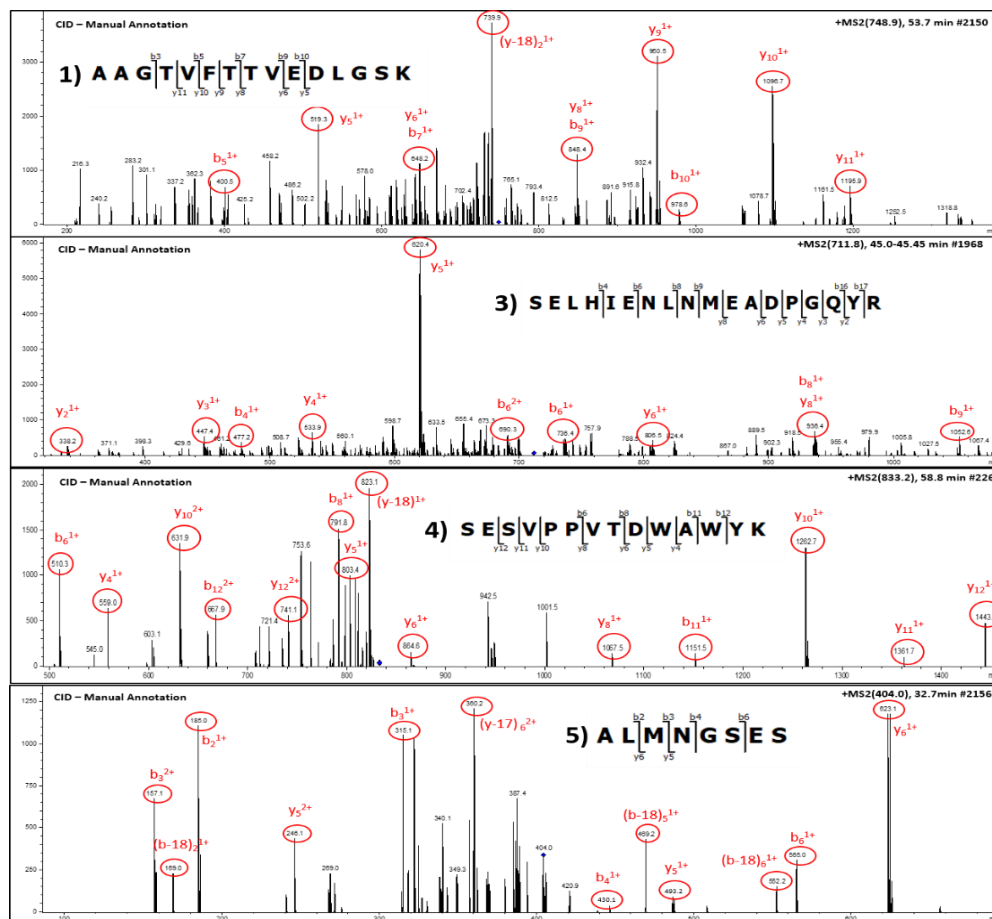
Due to the limited possibility for conducting customised searches using the Mascot search engine, this basigin-containing LC-MS/MS data from band 2 was re-interrogated using the more flexible Byonic search engine and by manually curation, e.g. only searching against the four basigin isoforms and including different variable modifications, **Table 5** and **Appendix Table S3**.

#	Peptide Sequence	Byonic Score	Modification Types	m/z	Charge (z)	Observed [M+H] ⁺ (Da)	Calculated [M+H] ⁺ (Da)	MS/MS	Start-End AA	Cleavage	Scan Time (min)	Comment
1A	G.AAGTVFTTVEDLGSK.I	425.6		748.3	2	1495.8	1495.8	Yes	22-36	NRagged	53.8	Isoform-2 Specific
1B	G.AAGTVFTTVEDLGSK.I	227.3		499.4	3	1495.3	1495.7	Yes*	22-36	NRagged	53.9	Isoform-2 Specific
2	K.GGVVLK.E			572.4	1	572.7	572.5		58-63	Specific	48.2	Manually Identified, No MS/MS
3	K.SESVPPVTDWAWYK.I	137.2		833.1	2	1665.3	1664.7	Yes	128-142	Specific	58.8	Common in all Isoforms
4	K.ALMNGSES.R	104.7		404.5	2	808.0	808.3	Yes	149-158	CRagged	32.9	Asn268-peptide
5	R.SELHIENLNMEADPGQYR.C	134.8	N[+1], M[+16]	712.1	3	2134.2	2132.9	Yes	167-184	Specific	45.3	Obtained in Mascot search also
6	K.GSDQAIITLR.V			536.1	2	1074.2	1074.2		185-194	Specific	57.6	Manually Identified, No MS/MS

Table 5: Identified peptides of cancer basigin isoform 2 from tailored Byonic searches and manual interrogation of LC-MS/MS data of the IP fraction (band 2) of HepG2 lysates. *MS/MS not shown.

This confirmed the presence of a cancer associated basigin isoform 2 (Byonic score 425.60) by not only recognising the same peptide sequence identified by Mascot, but also detecting additional peptide sequences exclusive to isoform 2 (see **Table 2**). Specifically, high quality MS/MS spectra for four peptides originating from this cancer source of basigin isoform 2 were observed, **Figure 20**.

Manual interrogation was performed by searching for the presence of theoretical tryptic peptides from basigin isoform 2. Only two additional peptides (K.GGVVLK.E and K.GSDQAII.TLR.V) could be observed but at low abundance without any MS/MS confirmation. The polypeptide sequence coverage of cancer basigin isoform 2 obtained from these search strategies was 28.60%, **Figure 21**.



ZIC-HILIC-based enrichment of the basigin peptide mixture (band 2) was performed to target and analyse the enriched glycosylated peptides of cancer basigin isoform 2 using ion trap LC-MS/MS. Unfortunately, no glycosylated peptides of human basigin were observed by automatic (Byonic) or manual interrogation of the data. A very weak TIC trace with no oxonium ions (m/z 204.08, 366.14, 512.10, 528.12, 657.23) on manual investigation at the MS/MS level indicated an absence of intact glycopeptides in the enriched sample. The non-enriched sample in contrast contained a few MS/MS that clearly corresponded to *N*-glycopeptides (detection of m/z 366.14 and 512.10), but these could not be matched by Byonic or manual assignment to basigin isoform 2 glycopeptides and were therefore concluded to arise from a contaminating glycoprotein of unknown identity.

3.4. *N*-Glycome Profiling of Neutrophils and HepG2

A separate aim of the thesis involved defining the cell- and subcellular-specific *N*-glycosylation of key cellular components of the liver tumour micro-environment. To this end, entire populations of *N*-glycans from various protein fractions of isolated neutrophils and HepG2 cells were enzymatically released using PNGase F, and profiled in their reduced native (non-derivatised) form using an optimised PGC-LC-ESI-CID-MS/MS method [73], **Figure 22**. Firstly, the *N*-glycome from lysates and isolated microsomes of HepG2 liver cancer cells were visually compared, **Figure 22 A-B**. In total, 15 *N*-glycans spanning mostly the high mannose type, but also covering the complex and paucimannosidic types, were confidently identified. The neutrophil *N*-glycome showed markedly different features by being rich in the unusual paucimannosidic *N*-glycans in particular the M2F *N*-glycan ($\text{Man}_2\text{GlcNAc}_2\text{Fuc}_1$), **Figure 22 C-D**. Although many *N*-glycans were common to both cell types and their subcellular compartments (see *N*-glycans in green in **Appendix Figure S5/Table S4**), several HepG2 and neutrophil-specific *N*-glycans were observed. EIC-based quantitative analyses of HepG2 glycans showed that the complex *N*-glycans were significantly higher and that the paucimannosidic/high mannose *N*-glycan types were lower in HepG2 microsomes compared to the lysate (all $p < 0.05$), **Figure 23 A**. Furthermore, the soluble (peripheral) and insoluble (membranous glycoproteins) part of the plasma membrane differed quantitatively in their *N*-glycosylation signatures with the latter being significantly higher in high mannose glycans and lower in paucimannose than the former ones (all $p < 0.05$), **Figure 23 A**. The complex biantennary sialylated glycan (m/z 1111.4²⁻) was intriguingly found exclusively in the soluble plasma membrane part, whereas a biantennary fucosylated glycan (m/z 1112.0²⁻) was identified exclusively in the membrane component of neutrophil plasma membrane.

Quantitative comparisons of the *N*-glycomics data revealed that neutrophils and HepG2 regardless of the observed subcellular differences differ dramatically in their *N*-glycosylation, **Figure 23 B**. Most

strikingly, the neutrophils were rich in paucimannose and had significant more hybrid type *N*-glycans whereas high mannose type features were the dominant signatures of the cancer cells (all $p < 0.01$). The presence of complex glycans proved to be statistically different for both cell types ($p \geq 0.05$).

Collectively, the presented *N*-glycomics experiments have demonstrated a pronounced cell- and subcellular-specific *N*-glycosylation of two key cellular components of the liver tumour micro-environment. These powerful structural methods for glycan mapping have provided insight not only of the exact glycans associated with the cancer- and immune-components of liver tumours, but also how they are spatially residing in the different compartments within these tumour cell types.

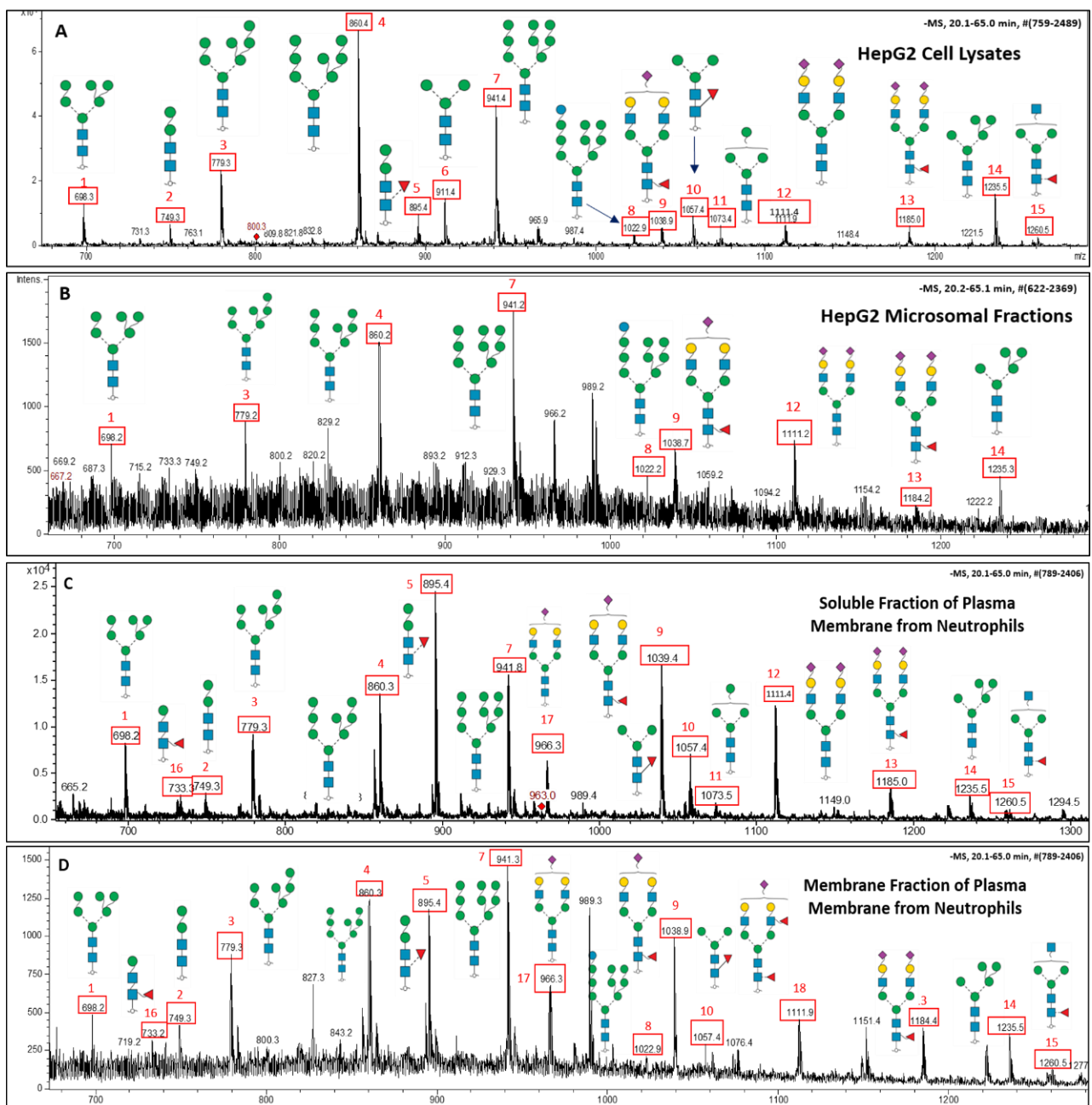


Figure 22: PGC-LC-ESI-MS profile of the *N*-glycome of HepG2 (**A**) lysates and (**B**) microsomes and of (**C**) the soluble and (**D**) membranous part of enriched plasma membranes of human neutrophils. The HepG2 *N*-

glycomes were profiled in three technical replicates. Four biological replicates were performed for the neutrophil *N*-glycomes. Representative base peak chromatograms (BPCs) from 20-65 min are provided to enable a visual comparison of the observed monosaccharide compositions (in red boxes). Proposed glycan structures are indicated. See **Appendix Figure S5/Table S4** for more information of the identified *N*-glycans.

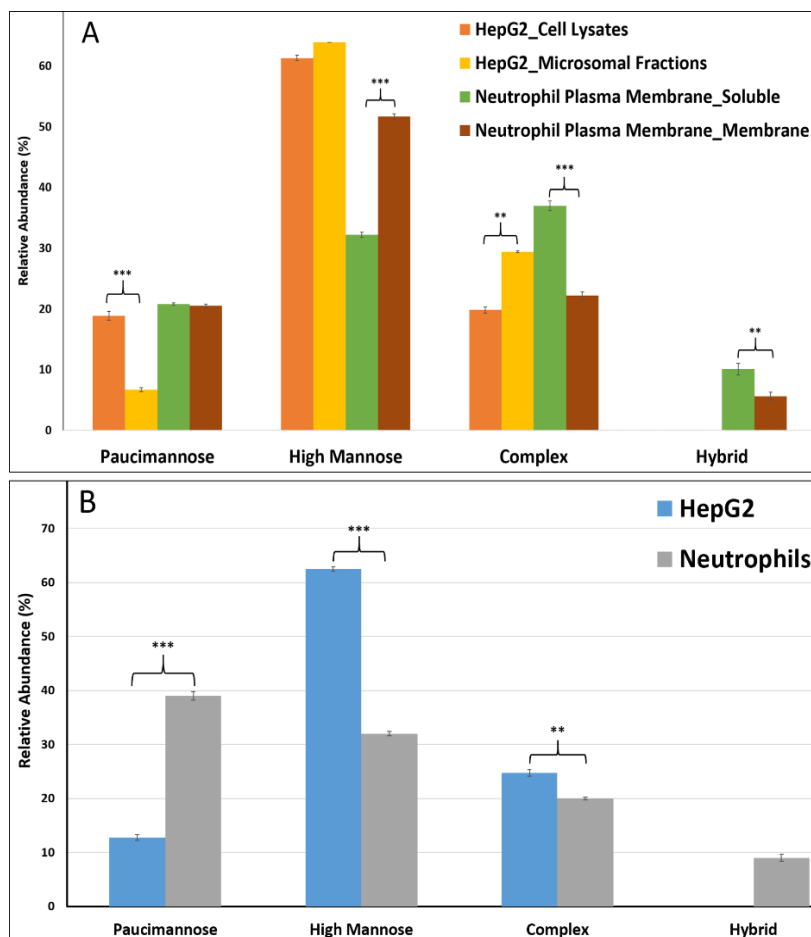


Figure 23: (A) Subcellular-specific *N*-glycosylation of the glycan types in lysates (orange) and microsomes (yellow) of HepG2 cells and in soluble (green) and membrane (brown) fractions of the plasma membrane of neutrophils. **(B)** Cell-specific *N*-glycosylation as evaluated by *N*-glycan type comparison of neutrophils (grey bars) and HepG2 liver cancer cells (blue bars). HepG2 cells were rich in high mannose and neutrophils in paucimannosidic glycans. For all graphs: $n = 3$, average \pm SD relative abundance of *N*-glycans is plotted, ($*p < 0.05$, $**p < 0.01$, $***p < 0.005$).

Chapter 4: Discussion

Literature has established that human basigin is expressed in many tissues and cell types as four distinct isoforms [41, 42, 126, 127]. However, remarkable little is still known of the detailed molecular features of this key glycoprotein that is recognised as an important driver of cancer [128]. This thesis aimed, in part, at filling this knowledge gap. Interestingly, the so-called retina-specific isoform 2 of human basigin was successfully identified and large parts of the polypeptide chain mapped from both neutrophil plasma membranes and lysates of HepG2 cells. *N*-glycome mapping at high structural resolution further served to provide novel insight into the cell- and sub-cellular specific *N*-glycosylation of these key immune and cancer components of the liver tumour micro-environment. The obtained results and the underlying methods used are discussed below in the context of the thesis aims and the existing literature.

4.1. Antibody-Based Enrichment and Detection of Basigin

Modern proteomics has demonstrated unrivalled potential to identify proteins in complex mixtures. However, conventional proteomics rarely provides deep structural information of proteins features beyond the confident identification of one or more peptides per polypeptide chain. In addition, particular proteome subsets e.g. proteins of low abundance, heavily modified (e.g. glycosylated) and hydrophobic (e.g. transmembrane) proteins can be extremely challenging to identify from unenriched biological protein mixtures even when using the most advanced and sensitive mass spectrometers [103]. In realising that basigin displays these “proteomics-unfriendly” characteristics, strategies to enrich basigin prior to downstream LC-MS/MS detection were explored.

One way to confirm IP'ed proteins is by WB, a widely used method in molecular biology for protein identification and to study protein function [101]. Hence, basigin-focused IP experiments confirmed by WB and LC-MS/MS were included in the experimental design, but limited success was achieved; only the IP of HepG2 lysates yielded clear evidence of basigin (discussed in **Section 4.2**). The IP of neutrophil lysates did not provide convincing identification of basigin, possibly because of low neutrophil protein starting material. The WBs yielded anti-basigin reactive bands, but lack of corresponding LC-MS/MS-based confirmation of basigin in the matching SDS-PAGE gel bands led to inconclusive results despite some of the IP'ed gel bands migrating in the expected basigin mass region (~40-60 kDa). Relatively weak identification of the abundant neutrophil-specific MPO and albumin in the unbound IP fraction by LC-MS/MS confirmed a working of IP protocol [124]. WB and modern LC-MS/MS based proteomics reportedly have the same sensitivity range for protein identification (low ng) [99, 100], but the sensitivity limit of the latter may be influenced dramatically if highly abundant proteins e.g. keratins are present in the same protein mixture due to masking and suppression effects during ionisation within the mass spectrometer.

Both WB and IP are highly dependent on the availability and appropriate application of high-affinity and specific antibodies to provide sensitive and accurate reactivity to the protein(s) of interest. The antibodies optimised for WB (detection of SDS-denatured antigens) may be designed differently for IP (typical detection of native epitopes of maturely folded proteins) and therefore careful antibody selection is necessary. Antibodies with those characteristics can be difficult to obtain for proteins in particular for basigin that is part of the IgSF containing many protein members with similar folds (and hence similar epitopes/antigens). Unspecific IgG reactivity naturally represent a potential issue since auto-reactivity of the antibodies may occur or enriched IgG, which is often abundant in biological mixtures, may react unintendedly with secondary antibodies in WB.

In this thesis, mouse monoclonal anti-human basigin antibody was selected due to reported suitability for both WB/IP [129]. Reactivity to rh basigin was ensured using dot-blotting, a fast and simple method that bypasses the gel electrophoresis step, but therefore also provides no size information of the targeted protein(s) [130]. Another primary rabbit polyclonal anti-CD147 (ab64616) antibody was tested in parallel for basigin antigen reactivity, but no rh basigin-reactivity was observed, which may reflect that this antibody was raised against a different basigin isoform. The rh basigin was confirmed to be the isoform 2 upon contacting the manufacturer. Thus, this proved to be a good “model antigen” to test anti-human basigin mAB reactivity since isoform 2 was also the identified basigin variant of neutrophils and HepG2. This thesis has demonstrated utility of the mouse monoclonal anti-basigin antibody to isolate human basigin from a complex mixture. However, a wider panel of antibodies, and further validation of their reactivity/specificity as well as optimisation of parameters like ionic strength, pH of lysis buffer, and type and concentrations of detergents used, factors which can dramatically influence the outcome of the IP experiment, are clearly required to enable more insightful IP and WB experiments of basigin in the future.

4.2 Basigin-Focused Cellular and Subcellular Isolation Strategies

Instead of targeting basigin directly after lysis of entire cell populations, strategies involving isolation of subcellular components such as the plasma membrane proteome potentially containing this low abundance membrane glycoprotein were explored. The rationale was that by significantly enriching basigin within the analysed protein mixture (increasing the stoichiometry of basigin molecules relative to other proteins), a deeper LC-MS/MS based characterisation would be facilitated.

This thesis first demonstrated an efficient isolation of primary human neutrophils using common blood cell isolation strategies. Neutrophils are abundant granular leukocytes in circulation and were investigated here since they are recognised as being key immune components of the liver tumour micro-environment [87]. Very high purity, viability and yields of resting (non-activated) neutrophils were achieved using two slightly different density gradient-based methods i.e. Polymorphprep and Lymphoprep [20]. The characteristic multi-lobed nuclei morphology indicative of fully mature neutrophils was observed by microscopy after Giemsa-Wright staining. This validation was important since immature neutrophils can appear in circulation under certain conditions e.g. during infection and inflammation [3]. Polymorphprep isolation is a one-step process in which MNCs and PMNs are directly separated by density. In contrast, the Lymphoprep isolation procedure is a two-step process whereby RBCs are first eliminated by dextran sedimentation followed by separation of MNCs from PMNs by density. Although Polymorphprep isolations are faster (~3 h), Lymphoprep

isolations (~5 h) evidently provided higher (~two times) yield while maintaining very high purity and viability of neutrophils, and were thus used prior to the downstream subcellular fractionation.

The basigin-rich plasma membrane was separated from granule components of neutrophils by differential density in Percoll gradients after gentle cell disruption using nitrogen cavitation [86, 87]. The high-efficient disruption critically maintained full integrity of the granules to avoid contamination of the temporally disrupted plasma membrane, which resealed into low-density vesicles that were harvested on top of the density gradient [86, 131]. Percoll is a colloidal silica-based PVP-coated, non-toxic, chemically inert and low-viscosity media for density separation of cells and sub-cellular particles with 95% recovery of plasma membrane components of neutrophils [132].

LAP activity was used to identify the low-density plasma membrane in fractions 27-35. Alkaline phosphatase is present on both the plasma membrane (30%) and within secretory vesicles (70%) of unperturbed neutrophils [86, 87]. Since plasma membrane fragments reseal “right-side out” after the disruption by nitrogen cavitation, the LAP enzyme activity is present outside of plasma membrane vesicles and inside of the secretory vesicles. [86, 88, 133]. The downstream (glyco)-proteomics applications were crucially dependent on the complete elimination of LC-MS/MS incompatible Percoll from the samples to avoid masking the presence of low-abundance proteins. In this two-step procedure, the bulk of the Percoll was first removed by ultra-centrifugation and then by short introduction of proteins into SDS-PAGE gels, washing and *en bloc* in-gel digestion. Despite these multiple handling steps, 60 µg total protein from the isolated neutrophil plasma membrane, sufficient for a detailed characterisation of neutrophilic basigin, could be harvested.

In this thesis, HepG2 was used to represent the liver cancer cells of the tumour micro-environment. HepG2 is a well-studied immortalised human liver carcinoma (epithelial) cell line from well-differentiated hepatocellular carcinoma liver tissue. Although these cells are non-tumourigenic in nature, they have a high growth potential and secrete many plasma proteins e.g. plasminogen, albumin, transferrin, and fibrinogen making them a good representative for HCC [11, 134, 135]. Several conventional cell-based methods were used to grow and monitor the HepG2 to ensure high viability, correct morphology and absence of microbial contamination. Several rounds of passaging led to the harvest of a total of 9.2×10^9 viable cells over an extended culture period of four weeks.

To promote basigin enrichment and characterisation, the cell surface proteome of HepG2 was targeted using selective biotinylation of proteins residing on the surfaces of HepG2 cells [105]. Cell-surface protein capture is ideal for adherent cells, and hence was only applied to HepG2 (not

neutrophils). Unlike other capture strategies, the captured cell surface proteins could be efficiently released with a 98% reported recovery rate using this protocol [136] due to the use of a thiol-cleavable amine-reactive trifunctional linker (sulfo-NHS-SS-biotin, mass 607 Da) [137]. The labelling and capture of cell surface proteins was clearly successful as evaluated by the isolation of a specific and much smaller subset of proteins. However, as confirmed by repeating the experiment twice, only keratin and no basigin was identified using ion trap LC-MS/MS proteomics. The absence of basigin may be explained by 1) its low abundance, 2) there are a lack of available primary amines or steric hindrances of amines on the extracellular domains of basigin hindering or lowering the efficiency of biotinylation and/or 3) high confluency of HepG2 cells, which lowers the labelling and capture efficiency [136]. Considering the expense of labelling reagents and the significant number of cells needed for each experiment, the cell surface capture approach was not explored further.

4.3. Molecular Characterisation of Neutrophil and Liver Cancer Basigin

The Byonic search engine confidently identified 795 different proteins (peptide FDR <1%) from a proteomics analysis of the tryptic peptide mixtures of enriched neutrophil plasma membrane proteins using a high resolution/high mass accuracy Orbitrap Fusion Tribrid mass spectrometer (**Table 3**). Typical plasma membrane-resident proteins i.e. actin, haptoglobin, serotransferrin, and myosin-9 were confidently identified providing support for a highly efficient enrichment of plasma

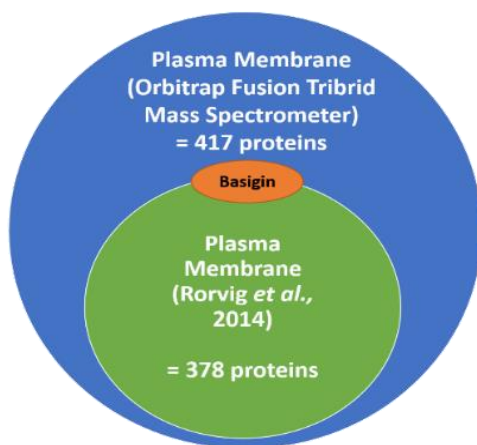


Figure 24: Neutrophil proteins from the plasma membrane identified in this thesis including basigin was compared with a previous study [9].

membrane proteins. On performing a gene ontology enrichment analysis, 86.1% of all proteins classified as being plasma membrane resident [138]. A landmark paper by Rorvig *et al.*, 2014 published a list of 378 identified plasma membrane proteins using LTQ Orbitrap XL MS/MS. The plasma membrane proteome analysis on the Orbitrap not only identified all the previously reported plasma membrane proteins, but identified an additional 417 proteins indicating the importance of utilising a state-of-the-art mass spectrometer to increase the proteome coverage. Some previously unknown plasma membrane proteins included apolipoprotein, fibrinogen, amyloid,

quinone oxidoreductase, voltage-gated potassium channel proteins, and ceramidase. Importantly, human basigin-1 (isoform 2), ranked as 275th protein, was confidently identified by the automated (Byonic-based) detection of five non-redundant peptides belonging to basigin. Although not a highly accurate measure for relative protein abundance, this intermediate ranking indicated that basigin

is amongst the low-to-medium abundance proteins within the plasma membrane sub-proteome. The detection of the isoform 2 specific *N*-terminal peptide AAGFVQAPLSQQR.W not only confirmed the isoform variant of human basigin, but also validated that neutrophilic basigin appears with an intact (non-truncated) *N*-terminus in a fully mature form without the preceding signal peptide. Upon manual investigation, four additional tryptic peptides of basigin were observed that further increased the overall polypeptide sequence coverage to 41.5%.

No PTMs other than a methionine oxidation (potentially introduced by sample handling) were detected on the mapped region of neutrophil basigin. Inspecting the literature for other potential types of PTMs, it was observed that only Ser362 and Ser368 human isoform 1 may appear with phosphorylation [139, 140]. Although theoretically giving rise to four LC-MS/MS compatible tryptic peptides (see **Figure 17**), none of the observed non-modified peptides covered any of the three potential *N*-glycosylation sites of human basigin. This indicated a high degree of occupancy of *N*-glycans on the three corresponding sequon-located asparagine residues of basigin. Since the heterogeneous *N*-glycosylated peptides are typically only weakly represented in LC-MS/MS proteomics data due to sub-stoichiometry and less ionisation strength [72], a glycopeptide enrichment strategy was performed to enable structural characterisation of protein *N*-glycosylation in glycoproteomics [114, 122, 141].

Glycopeptides of the neutrophil plasma membrane peptide mixture were enriched using an optimised ZIC-HILIC SPE protocol [114]. In short, ZIC-HILIC, a variant of normal phase liquid chromatography, consists of a sulfoalkylbetaine derivatised stationary phase containing both strongly acidic sulfonic groups and basic quaternary ammonium groups separated by short alkyl spacers. This immobilised zwitterion adsorbs water by hydrogen bonding, which facilitates selective partitioning of the hydrophilic glycopeptides from the organic (more hydrophobic) mobile phase. Critically, TFA acts as an ion-pairing reagent in the mobile phase, which increases the hydrophilicity difference between glycopeptides (retained) and non-glycosylated peptides (non-retained) [114]. This enrichment method favours *N*-glycopeptides over *O*-glycopeptides due to the larger and more hydrophilic conjugated glycans; it was shown that non-biased *N*-glycopeptide enrichment can be achieved using ZIC-HILIC SPE [114, 141]. Since HILIC does not involve any irreversible chemical or enzymatic alterations of the peptide or glycan moieties it is widely recognised as the preferred (but not the only) enrichment method to facilitate analysis of native glycopeptides [115, 122, 141].

The high resolution of Orbitrap in both the MS and MS/MS levels and the alternating HCD, EThcD and CID fragmentation modes to yield orthogonal structural information of different parts of the

intact glycopeptides are crucial to enable efficient and confident glycopeptide identification directly from complex mixtures [125]. Manual interrogation of the raw data showed the presence of some HCD-MS/MS spectra containing abundant oxonium ions (m/z 204.08, 366.14, 512.10, 528.12, 657.23) that are useful diagnostic ions for identifying *N*-glycopeptide containing fragment spectra. Oxonium ions can even provide information of the monosaccharide composition and structure of the *N*-glycopeptide [115, 122, 125]. Using tailored parameters in the glycopeptide-centric search engine Byonic and by using a variety of strict (narrow proteome and glycome search space) and more relaxed search criteria specific *N*-glycopeptides such as K.VVLHPNYSQVDIGLIK.L (Hex₄HexNAc₂), R.TNLTDR.Q (Hex₃HexNAc₂) and R.QQQHLFGSNVTDCSGNFCLFR.S (Hex₃HexNAc₂Fuc₁) from the *N*-glycoproteins haptoglobin, MMP-9 and serotransferrin, respectively, were confidently identified from the neutrophil plasma membrane fraction (**Table 3**). However, the plasma membrane fractions mostly contained non-glycosylated membrane proteins like actin and myosin. Some *N*-glycosylated proteins were observed without the conjugated *N*-glycans. The identified *N*-glycans of the actually glycosylated plasma membrane glycoproteins spanned all glycan types including the complex and high mannose types, but were mainly of the paucimannosidic type. The observation of paucimannose-rich plasma membranes of neutrophils was later confirmed by glycomics of the neutrophil plasma membrane; paucimannose was found to be the most abundant *N*-glycan type on the cell surface (~39%). The agreement between the glycopeptide analysis and the glycomics also indicated that the glycopeptide enrichment was unbiased for the entire *N*-glycopeptide population. No *N*-glycopeptides belonging any isoform of human basigin were detected by Byonic possibly reflecting too low abundance of these glycopeptides even after multiple rounds of glycoprotein and glycopeptide enrichment. As a last report, the *N*-glycoproteomics LC-MS/MS data was manually interrogated by searching the HCD fragment spectra for the presence of the characteristic Y₁ ions (peptide + GlcNAc fragment) from the three theoretical basigin tryptic glycopeptides. This indirect approach, which builds on systematic studies of the fragmentation pattern of *N*-glycopeptides [142] and relies on the high MS/MS mass accuracy of the Orbitrap, did, however, not yield evidence of neutrophil basigin *N*-glycosylation.

Basigin was also identified from IP'ed protein extracts from HepG2 lysates using ion trap-based LC-MS/MS proteomics. However, the isoform of human basigin could not be specified from the isoform-unspecific R. SELHIENLNMEADPGQYR.C peptide (Met oxidation probably due to sample handling) identified by Mascot. However, the more flexible Byonic search engine, which is able to detect semi- and non-tryptic peptides, identified an additional semi-tryptic (*N*-terminal) peptide AAGFVQAPLSQQR.W specific to isoform 2 of basigin increasing the polypeptide sequence coverage

to 28.60%. These two peptides were also identified for neutrophil basigin, however, the remainder of the peptides covered different regions of the polypeptide chain. Although the peptide K. GGVVLKEDALPGQK.T was a missed tryptic cleavage observed for neutrophil basigin, the fully cleaved peptide K. GGVVLK.E was observed for HepG2 basigin. Interestingly, for HepG2 basigin, K. ALMNGSES.R (C-ragged) containing the Asn268 glycosylation site was identified, but without any conjugated glycans. Neither other PTMs nor the C-terminal peptide were detected for neutrophil and HepG2 basigin. Upon manual data interrogation, a few CID-MS/MS spectra were observed to contain oxonium ions e.g. m/z 366.14 and 512.10 indicating the presence of intact glycopeptides. However, even after thorough spectral annotation, these glycopeptides could not be assigned to the isoform 2 of human basigin (or any other glycoprotein in the sample). LC-MS/MS analysis after glycopeptide enrichment also did not uncover the site-specific glycosylation of HepG2 basigin.

Due to low abundance and the inherent analytical challenges associated with glycopeptide analysis, in particular from hydrophobic membrane glycoproteins, no convincing site-specific data of basigin *N*-glycosylation was obtained despite the application of multiple orthogonal approaches. Previously, researchers from our group quantitatively mapped the *N*-glycan types of isolated HEK-derived rh basigin (isoform 2) by *N*-glycome profiling; they found that ~97% of the *N*-glycans belonged to the complex/hybrid type. These glycans were highly α 1,6-linked core fucosylated. Few glycans terminated with NeuAc-type sialylation whereas the majority of structures terminated with the β 1,2/6- (mono-, bi- and tri-antennary) or β 1,4-(bisecting) GlcNAc typical for recombinant HEK glycoproteins [143, 144]. The site occupancies of Asn160 and Asn268 were ~88% and ~21%, respectively, and differences in the glycan micro-heterogeneity were observed between the sites [74, 93]. The challenges of site-specific mapping the *N*-glycosylation of basigin when the recombinant protein was added to a complex protein mixture in a ratio resembling biological samples studied in this thesis were also demonstrated [74]. Given the relative high sequence coverage of neutrophil and HepG2 sequence coverage and the LC-MS-friendly tryptic glycopeptides covering the Asn160 (K.ILLTCSLNDSATEVTGHR.W) and Asn268 (K.ALMNGSESR.F) sites, it was rather surprising that no glycopeptides could be detected with or without prior HILIC enrichment. Potential reasons that may explain their absence in the analyses include unusual glycosylation signatures on basigin, for example, 1) extremely branched (e.g. tetra/penta-antennary) and high negatively charged (e.g. poly-sialic acid) glycoforms that are very difficult to detect in the mass spectrometer; such structures are more likely occur on HepG2 basigin [145], 2) extremely small (e.g. paucimannose) and even further truncated (chitobiose type) glycans [146] that may not be sufficiently hydrophilic to be retained by ZIC-HILIC; such structures are more likely to be relevant for

neutrophil basigin [16, 17, 147], and finally 3) an extreme degree of micro-heterogeneity that would significantly lower the stoichiometry of the individual glycopeptides and thus be more difficult to detect.

For better understanding the (glyco)peptides of liver cancer-associated basigin (and to identify the isoforms), the proteome data repositories and databases were searched such as PRIDE and Peptide Atlas to explore if already acquired/deposited/published data contained evidence of basigin *N*-glycopeptides. Relevant key words were used to narrow down data of interest including “HCC” and “basigin” and relevant parameters were specified including high resolution mass spectrometers such as Orbitrap QE or Fusion and specific glycopeptides. Although human basigin was identified in many deposited HCC (and breast cancer) data sets, neither the exact isoform of basigin could be specified nor were any glycosylated peptides observed from the explored LC-MS/MS data.

Convincing empirical evidence has shown that the canonical isoform 1 of human basigin is associated with HCC progression and invasion by its stimulation of the expression and secretion of MMPs and growth factor of stromal cells surrounding cancer cells [126, 148, 149]. Basigin isoform 3 and 4 are expressed widely in normal (non-cancerous) human tissues, but are also upregulated in HCC tissues [146]. However, the levels of these isoforms were not assessed at the protein level. Basigin isoform 3 was investigated using nuclear magnetic resonance and fluorescence complementation assays revealing that it is capable of inhibiting HCC proliferation and invasion through interaction with isoform 1 as an endogenous inhibitor via hetero-oligomerisation [126, 148]. In fact, oligomerisation (homo- or heterodimeric) was an aspect not studied in this thesis due to the used bottom-up LC-MS/MS (glyco)proteomics. Top down mass spectrometry and native PAGE are methods that could be applied to investigate the monomer/dimer status of basigin in the future.

In conclusion, isoform 2 (basigin-1), expressed in neutrophils and HepG2 cells, is not a commonly reported variant of basigin particularly not in a tumour micro-environment context. Basigin isoform 2 has a similar polypeptide sequence to isoform 1 containing three and four potential *N*-glycosylation sites, but isoform 2 is a splice variant of the canonical form that lacks a large part of the polypeptide chain i.e. amino acid 24-139. Isoform 2 is reportedly expressed mainly by photoreceptor cells of the retina (hence “retina-specific”) and absence of this isoform in eyes can be an underlying cause of blindness [42, 43, 57] and is known to weakly interact with other proteins similar to isoform 1 for example by acting as a chaperone for the MCTs [42, 43, 57]. Although many molecular aspects including the site-specific glycosylation of basigin still needs to be defined, the

presented technology and biological findings enable further structural and functional characterisation of basigin in the future.

4.4. N-Glycome Signatures of Key Cellular Components of Liver Tumours

The *N*-glycomic profiling of sub-cellular fractions from neutrophils and liver cancer cells were done to investigate the differences in *N*-glycosylation between tumour and immune components of liver tumour as well as define, for the first time, the subcellular-specific glycan signatures in these cells. Altered glycans on the tumour and immune cell surfaces in the tumour micro-environment play critical roles in cancer pathogenesis and progression [14]. Furthermore, an increasing number of studies indicate the relevance of protein glycosylation in the modulation of the innate immune system in processes related to cell homeostasis and inflammation, processes known to be associated with tumourigenesis and metastasis [14, 150]. Thus, a broader understanding of the *N*-glycome signatures is necessary for generating a more complete picture of the glycosylation patterns and their spatial arrangement and regulation in tumour and immune cells.

Significant cell-specific differences in the glycan type distribution were observed; neutrophils were primarily presenting paucimannosidic glycans whereas HepG2 cells were abundantly expressing high mannose glycans. Literature hosts recent but highly convincing evidence supporting the existence of the truncated paucimannosidic glycans on neutrophil glycoproteins e.g. elastase and cathepsin G [16-18, 151]. Thus, it was not too surprising to detect dominant signatures of these unusual short *N*-glycans from isolated human neutrophils. Similarly, the literature contains support of an abundant expression high mannosidic glycans on HCC cell surfaces. In these studies, the entire high mannose series (4-11 mannosyl residues) was observed and reported to be correlating with cancer progression [14, 150, 152]. Thus, a crude (and admittedly an overly simplified) distinction can be suggested with respect to the *N*-glycosylation of liver tumours; the immune cells contribute paucimannosidic glycans and the tumour component brings high mannosidic glycans to this heterogeneous cellular environment. As discussed in a recent review, these two glycan type may carry out different functions within tumours by reacting with different mannose-based receptors e.g. C-type lectins including MBL and DC-SIGN [153]. Other lectin receptors e.g. galectin-3 or E-selectins reacting with different glycan epitopes have also been associated with tumour processes indicating that manno-glycans are not the only important glyco-features in tumours [154-157].

Furthermore, significant subcellular differences in the *N*-glycosylation were observed; the insoluble (integral glycoproteins) component of the neutrophil plasma membrane showed a significantly

higher abundance of high mannose compared to the soluble component (peripheral glycoproteins). Subcellular-specific (granule-specific) *N*-glycan differences have been suggested before in the literature [16], but this is the first report indicating glycans differences within the plasma membrane environment. Although both integral and peripheral glycoproteins use the same ER-Golgi-based biosynthetic pathway, their glycosylation differences may be explained by minor differences in how they are processed within the glycosylation machinery [4, 59]. Due to less accessibility by steric hindrance from the lipid bilayer and interactions in lipid rafts etc, the integral membrane proteins are predictably less prone to extensive processing by the glycosylation enzymes (glycosidases and glycosyltransferases) than the soluble (and more accessible) peripheral glycoproteins during biosynthesis [93, 158]. This differential glycosylation site accessibility could serve to explain the higher proportion of complex/hybrid type (more processed) glycans of peripheral glycoproteins and the higher proportion of high mannose (unprocessed) glycans of integral glycoproteins.

The HepG2 microsomes, vesicle-like re-formed pieces of predominantly ER membranes [133, 159], were as expected rich in the immature high mannose type glycans due to the sampling from this early part of the glycosylation machinery. The HepG2 lysate containing all membrane and soluble glycoproteins alike were also rich in high mannose, but had a significantly higher proportion of paucimannosidic glycans. This supports findings in other papers indicating that paucimannose is a glycan type that arise from significant processing [16, 153]. The HepG2 microsomes were also rich in complex sialyl- and fucosyl-glycans. Complex sialylated structures in the HepG2 glycome are known to contribute to tumour progression through their *cis*- and *trans*-interactions with glycans and receptors on the surfaces of their own or other cells, respectively [160, 161].

Decoding the glycosylation structures and regulation in liver tumours is not only important to understand cancer-associated functions of glycoproteins, but may also be important as a biomarker to detect and monitor cancer initiation, progression and recurrence. For example, glycome profiling on different cancerous cell types including in breast cancer, colorectal and ovarian cancer tumours identified some potential glycan cancer markers [162]. In addition, *N*-glycome analysis of serum and liver tissues provided a list of potential prognostic glycan markers for HCC [163]. These studies were not performed on the individual cell types of the tumour micro-environment. In fact, it has not been possible to find any studies that have attempted to unravel the cell- and subcellular-specific *N*-glycosylation of the individual components of the tumour micro-environment.

Despite the fact that this thesis has only addressed the *N*-glycosylation of two components of the extremely diverse tumour micro-environment and has done this in a rather simplistic reductionistic fashion by investigating the cell types separately in an *ex vivo* and *in vitro* fashion, the findings presented here are novel and will contribute to further, more advanced, studies encompassing all the cell types from real liver tumour tissues (discussed further below). Such holistic and more biological relevant experiments are required to ultimately generate a more the complete understanding of the involvement of *N*-glycosylation in the tumour micro-environment.

Chapter 5: Conclusion and Future Directions

5.1. Conclusions

The aim of this research study was, in part, to isolate and characterise the molecular features of human basigin from primary neutrophils and HepG2 cells representing key immune and cancer cells forming parts of the heterogeneous liver tumour micro-environment and, in part, to profile their cell- and subcellular-specific *N*-glycosylation signatures using glycomics. Little is known of the structure of this key cancer and immune glycoprotein despite convincing evidence supporting the involvement of basigin and the *N*-glycosylation in a plethora of functions in liver tumourigenesis.

Similar to other cell surface glycoproteins including the members of the CD family, basigin (CD147) is a challenging low-abundance transmembrane protein; a variety of cellular approaches were applied during my candidature to directly enrich basigin or isolate basigin-rich cell-surface fractions from neutrophils and liver cancer cells with an emphasis on using powerful LC-MS/MS-based proteomics and glycoproteomics. HEK-derived rh basigin (reportedly isoform 2) was used to optimise the isolation and characterisation techniques, however, other cell-dependent factors e.g. dimerisation status and PTMs including glycosylation potentially made this model glycoprotein an inaccurate representation of the naturally occurring basigin. This shortcoming hampered especially the antibody-based isolation (IP) and detection (WB) strategies, which were only partially successful. Access to native forms of human basigin from commercial sources may have enabled more tailored methods to be developed and validated. However, since glycosylation unfortunately remains somewhat overlooked in cell biology, native glycoprotein products are unfortunately rarely available. Cell surface-centric subcellular fractionation with downstream mass spectrometry-based detection proved to be a fruitful technology match up to facilitate characterisation of neutrophilic basigin. In fact, subcellular-specific proteomics using powerful modern mass spectrometers is gaining popularity these years due to the additional spatial insight of the proteome [86, 131].

This thesis has for the first time documented that the less common human basigin-1 (isoform 2, retina-specific) is expressed in a fully mature form (without signal peptide) on the cell surface of human neutrophils and HepG2 liver cancer cells. Large parts of the basigin isoform 2 polypeptide chain of both cell sources (~35-50%) were mapped using high confidence LC-MS/MS data demonstrating an absence of common PTMs such as phosphorylation, deamidation and methylation. Documenting this “retina-specific” basigin isoform 2 of immune and cancer components in the tumour micro-environment is novel and rather surprising since the common (canonical) isoform 1 has been associated more closely with liver cancer [164, 165]. Future work will confirm if these *ex vivo* (neutrophil) and *in vitro* (HepG2) observations of basigin isoform 2 are indeed reflected in liver tumour tissues obtained from HCC patients. Although no direct site-specific evidence of *N*-glycosylation of basigin was obtained using multiple tailored LC-MS/MS strategies, the absence of any non-glycosylated tryptic peptides covering the three potential *N*-glycosylation sites and the significant discrepancy between the observed (~42 kDa) and the predicted (29 kDa) molecular mass of basigin isoform 2 strongly indicated heavy *N*-glycosylation. Intact glycopeptide mapping from complex mixtures remains extremely challenging in particularly for low abundance membrane glycoproteins. As glycoproteomics is rapidly maturing [72, 122], it is expected that basigin *N*-glycopeptides will be profiled in the near future. The methods developed in this thesis assist in that important task.

Detailed cell- and subcellular *N*-glycome profiling of neutrophils and HepG2 using glycomics showed that not only did the glycans differ significantly between the two cell types, distinct *N*-glycosylation signatures were observed between isolated components of the cells. Recapitulating findings from similar glycome-centric studies performed in other biological systems [161], the primary immune cells studied here were rich in the unusual paucimannose whereas high mannose *N*-glycosylation was a dominating feature of the epithelial cancer cells. High mannose and paucimannose glycans are likely to have different receptors and hence may display different signalling functions [153]. Interestingly, the integral and peripheral membrane glycoproteins of the neutrophil plasma membrane demonstrated significant differences in the glycan type distribution and even displayed unique *N*-glycan structures. Thus, spatial *N*-glycome mapping of neutrophils and HepG2 generated novel insight into the variety of *N*-glycans forming the liver tumour micro-environment. Future experiments are required to confirm that activated neutrophils and liver cancer cells extracted from HCC tumours are displaying these glycan features. However, the presented observations provide an important platform to hypothesise on the functions and the underpinning mechanisms of the cell- and subcellular-specific *N*-glycosylation that are known features of liver tumours.

5.2. Future Directions

In reaching the three outlined aims, this research study has created an important technology platform and a knowledge base that can be used to further advance our understanding of the intriguing structures and functions of the *N*-glycosylated basigin in the future.

It should be stressed that the presented experiments, as discussed above, have several limitations and that many structural and biological questions remain unanswered. A rather simplistic model of the liver tumour micro-environment was chosen due to strict time constraints of this project. Future studies with a longer time horizon could benefit from investigating the individual cells of the actual liver tumour micro-environment by obtaining biological tumour tissues from HCC patients to accurately explore the structure and function of basigin *N*-glycosylation. Experimental designs using biological relevant specimen would also allow isolation and analysis of basigin from other cell types forming other critical parts of the liver tumour micro-environment including haemopoietic cells such as B- and T- cells, cells of mesenchymal origin such as fibroblasts and adipocytes, tumour stem cells and even non-cellular compartments e.g. extracellular matrix in addition to neutrophils and epithelial cells. Holistic approaches having attention to the entire suite of cell types in the heterogeneous liver tumours are challenging to carry out even with present day 'omics technologies, but offer more complete biological insight and is consequently well worth pursuing when aiming for delineating the cellular and molecular processes associated with liver cancer.

Similar to the basigin-focused studies, the cell- and subcellular-specific *N*-glycome profiling would benefit from being carried out using cells derived from HCC tumours. This is particularly pertinent since *N*-glycosylation is dynamic and highly context dependent. Inclusion of the additional cell types of the liver tumour micro-environment and by isolating the entire suite of cellular compartments prior to glycomics would provide higher spatial resolution and a more accurate picture of the *N*-glycome landscape in tumour tissues. Finally, including tissues of larger HCC patient cohorts of different stages and non-cancerous damaged liver tissues from liver transplantations (e.g. cirrhotic or alcohol livers), would provide valuable insight into inter-individual variation and the regulation of basigin and the cell- and subcellular-specific *N*-glycosylation in a disease development context.

Exploring the above follow-up experiments and aiming for dissecting central aspects of the glycobiology associated with liver cancer are ambitious, but highly relevant, goals to pursue in a PhD. Addressing such aims is critically important to advance our understanding of the complex molecular and cellular processes associated with the tumour micro-environment and the exact involvement of basigin and the *N*-glycosylation in liver damage and cancer.

References

1. Beasley, R.P., *Hepatitis B virus. The major etiology of hepatocellular carcinoma*. Cancer, 1988. **61**(10): p. 1942-1956.
2. Grivennikov, S.I., F.R. Greten, and M. Karin, *Immunity, inflammation, and cancer*. Cell, 2010. **140**(6): p. 883-99.
3. Malech, H.L., F.R. DeLeo, and M.T. Quinn, *The role of neutrophils in the immune system: an overview*. Neutrophil Methods and Protocols, 2014: p. 3-10.
4. in *Essentials of Glycobiology*, A. Varki, et al., Editors. 2009, Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).
5. Muramatsu, T., *Basigin (CD147), a multifunctional transmembrane glycoprotein with various binding partners*. The Journal of Biochemistry, 2015. **159**(5): p. 481-490.
6. *UniProt: the universal protein knowledgebase*. Nucleic Acids Research, 2017. **45**(D1): p. D158-D169.
7. Hattangadi, S.M., et al., *From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications*. Blood, 2011. **118**(24): p. 6258-6268.
8. Bai, Y., et al., *Importance of N-glycosylation on CD147 for its biological functions*. International journal of molecular sciences, 2014. **15**(4): p. 6356-6377.
9. Rørvig, S., et al., *Proteome profiling of human neutrophil granule subsets, secretory vesicles, and cell membrane: correlation with transcriptome profiling of neutrophil precursors*. Journal of Leukocyte Biology, 2013. **94**(4): p. 711-721.
10. Grass, G.D. and B.P. Toole, *How, with whom and when: an overview of CD147-mediated regulatory networks influencing matrix metalloproteinase activity*. Bioscience reports, 2016. **36**(1): p. e00283.
11. Japan, L.C.S.G.O., *The general rules for the clinical and pathological study of primary liver cancer*. The Japanese journal of surgery, 1989. **19**(1): p. 98-129.
12. Ramalho, M., et al., *Magnetic resonance imaging of the cirrhotic liver: diagnosis of hepatocellular carcinoma and evaluation of response to treatment - Part 1*. Radiol Bras, 2017. **50**(1): p. 38-47.
13. Okuda, K., et al., *Natural history of hepatocellular carcinoma and prognosis in relation to treatment study of 850 patients*. Cancer, 1985. **56**(4): p. 918-928.
14. Blomme, B., et al., *Alteration of protein glycosylation in liver diseases*. J Hepatol, 2009. **50**(3): p. 592-603.
15. Udby, L. and N. Borregaard, *Subcellular fractionation of human neutrophils and analysis of subcellular markers*. Methods Mol Biol, 2007. **412**: p. 35-56.
16. Thaysen-Andersen, M., et al., *Human neutrophils secrete bioactive paucimannosidic proteins from azurophilic granules into pathogen-infected sputum*. Journal of Biological Chemistry, 2015. **290**(14): p. 8789-8802.
17. Loke, I., et al., *Paucimannose-Rich N-glycosylation of Spatiotemporally Regulated Human Neutrophil Elastase Modulates Its Immune Functions*. Molecular & Cellular Proteomics, 2017. **16**(8): p. 1507-1527.
18. Loke, I., N.H. Packer, and M. Thaysen-Andersen, *Complementary LC-MS/MS-based N-glycan, N-glycopeptide, and intact N-glycoprotein profiling reveals unconventional Asn71-glycosylation of human neutrophil cathepsin G*. Biomolecules, 2015. **5**(3): p. 1832-1854.
19. Jiang, J.L., et al., *The involvement of HAb18G/CD147 in regulation of store-operated calcium entry and metastasis of human hepatoma cells*. J Biol Chem, 2001. **276**(50): p. 46870-7.
20. Allen, R.C., R.L. Stjernholm, and R.H. Steele, *Evidence for the generation of an electronic excitation state (s) in human polymorphonuclear leukocytes and its participation in*

- bactericidal activity*. Biochemical and biophysical research communications, 1972. **47**(4): p. 679-684.
21. Chang, H.H., et al., *Multistable and multistep dynamics in neutrophil differentiation*. BMC cell biology, 2006. **7**(1): p. 11.
 22. Maallem, H., K. Sheppard, and J. Fletcher, *The discharge of primary and secondary granules during immune phagocytosis by normal and chronic granulocytic leukaemia polymorphonuclear neutrophils*. Br J Haematol, 1982. **51**(2): p. 201-8.
 23. Biswas, C., et al., *The human tumor cell-derived collagenase stimulatory factor (renamed EMMPRIN) is a member of the immunoglobulin superfamily*. Cancer research, 1995. **55**(2): p. 434-439.
 24. Biswas, C., *Tumor cell stimulation of collagenase production by fibroblasts*. Biochemical and biophysical research communications, 1982. **109**(3): p. 1026-1034.
 25. Miyauchi, T., Y. Masuzawa, and T. Muramatsu, *The basigin group of the immunoglobulin superfamily: complete conservation of a segment in and around transmembrane domains of human and mouse basigin and chicken HT7 antigen*. J Biochem, 1991. **110**(5): p. 770-4.
 26. Kasinrerk, W., et al., *Human leukocyte activation antigen M6, a member of the Ig superfamily, is the species homologue of rat OX-47, mouse basigin, and chicken HT7 molecule*. J Immunol, 1992. **149**(3): p. 847-54.
 27. Fossum, S., S. Mallett, and A.N. Barclay, *The MRC OX-47 antigen is a member of the immunoglobulin superfamily with an unusual transmembrane sequence*. Eur J Immunol, 1991. **21**(3): p. 671-9.
 28. Nehme, C.L., et al., *Breaching the diffusion barrier that compartmentalizes the transmembrane glycoprotein CE9 to the posterior-tail plasma membrane domain of the rat spermatozoon*. The Journal of Cell Biology, 1993. **120**(3): p. 687-694.
 29. Altruda, F., et al., *Cloning of cDNA for a novel mouse membrane glycoprotein (gp42): shared identity to histocompatibility antigens, immunoglobulins and neural-cell adhesion molecules*. Gene, 1989. **85**(2): p. 445-51.
 30. Ochrietor, J.D., et al., *Retina-specific expression of 5A11/Basigin-2, a member of the immunoglobulin gene superfamily*. Invest Ophthalmol Vis Sci, 2003. **44**(9): p. 4086-96.
 31. Fadool, J.M. and P.J. Linser, *5A11 antigen is a cell recognition molecule which is involved in neuronal-glial interactions in avian neural retina*. Dev Dyn, 1993. **196**(4): p. 252-62.
 32. Schlosshauer, B. and K.H. Herzog, *Neurothelin: an inducible cell surface glycoprotein of blood-brain barrier-specific endothelial cells and distinct neurons*. J Cell Biol, 1990. **110**(4): p. 1261-74.
 33. Kanekura, T., et al., *Basigin, a new member of the immunoglobulin superfamily: genes in different mammalian species, glycosylation changes in the molecule from adult organs and possible variation in the N-terminal sequences*. Cell Struct Funct, 1991. **16**(1): p. 23-30.
 34. Fan, Q.-W., et al., *Expression of basigin, a member of the immunoglobulin superfamily, in the mouse central nervous system*. Neuroscience research, 1998. **30**(1): p. 53-63.
 35. Miyauchi, T., et al., *Basigin, a new, broadly distributed member of the immunoglobulin superfamily, has strong homology with both the immunoglobulin V domain and the β -chain of major histocompatibility complex class II antigen*. Journal of biochemistry, 1990. **107**(2): p. 316-323.
 36. Kaname, T., et al., *Mapping basigin (BSG), a member of the immunoglobulin superfamily, to 19p13.3*. Cytogenetic and Genome Research, 1993. **64**(3-4): p. 195-197.
 37. Muramatsu, T., *Basigin (CD147), a multifunctional transmembrane glycoprotein with various binding partners*. J Biochem, 2016. **159**(5): p. 481-90.
 38. Bai, Y., et al., *Importance of N-glycosylation on CD147 for its biological functions*. Int J Mol Sci, 2014. **15**(4): p. 6356-77.

39. Yu, X.L., et al., *Crystal structure of HAb18G/CD147: implications for immunoglobulin superfamily homophilic adhesion*. J Biol Chem, 2008. **283**(26): p. 18056-65.
40. Redzic, J.S., et al., *The retinal specific CD147 IgO domain: from molecular structure to biological activity*. Journal of molecular biology, 2011. **411**(1): p. 68-82.
41. Muramatsu, T. and T. Miyauchi, *Basigin (CD147): a multifunctional transmembrane protein involved in reproduction, neural function, inflammation and tumor invasion*. Histology and histopathology, 2003. **18**(3): p. 981-987.
42. Ochrietor, J.D., et al., *Retina-specific expression of 5A11/Basigin-2, a member of the immunoglobulin gene superfamily*. Investigative ophthalmology & visual science, 2003. **44**(9): p. 4086-4096.
43. Fadool, J.M. and P.J. Linser, *Differential glycosylation of the 5A11/HT7 antigen by neural retina and epithelial tissues in the chicken*. Journal of neurochemistry, 1993. **60**(4): p. 1354-1364.
44. Huang, W., et al., *Modulation of CD147-induced matrix metalloproteinase activity: role of CD147 N-glycosylation*. Biochemical Journal, 2013. **449**(2): p. 437-448.
45. Gabison, E.E., et al., *EMMPRIN/CD147, an MMP modulator in cancer, development and tissue repair*. Biochimie, 2005. **87**(3): p. 361-368.
46. Wright, K.E., et al., *Structure of malaria invasion protein RH5 with erythrocyte basigin and blocking antibodies*. Nature, 2014. **515**(7527): p. 427.
47. Ait-Ali, N., et al., *Rod-derived cone viability factor promotes cone survival by stimulating aerobic glycolysis*. Cell, 2015. **161**(4): p. 817-832.
48. WEIDLE, U.H., et al., *Cancer-related issues of CD147*. Cancer Genomics-Proteomics, 2010. **7**(3): p. 157-169.
49. Knutti, N., M. Kuepper, and K. Friedrich, *Soluble extracellular matrix metalloproteinase inducer (EMMPRIN, EMN) regulates cancer-related cellular functions by homotypic interactions with surface CD147*. FEBS journal, 2015. **282**(21): p. 4187-4200.
50. Kluger, R. and A. Alagic, *Chemical cross-linking and protein-protein interactions—a review with illustrative protocols*. Bioorganic chemistry, 2004. **32**(6): p. 451-472.
51. Berditchevski, F., et al., *Generation of Monoclonal Antibodies to Integrin-associated Proteins EVIDENCE THAT $\alpha 3 \beta 1$ COMPLEXES WITH EMMPRIN/BASIGIN/OX47/M6*. Journal of Biological Chemistry, 1997. **272**(46): p. 29174-29180.
52. Schlegel, J., et al., *Solution characterization of the extracellular region of CD147 and its interaction with its enzyme ligand cyclophilin A*. Journal of molecular biology, 2009. **391**(3): p. 518-535.
53. Hibino, T., et al., *S100A9 is a novel ligand of EMMPRIN that promotes melanoma metastasis*. Cancer research, 2013. **73**(1): p. 172-183.
54. Seizer, P., et al., *EMMPRIN (CD147) is a novel receptor for platelet GPVI and mediates platelet rolling via GPVI-EMMPRIN interaction*. Thrombosis and haemostasis, 2009. **101**(4): p. 682-686.
55. Igakura, T., et al., *A null mutation in basigin, an immunoglobulin superfamily member, indicates its important roles in peri-implantation development and spermatogenesis*. Dev Biol, 1998. **194**(2): p. 152-65.
56. Kuno, N., et al., *Female sterility in mice lacking the basigin gene, which encodes a transmembrane glycoprotein belonging to the immunoglobulin superfamily*. FEBS letters, 1998. **425**(2): p. 191-194.
57. Philp, N.J., et al., *Loss of MCT1, MCT3, and MCT4 expression in the retinal pigment epithelium and neural retina of the 5A11/basigin-null mouse*. Investigative ophthalmology & visual science, 2003. **44**(3): p. 1305-1311.

58. Grivennikov, S.I., F.R. Greten, and M. Karin, *Immunity, inflammation, and cancer*. Cell, 2010. **140**(6): p. 883-899.
59. Ge, H., et al., *Functional relevance of protein glycosylation to the pro-inflammatory effects of extracellular matrix metalloproteinase inducer (EMMPRIN) on monocytes/macrophages*. PloS one, 2015. **10**(2): p. e0117463.
60. Mann, M., et al., *Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome*. Trends in biotechnology, 2002. **20**(6): p. 261-268.
61. Ham, B.M., *Proteomics of biological systems: protein phosphorylation using mass spectrometry techniques*. 2011: John Wiley & Sons.
62. Rudd, P.M. and R.A. Dwek, *Glycosylation: heterogeneity and the 3D structure of proteins*. Critical reviews in biochemistry and molecular biology, 1997. **32**(1): p. 1-100.
63. Wormald, M.R., et al., *Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling*. Chemical Reviews, 2002. **102**(2): p. 371-386.
64. Yuriev, E., et al., *Differential site accessibility mechanistically explains subcellular-specific N-glycosylation determinants*. 2014.
65. Waechter, C.J. and W.J. Lennarz, *The role of polyprenol-linked sugars in glycoprotein synthesis*. Annual review of biochemistry, 1976. **45**(1): p. 95-112.
66. Herscovics, A., *Importance of glycosidases in mammalian glycoprotein biosynthesis*. Biochimica et Biophysica Acta (BBA)-General Subjects, 1999. **1473**(1): p. 96-107.
67. Kornfeld, R. and S. Kornfeld, *Assembly of asparagine-linked oligosaccharides*. Annual review of biochemistry, 1985. **54**(1): p. 631-664.
68. Lowe, J.B. and J.D. Marth, *A genetic approach to mammalian glycan function*. Annual review of biochemistry, 2003. **72**(1): p. 643-691.
69. Yang, L., et al., *Targeted identification of metastasis-associated cell-surface sialoglycoproteins in prostate cancer*. Molecular & Cellular Proteomics, 2011. **10**(6): p. M110. 007294.
70. Tang, W., S.B. Chang, and M.E. Hemler, *Links between CD147 function, glycosylation, and caveolin-1*. Molecular biology of the cell, 2004. **15**(9): p. 4043-4050.
71. Grass, G.D. and B.P. Toole, *How, with whom and when: an overview of CD147-mediated regulatory networks influencing matrix metalloproteinase activity*. Biosci Rep, 2015. **36**(1): p. e00283.
72. Thaysen-Andersen, M. and N.H. Packer, *Advances in LC-MS/MS-based glycoproteomics: getting closer to system-wide site-specific mapping of the N-and O-glycoproteome*. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2014. **1844**(9): p. 1437-1452.
73. Jensen, P.H., et al., *Structural analysis of N-and O-glycans released from glycoproteins*. Nature protocols, 2012. **7**(7): p. 1299.
74. Lee, L.Y., et al., *Toward automated N-glycopeptide identification in glycoproteomics*. Journal of proteome research, 2016. **15**(10): p. 3904-3915.
75. Zucker, M.B. and J. Borrelli, *Some effects of divalent cations on the clotting mechanism and the platelets of EDTA blood*. Journal of applied physiology, 1958. **12**(3): p. 453-460.
76. Zhou, L., et al., *Impact of human granulocyte and monocyte isolation procedures on functional studies*. Clinical and Vaccine Immunology, 2012. **19**(7): p. 1065-1074.
77. Bøyum, A., *Isolation of lymphocytes, granulocytes and macrophages*. Scandinavian Journal of Immunology, 1976. **5**(s5): p. 9-15.
78. Seeger, F.H., et al., *Cell isolation procedures matter: a comparison of different isolation protocols of bone marrow mononuclear cells used for cell therapy in patients with acute myocardial infarction*. European heart journal, 2007. **28**(6): p. 766-772.

79. Khan, A., et al., *Simplified evaluation of apoptosis using the Muse cell analyzer*. *Postepy biochemii*, 2012. **58**(4): p. 492-496.
80. Natt, M.P. and C.A. Herrick, *A new blood diluent for counting the erythrocytes and leucocytes of the chicken*. *Poultry Science*, 1952. **31**(4): p. 735-738.
81. Woronzoff-Dashkoff, K.K., *The Wright-Giemsa stain: secrets revealed*. *Clinics in laboratory medicine*, 2002. **22**(1): p. 15-23.
82. Bermejo, P., et al., *Enzymatic digestion and ultrasonication: a powerful combination in analytical chemistry*. *TrAC Trends in Analytical Chemistry*, 2004. **23**(9): p. 654-663.
83. Rose, H.G. and M. Oklander, *Improved procedure for the extraction of lipids from human erythrocytes*. *Journal of lipid research*, 1965. **6**(3): p. 428-431.
84. Walker, J.M., *The bicinchoninic acid (BCA) assay for protein quantitation*. *Basic protein and peptide protocols*, 1994: p. 5-8.
85. Brown, R.E., K.L. Jarvis, and K.J. Hyland, *Protein measurement using bicinchoninic acid: elimination of interfering substances*. *Analytical biochemistry*, 1989. **180**(1): p. 136-139.
86. Clemmensen, S.N., L. Udby, and N. Borregaard, *Subcellular fractionation of human neutrophils and analysis of subcellular markers*. *Neutrophil Methods and Protocols*, 2014: p. 53-76.
87. Borregaard, N. and J.B. Cowland, *Granules of the human neutrophilic polymorphonuclear leukocyte*. *Blood*, 1997. **89**(10): p. 3503-3521.
88. Borregaard, N., et al., *Identification of a highly mobilizable subset of human neutrophil intracellular vesicles that contains tetranectin and latent alkaline phosphatase*. *Journal of Clinical Investigation*, 1990. **85**(2): p. 408.
89. Lominadze, G., et al., *Proteomic analysis of human neutrophil granules*. *Molecular & Cellular Proteomics*, 2005. **4**(10): p. 1503-1521.
90. Fellig, Y., et al., *A hepatocellular carcinoma cell line producing mature hepatitis B viral particles*. *Biochemical and biophysical research communications*, 2004. **321**(2): p. 269-274.
91. Hiraoka, Y., J.W. Sedat, and D.A. Agard, *Determination of three-dimensional imaging properties of a light microscope system. Partial confocal behavior in epifluorescence microscopy*. *Biophysical journal*, 1990. **57**(2): p. 325-333.
92. Lu, H., M.A. Schmidt, and K.F. Jensen, *A microfluidic electroporation device for cell lysis*. *Lab on a Chip*, 2005. **5**(1): p. 23-29.
93. Lee, L.Y., et al., *Differential site accessibility mechanistically explains subcellular-specific N-glycosylation determinants*. *Frontiers in immunology*, 2014. **5**.
94. Crowell, A.M., M.J. Wall, and A.A. Doucette, *Maximizing recovery of water-soluble proteins through acetone precipitation*. *Analytica chimica acta*, 2013. **796**: p. 48-54.
95. Wyckoff, M., D. Rodbard, and A. Chrambach, *Polyacrylamide gel electrophoresis in sodium dodecyl sulfate-containing buffers using multiphasic buffer systems: Properties of the stack, valid R_f– measurement, and optimized procedure*. *Analytical biochemistry*, 1977. **78**(2): p. 459-482.
96. Swank, R. and K. Munkres, *Molecular weight analysis of oligopeptides by electrophoresis in polyacrylamide gel with sodium dodecyl sulfate*. *Analytical biochemistry*, 1971. **39**(2): p. 462-477.
97. Rabilloud, T., et al., *Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients*. *Electrophoresis*, 1997. **18**(3-4): p. 307-316.
98. Zehr, B.D., T.J. Savin, and R.E. Hall, *A one-step, low background Coomassie staining procedure for polyacrylamide gels*. *Analytical biochemistry*, 1989. **182**(1): p. 157-159.
99. Shevchenko, A., et al., *In-gel digestion for mass spectrometric characterization of proteins and proteomes*. *Nature protocols*, 2007. **1**(6): p. 2856-2860.

100. Stott, D., *Immunoblotting and dot blotting*. Journal of immunological methods, 1989. **119**(2): p. 153-187.
101. Eisenthal, R. and M.J. Danson, *Enzyme assays: a practical approach*. 2002: Practical Approach (Paperback).
102. Karhemo, P.-R., et al., *An optimized isolation of biotinylated cell surface proteins reveals novel players in cancer metastasis*. Journal of proteomics, 2012. **77**: p. 87-100.
103. Burgess, A., *Affinity purification of plasma membranes*. Journal of biomolecular techniques: JBT, 1999. **10**(2): p. 64.
104. Jang, J.H. and S. Hanash, *Profiling of the cell surface proteome*. Proteomics, 2003. **3**(10): p. 1947-1954.
105. Shin, B.K., et al., *Global profiling of the cell surface proteome of cancer cells uncovers an abundance of proteins with chaperone function*. Journal of Biological Chemistry, 2003. **278**(9): p. 7607-7616.
106. Tarentino, A.L., C.M. Gomez, and T.H. Plummer Jr, *Deglycosylation of asparagine-linked glycans by peptide: N-glycosidase F*. Biochemistry, 1985. **24**(17): p. 4665-4671.
107. Domon, B. and C.E. Costello, *A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates*. Glycoconjugate journal, 1988. **5**(4): p. 397-409.
108. Cooper, C.A., E. Gasteiger, and N.H. Packer, *GlycoMod—a software tool for determining glycosylation compositions from mass spectrometric data*. Proteomics, 2001. **1**(2): p. 340-349.
109. Ceroni, A., et al., *GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans*. Journal of proteome research, 2008. **7**(4): p. 1650-1659.
110. Everest-Dass, A.V., et al., *Structural feature ions for distinguishing N-and O-linked glycan isomers by LC-ESI-IT MS/MS*. Journal of The American Society for Mass Spectrometry, 2013. **24**(6): p. 895-906.
111. Pabst, M., et al., *Mass+ retention time= structure: a strategy for the analysis of N-glycans by carbon LC-ESI-MS and its application to fibrin N-glycans*. Analytical chemistry, 2007. **79**(13): p. 5051-5057.
112. Varki, A., et al., *Symbol nomenclature for graphical representations of glycans*. Glycobiology, 2015. **25**(12): p. 1323-1324.
113. Abrahams, J.L., M.P. Campbell, and N.H. Packer, *Building a PGC-LC-MS N-glycan retention library and elution mapping resource*. 2017.
114. Mysling, S., et al., *Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics*. Analytical chemistry, 2010. **82**(13): p. 5598-5609.
115. Stadlmann, J., et al., *Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides*. Proteomics, 2008. **8**(14): p. 2858-2871.
116. Peri, S., H. Steen, and A. Pandey, *GPMAW—a software tool for analyzing proteins and peptides*. 2001, Elsevier.
117. Elias, J.E., et al., *Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations*. Nature methods, 2005. **2**(9): p. 667.
118. Bern, M., Y.J. Kil, and C. Becker, *Byonic: advanced peptide and protein identification software*. Current protocols in bioinformatics, 2012: p. 13.20. 1-13.20. 14.
119. Colaert, N., et al., *Thermo-msf-parser: an open source Java library to parse and visualize Thermo Proteome Discoverer msf files*. Journal of proteome research, 2011. **10**(8): p. 3840-3843.
120. Storey, J.D., *False discovery rate*, in *International encyclopedia of statistical science*. 2011, Springer. p. 504-508.

121. Huddleston, M.J., M.F. Bean, and S.A. Carr, *Collisional fragmentation of glycopeptides by electrospray ionization LC/MS and LC/MS/MS: methods for selective detection of glycopeptides in protein digests*. Analytical chemistry, 1993. **65**(7): p. 877-884.
122. Wuhrer, M., et al., *Glycoproteomics based on tandem mass spectrometry of glycopeptides*. Journal of Chromatography B, 2007. **849**(1): p. 115-128.
123. Jebanathirajah, J., H. Steen, and P. Roepstorff, *Using optimized collision energies and high resolution, high accuracy fragment ion selection to improve glycopeptide detection by precursor ion scanning*. Journal of the American Society for Mass Spectrometry, 2003. **14**(7): p. 777-784.
124. Borregaard, N., O.E. Sørensen, and K. Theilgaard-Mönch, *Neutrophil granules: a library of innate immunity proteins*. Trends in immunology, 2007. **28**(8): p. 340-345.
125. Parker, B.L., et al., *Terminal galactosylation and sialylation switching on membrane glycoproteins upon TNF-alpha-induced insulin resistance in adipocytes*. Molecular & Cellular Proteomics, 2016. **15**(1): p. 141-153.
126. Zhao, S.-H., et al., *Basigin-2 is the predominant basigin isoform that promotes tumor cell migration and invasion and correlates with poor prognosis in epithelial ovarian cancer*. Journal of Translational Medicine, 2013. **11**(1): p. 92.
127. Apweiler, R., et al., *UniProt: the universal protein knowledgebase*. Nucleic acids research, 2004. **32**(suppl_1): p. D115-D119.
128. Weidle, U.H., et al., *Cancer-related issues of CD147*. Cancer Genomics Proteomics, 2010. **7**(3): p. 157-69.
129. Kanekura, T. and X. Chen, *CD147/basigin promotes progression of malignant melanoma and other cancers*. J Dermatol Sci, 2010. **57**.
130. Verlaan-de Vries, M., et al., *A dot-blot screening procedure for mutated ras oncogenes using synthetic oligodeoxynucleotides*. Gene, 1986. **50**(1): p. 313-320.
131. Udby, L. and N. Borregaard, *Subcellular fractionation of human neutrophils and analysis of subcellular markers*. Neutrophil Methods and Protocols, 2007: p. 35-56.
132. Dunkley, P.R., et al., *A rapid Percoll gradient procedure for isolation of synaptosomes directly from an S 1 fraction: homogeneity and morphology of subcellular fractions*. Brain research, 1988. **441**(1): p. 59-71.
133. Moya-Quiles, M.R., et al., *Alkaline treatment of muscle microsomes releases amphiphilic and hydrophilic forms of acetylcholinesterase*. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology, 1992. **1121**(1-2): p. 88-96.
134. Van Pelt, J.F., et al., *Identification of HepG2 variant cell lines by short tandem repeat (STR) analysis*. Molecular and cellular biochemistry, 2003. **243**(1): p. 49-54.
135. Ramalho, M., et al., *Magnetic resonance imaging of the cirrhotic liver: diagnosis of hepatocellular carcinoma and evaluation of response to treatment - Part 1*. Radiologia Brasileira, 2017. **50**(1): p. 38-47.
136. Wollscheid, B., et al., *Mass-spectrometric identification and relative quantification of N-linked cell surface glycoproteins*. Nature biotechnology, 2009. **27**(4): p. 378-386.
137. Deller, M.C. and E.Y. Jones, *Cell surface receptors*. Current opinion in structural biology, 2000. **10**(2): p. 213-219.
138. Al-Shahrour, F., R. Diaz-Uriarte, and J. Dopazo, *FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes*. Bioinformatics, 2004. **20**.
139. Bian, Y., et al., *An enzyme assisted RP-RPLC approach for in-depth analysis of human liver phosphoproteome*. Journal of proteomics, 2014. **96**: p. 253-262.
140. Mann, M., et al., *Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome*. Trends Biotechnol, 2002. **20**(6): p. 261-8.

141. Di Palma, S., et al., *Zwitterionic hydrophilic interaction liquid chromatography (ZIC-HILIC and ZIC-cHILIC) provide high resolution separation and increase sensitivity in proteome analysis*. Analytical chemistry, 2011. **83**(9): p. 3440-3447.
142. Cuyckens, F., et al., *Tandem mass spectral strategies for the structural characterization of flavonoid glycosides*. Analisis, 2000. **28**(10): p. 888-895.
143. Ito, H., et al., *Comparison of analytical methods for profiling N-and O-linked glycans from cultured cell lines*. Glycoconjugate journal, 2016. **33**(3): p. 405-415.
144. Thaysen-Andersen, M., et al., *Recombinant human heterodimeric IL-15 complex displays extensive and reproducible N*. Glycoconjugate journal, 2016. **33**(3): p. 417-433.
145. Wang, Q., et al., *Glycoproteomic Analysis of Human Hepatoblastoma Cell Lines Using Glycopeptide Capture and Mass Spectrometry*. Journal of Glycomics & Lipidomics, 2015. **5**(3): p. 1.
146. Liao, C.G., et al., *Characterization of basigin isoforms and the inhibitory function of basigin-3 in human hepatocellular carcinoma proliferation and invasion*. Mol Cell Biol, 2011. **31**.
147. Dahmen, A.-C., et al., *Paucimannosidic glycoepitopes are functionally involved in proliferation of neural progenitor cells in the subventricular zone*. Glycobiology, 2015. **25**(8): p. 869-880.
148. Belton, R.J., et al., *Basigin-2 is a cell surface receptor for soluble basigin ligand*. J Biol Chem, 2008. **283**.
149. Muramatsu, T. and T. Miyauchi, *Basigin (CD147): a multifunctional transmembrane protein involved in reproduction, neural function, inflammation and tumor invasion*. Histol Histopathol, 2003. **18**.
150. Bekesova, S., et al., *N-glycans in liver-secreted and immunoglobulin-derived protein fractions*. Journal of proteomics, 2012. **75**(7): p. 2216-2224.
151. Loke, I., et al., *Structure, function and biosynthesis of a new class of humann-glycosylated neutrophilic proteins in pathogen-infected sputum*. The Febs Journal, 2015. **282**: p. 47.
152. Nardy, A.F.F.R., et al., *The Sweet Side of Immune Evasion: Role of Glycans in the Mechanisms of Cancer Progression*. Frontiers in Oncology, 2016. **6**: p. 54.
153. Loke, I., et al., *Emerging roles of protein mannosylation in inflammation and infection*. Molecular aspects of medicine, 2016. **51**: p. 31-55.
154. Liu, F.-T., et al., *Expression and function of galectin-3, a beta-galactoside-binding lectin, in human monocytes and macrophages*. The American journal of pathology, 1995. **147**(4): p. 1016.
155. Rabinovich, G.A. and M.A. Toscano, *Turning'sweet'on immunity: galectin-glycan interactions in immune tolerance and inflammation*. Nature reviews. Immunology, 2009. **9**(5): p. 338.
156. van Kooyk, Y. and G.A. Rabinovich, *Protein-glycan interactions in the control of innate and adaptive immune responses*. Nat Immunol, 2008. **9**(6): p. 593-601.
157. Läubli, H. and L. Borsig. *Selectins promote tumor metastasis*. in *Seminars in cancer biology*. 2010. Elsevier.
158. Thaysen-Andersen, M. and N.H. Packer, *Site-specific glycoproteomics confirms that protein structure dictates formation of N-glycan type, core fucosylation and branching*. Glycobiology, 2012. **22**(11): p. 1440-1452.
159. Mathias, R.A., et al., *Triton X-114 phase separation in the isolation and purification of mouse liver microsomal membrane proteins*. Methods, 2011. **54**(4): p. 396-406.
160. Christiansen, M.N., et al., *Cell surface protein glycosylation in cancer*. Proteomics, 2014. **14**(4-5): p. 525-546.
161. Lee, L.Y., et al., *Comprehensive N-glycome profiling of cultured human epithelial breast cells identifies unique secretome N-glycosylation signatures enabling tumorigenic subtype classification*. J Proteome Res, 2014. **13**(11): p. 4783-95.

162. Sethi, M.K., et al., *Comparative N-glycan profiling of colorectal cancer cell lines reveals unique bisecting GlcNAc and α -2, 3-linked sialic acid determinants are associated with membrane proteins of the more metastatic/aggressive cell lines*. Journal of proteome research, 2013. **13**(1): p. 277-288.
163. Kamiyama, T., et al., *Identification of novel serum biomarkers of hepatocellular carcinoma using glycomic analysis*. Hepatology, 2013. **57**(6): p. 2314-2325.
164. Calabro, S.R., et al., *Hepatocyte produced matrix metalloproteinases are regulated by CD147 in liver fibrogenesis*. PloS one, 2014. **9**(7): p. e90571.
165. Calvisi, D.F., *CD147/Basigin: a Warburg oncogene in hepatocellular carcinoma?* Chinese Journal of Cancer Research, 2016. **28**(3): p. 377.

Appendices

Office of the Deputy Vice-Chancellor
(Research)

Research Office
Research Hub, Building C5C East
Macquarie University
NSW 2109 Australia
T: +61 (2) 9850 4459
<http://www.research.mq.edu.au/>
ABN 90 952 801 237



MACQUARIE
University
SYDNEY · AUSTRALIA

10 June 2015

Dr Morten Andersen
Macquarie University NSW 2109

Dear Dr Andersen,

Reference No: 5201500409

Title: *White blood cell differentiation and activation markers: Their relationship to diabetic complications*

Thank you for submitting the above application for ethical and scientific review. Your application was considered by the Macquarie University Human Research Ethics Committee (HREC (Medical Sciences)) at its meeting on 28 May 2015, at which further information was requested to be reviewed by the Ethics Secretariat.



The requested information was received with correspondence on 5 June 2015.

I am pleased to advise that ethical and scientific approval has been granted for this project to be conducted at:

- Macquarie University

This research meets the requirements set out in the *National Statement on Ethical Conduct in Human Research* (2007 – Updated March 2014) (the *National Statement*).

This letter constitutes ethical and scientific approval only.

SECTION A			
 		<h2 style="text-align: center;">Biohazard Risk Assessment Form – NON-GMO</h2>	
		Notification Number: NIP041214BHA	
Department	Chemistry and Biomolecular Sciences	Date:	04 Dec 2014
Chief investigator:	Professor Nicole Packer		
Contact number/email:	9850 8176 / nicki.packer@mq.edu.au		
Title of research/practical	A study of glycosylation and its changes in healthy and diseased cells and in bacteria and fungi		
Is additional approval required?	Animal Ethics <input checked="" type="checkbox"/> Human Ethics <input checked="" type="checkbox"/> Fieldwork Manager <input type="checkbox"/> Other <input type="checkbox"/> (state) _____		
Exact location(s) of research:			
E8A 109 PC2/QC2 laboratory, E8C 320 laboratory, E8C 323 APAF PC2 cell culture laboratory, E8C 326 BMD laboratory, ASAM Level 1 PC2 laboratory, E8C 252 PC2 laboratory (due to complete in February 2015)			
Control measures: Eliminate risk <input type="checkbox"/> Substitute the hazard <input type="checkbox"/> Isolate the hazard <input checked="" type="checkbox"/> Implement engineering controls <input checked="" type="checkbox"/> Administration <input checked="" type="checkbox"/> (e.g. Training) PPE <input checked="" type="checkbox"/> E.g. Eliminate by irradiation prior to use, isolation by class II biological safety cabinets, administration by following SWP as below, PPE as listed below.			
Isolation by class II biological safety cabinets Training by experienced personnel, laboratory induction and safe laboratory practice training conducted by experienced and authorised personnel Following good safe operating procedures, good microbiology practices and current biological waste disposal practices Wearing appropriate PPE in accordance to safe operating procedures			
Supporting documents which must be read in conjunction with this assessment. (e.g. Safe Working Procedures, Safety Data Sheets, Guidelines/Protocols)			
Safe working procedures: 1) Working with potential human pathogens and known human pathogens in PC2 laboratory 2) Working with peripheral blood in PC2 laboratory Safe work procedure risk assessment: 1) Working with peripheral blood in PC2 laboratory			
What is the type of the biological material?			
Bacteria <input checked="" type="checkbox"/> Fungi <input checked="" type="checkbox"/> Virus <input type="checkbox"/> Cell Line <input checked="" type="checkbox"/> Tissue <input checked="" type="checkbox"/> Parasite <input type="checkbox"/> Animal <input type="checkbox"/> Plant <input type="checkbox"/> Soil <input type="checkbox"/> Toxin <input type="checkbox"/> Prions <input type="checkbox"/> Nucleic Acid <input type="checkbox"/> other <input checked="" type="checkbox"/> Peripheral blood cells			
List the Personal Protective Equipment required:			
Gloves <input checked="" type="checkbox"/> Latex or Nitrile <input checked="" type="checkbox"/> Eye protection <input checked="" type="checkbox"/> Safety glasses <input checked="" type="checkbox"/> Clothing <input checked="" type="checkbox"/> Button up lab coat, PC2 surgical gowns			
Footwear <input checked="" type="checkbox"/> Enclosed covered shoes <input checked="" type="checkbox"/> Respiratory Protection <input checked="" type="checkbox"/> Face shield, N95 surgical mask (e.g. PF2 face mask) Other <input type="checkbox"/> _____			
What are the risks associated with this Biological Agent. (Can be more than one risk group depending on method)			
Risk Group	Details of Biohazards including risks associate with biological agent http://www.absa.org/riskgroups/index.html	Biosafety level	Risk Reduction Measures (must be followed by the researcher)
Group 1- Low individual and community risk (Microorganism that is unlikely to cause human, plant or animal disease)	Well characterised human cell lines from ATCC collection, eg. [HCT116 (colon cancer), ATCC CCL-247] [A549 (Lung cancer cells), ATCC CCL-185] [HL60 (Leukemia cells), ATCC CCL-240] [OCI-AML/OCI-AML2 (Leukemia cells), from Ontario Cancer Institute, Toronto, Canada] [Jurkat (Leukemia cell), ATCC TIB-152] [K562 (Leukemia cells), ATCC-243] [OPM-2 (Myeloma cells)] [KARPAS (Lymphoma cells)] [MAVER (Lymphoma cells)]	Biosafety Level 1	1 Standard laboratory procedures will be followed in accordance with Laboratory Microbiological Standards AS/NZ 2243:3:2010 and university guidelines (see supporting documents - Section A above) and include spillage and emergency response. 2 Investigator has attended university Biosafety training course (see 3) 3 Chief Investigator identified in Section A confirms that the researchers have received appropriate training and instruction or has adequate supervision and understands safe laboratory practice according to AS/NZ 2243:3:2010 and university guidelines (see supporting documents - Section A above)
Group 2- Moderate individual risk, limited community risk (Microorganism that is unlikely to be a significant risk to laboratory workers, the community/livestock/environment. Laboratory exposures may cause infection but effective treatment and preventative measures are available and the risk of spread is limited).	Human tissue samples (tested negative for HIV, Hepatitis A and B) derived from healthy patients or tumour samples from cancer patients Peripheral blood cells and serum/plasma Urine and faecal samples Bacteria, yeast, filamentous fungi	Biosafety Level 2	1 Standard laboratory procedures will be followed in accordance with Laboratory Microbiological Standards AS/NZ 2243:3:2010 and university guidelines which are appropriate for Risk Group 2 (see supporting documents - Section A above) and include spillage and emergency response. 2 Researcher has attended university Biosafety training course (see 3) 3 Chief Investigator identified in Section A confirms that the researcher has received appropriate training and instruction or has adequate supervision and understands safe laboratory practice according to AS/NZ 2243:3:2010 and university guidelines (see supporting documents - Section A above)
Group 3 -High individual risk, limited community risk (Microorganisms that usually causes serious human or animal disease and may present a significant risk to laboratory workers. It could present a limited to moderate risk if spread in the community or the environment, but there are usually effective preventative measures or treatment available).	NA		1 Standard laboratory procedures will be followed in accordance with Laboratory Microbiological Standards AS/NZ 2243:3:2010 and university guidelines which are appropriate for Risk Group 3 (see supporting documents - Section A above) and include spillage and emergency response. 2 Researcher has attended university Biosafety training course (see 3) 3 Chief Investigator identified in Section A confirms that the investigator has received appropriate training and instruction or has adequate supervision and understands safe laboratory practice according to AS/NZ 2243:3:2010 and university guidelines (see supporting documents - Section A above)
Process and equipment to be used		You must include: - Brief description of work, control measures (including aerosols), sample storage, transport of samples, clean up procedures, disinfectant and waste disposal.	

SECTION B	
	Biohazard Safety Committee – Risk Assessment Decision
<p>Important Information</p> <p>For NON-GMO investigations email this assessment to biohazard@mq.edu.au for approval by the Biohazard Safety Committee.</p> <p>Individual Responsibilities</p> <p>By submitting this assessment, the Chief Investigator identified in Section A, confirms that any supporting documents, training, guidance, instruction or protocols issued by the University will be followed so far as reasonably practicable to ensure the work is carried out without risk to health, safety or the environment. The Chief investigator is responsible for ensuring, so far as reasonably practicable the safety of researchers and others who may be affected by the work described within this document.</p> <p>Decision to be completed by the Biohazard Safety Committee:</p> <p>The Committee has agreed that this risk assessment is sufficient for investigations to commence? Yes <input checked="" type="checkbox"/> No <input type="checkbox"/> Further action required <input type="checkbox"/></p> <div style="border: 1px solid black; padding: 5px; margin-top: 5px;"> <p>Further Action/Comments:</p> <p>The committee has revised your risk assessment and provided the following feedback to consider.</p> <p>Spills: It may be worth mentioning to use paper towels soaked in <u>Virkon</u> 1% and wait for ten minutes before <u>proceed</u> with the cleaning up. This may be in the supporting documents.</p> <p>Your risk assessment is approved to proceed.</p> </div>	
Name of Approver (Committee Rep):	Joanne Cuomo
Date Approved:	10 th of December 2014
This Risk Assessment must be approved for work to commence	

Figure S1: Human Ethics and Bio-Safety Approval Forms

LC/MSD Trap XCT Plus Series 1100 Acquisition Details
<ul style="list-style-type: none"> Smart fragmentation (start/end amplitude 30-200%) at 1.0 V Isolation window of 4 <i>m/z</i> with a maximum accumulation time of 200 ms. Smart ion charge control (ICC) with <i>m/z</i> 900 and a target isolation of 80,000 ions per scan event. ESI performed using a capillary voltage of +3.2 Kv <ul style="list-style-type: none"> ➤ a nitrogen drying gas flow of 6 l/min at 325°C ➤ a nitrogen-based nebuliser pressure of 12 psi. Dynamic exclusion inactivated. The mass spectrometer calibrated using a tune mix (Agilent Technologies) with mass accuracy better than 0.5 Da.

Table S1: Acquisition details of PGC-LC-MS/MS (Trap XCT Plus Series 1100) for analyses of *N*-glycans.

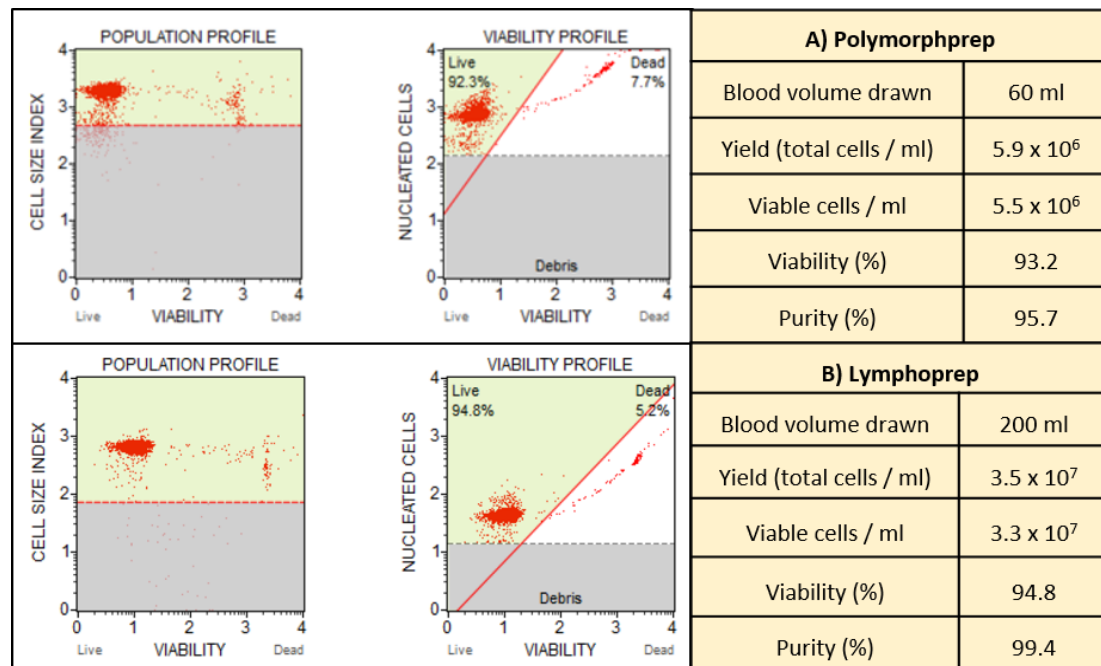


Figure S2: Comparison of neutrophil yield, viability and purity as determined by an automated cell counter (left graphs) and by microscopy of Giemsa-Wright stained neutrophils isolated using **(A)** Polymorphprep and **(B)** Lymphoprep. Averages of three technical replicates ($n = 3$) are given.

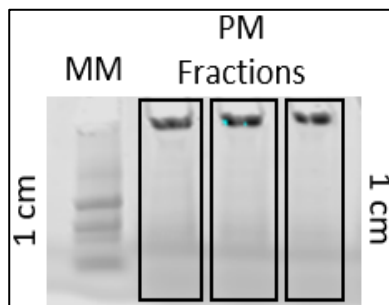


Figure S3: SDS-PAGE analysis (run up to 1 cm only) of proteins extracted from enriched plasma membrane fractions from neutrophils. MM is molecular marker (in kDa) and PM are plasma membrane fractions. The bands, as marked in black boxes were cut out for in-gel trypsin digestion for analysis on mass spectrometry.

Search Variables	Selected Search Parameters
Data Input	High-Resolution Orbitrap Fusion HCD-MS/MS
Proteome Database	a) Narrow: Human Basigin (P35613, Isoforms 1-4, UniProtKB) b) Broad: Entire Human Proteome (20,198 entries, UniProtKB)
N-Glycan Database	a) Broad: 309 Common Mammalian N-glycan Monosaccharide Compositions
Enzyme	Trypsin (KR)
Cleavage Residues	C-terminal cutter
Peptide Terminus	Fully Specific
Number of Missed Cleavages	2
Charge States	+1, +2, +3, +4
Precursor Tolerance	6.0 ppm
Fragment Tolerance (HCD)	10.0 ppm
Fragment Tolerance (ETD)	20.0 ppm
Peptide Variability	Standard Peptide Modification: Fixed Carbamidomethylation (Cys), Variable Oxidation (Met) and Deamidation (Asn/Gln)

Table S2: Byonic search parameters using raw fusion files for analyses of enriched plasma membrane fractions from neutrophils.

Search Variables	Selected Search Parameters
Data Input	Low-Resolution Bruker Ion-Trap CID-MS/MS
Proteome Database	a) Narrow: Human Basigin (P35613, Isoforms 1-4, UniProtKB) b) Broad: Entire Human Proteome (20,198 entries, UniProtKB)
N-Glycan Database	a) Broad: 309 Common Mammalian N-glycan Monosaccharide Compositions
Enzyme	Trypsin (KR)
Cleavage Residues	C-terminal cutter
Peptide Terminus	Fully Specific
Number of Missed Cleavages	-1 (Any)
Charge States	+1, +2, +3, +4
Precursor Tolerance	1.6 Da
Fragment Tolerance	1.6 Da
Peptide Variability	Standard Peptide Modification: Fixed Carbamidomethylation (Cys), Variable Oxidation (Met) and Deamidation (Asn/Gln)

Table S3: Byonic search parameters using raw ion-trap files for analyses of IP'ed fractions obtained from HepG2 cells.

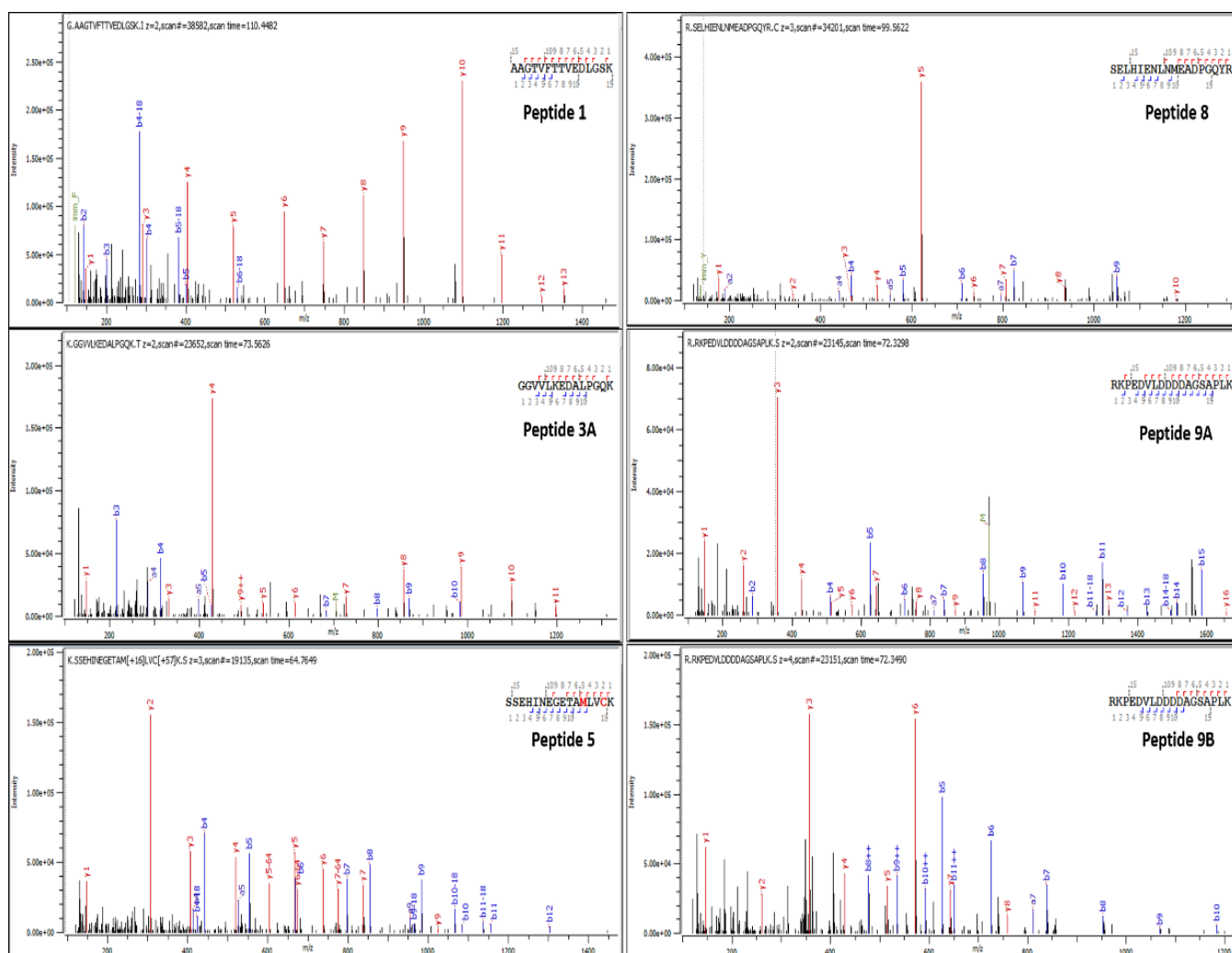


Figure S4: Annotated MS/MS spectrum obtained from Byonic search for six out of the total nine peptides originating from human basigin-1 isoform-2 from enriched plasma membrane fractions of neutrophils. The indicated peptide numbers match as provided in **Figure 17**.

Glycan #	m/z	Charge (z)	Observed [M,Da]	Calculated [M,Da]	Delta Mass (Da)	Monosaccharide Composition				Glycan Type	MS/MS	Observed In Figure 25
						Hex	Hex NAc	Fuc	NeuAc			
1	698.3	-2	1398.6	1398.5	0.10	6	2	0	0	High Mannose	Yes	(A-D)
2	749.4	-1	750.4	750.3	0.10	2	2	0	0	Paucimannose	Yes	A), (C), (D)
3	779.3	-2	1560.6	1560.5	0.10	7	2	0	0	High Mannose	Yes	(A-D)
4	860.3	-2	1722.6	1722.6	0.20	8	2	0	0	High Mannose	Yes*	(A-D)
5	895.4	-1	896.4	896.3	0.06	2	2	1	0	Paucimannose	Yes	(A), (C), (D)
6	911.4	-1	912.4	912.3	0.06	3	2	0	0	Paucimannose	Yes	(A)
7	941.4	-2	1884.8	1884.6	0.15	9	2	0	0	High Mannose	Yes	(A-D)
8	1022.3	-2	2046.6	2046.7	0.10	10	2	0	0	High Mannose	Yes	(A), (B), (D)
9	1038.8	-2	2079.6	2079.7	0.10	5	4	1	1	Complex	Yes	(A-D)
10	1057.4	-1	1058.4	1058.4	0.10	3	2	1	0	Paucimannose	Yes	(A), (C), (D)
11	1073.4	-1	1074.4	1074.4	0.10	4	2	0	0	High Mannose	Yes	(A), (C)
12	1111.4	-2	2224.8	2224.8	0.10	5	4	0	2	Complex	Yes	(A), (B), (C)
13	1184.5	-2	2371.0	2370.8	0.10	5	4	1	2	Complex	Yes	(A-D)
14	1235.5	-1	1236.5	1236.4	0.10	5	2	0	0	High Mannose	Yes	(A-D)
15	1260.5	-1	1261.5	1261.5	0.10	3	3	1	0	Complex	Yes	(A), (C), (D)
16	733.3	-1	734.3	734.3	0.10	1	2	1	0	Paucimannose	Yes	(C), (D)
17	966.3	-2	1934.7	1933.7	0.10	5	4	0	1	Complex	Yes	(C), (D)
18	1112	-2	2226.1	2226.1	0.10	5	4	2	1	Complex	Yes	(D)

Table S4: Overview of the observed *N*-glycans of HepG2 and neutrophils (see **Figures 22 and S5**). Glycans highlighted in green indicate the common structures observed in all fractions from both cell types. *Example of annotated MS/MS spectra of the identified *N*-glycan are presented in **Figure S6**.

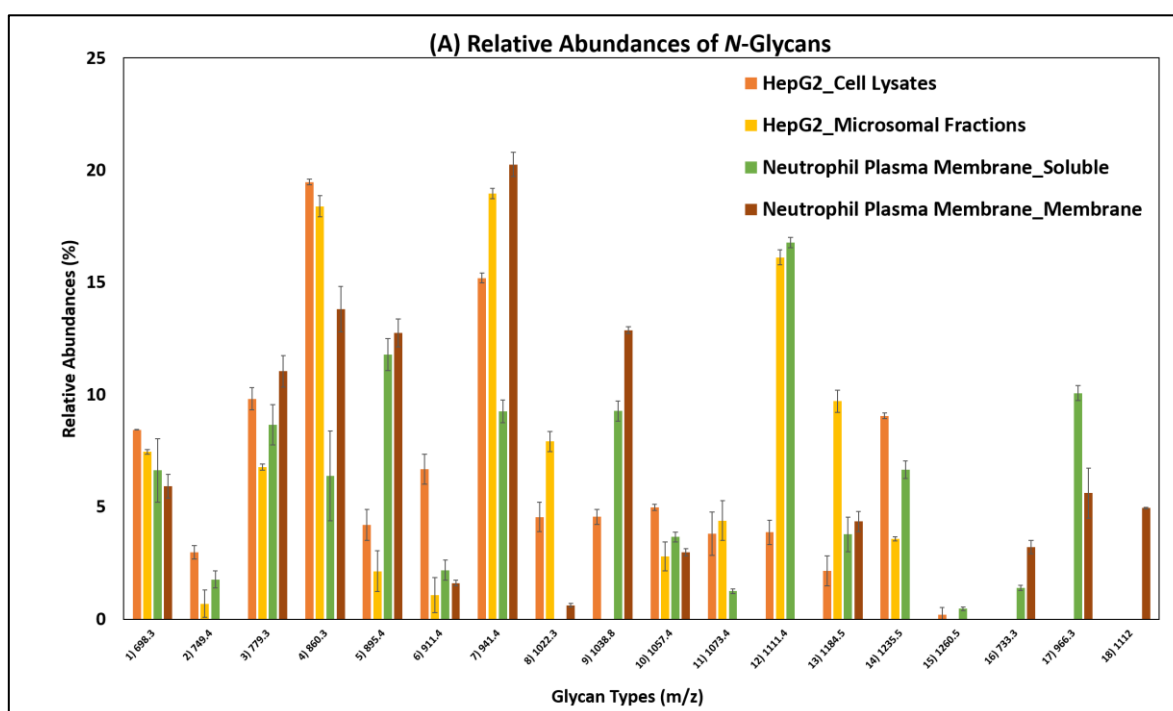


Figure S5: The relative distribution of each glycans present in each soluble or membrane fractions of the cell types are represented here. The quantitation of each glycan peak has been compared among the soluble fraction (orange) and membrane fraction (yellow) of HepG2 cells with soluble fraction (green) and membrane fraction (brown) of neutrophils.

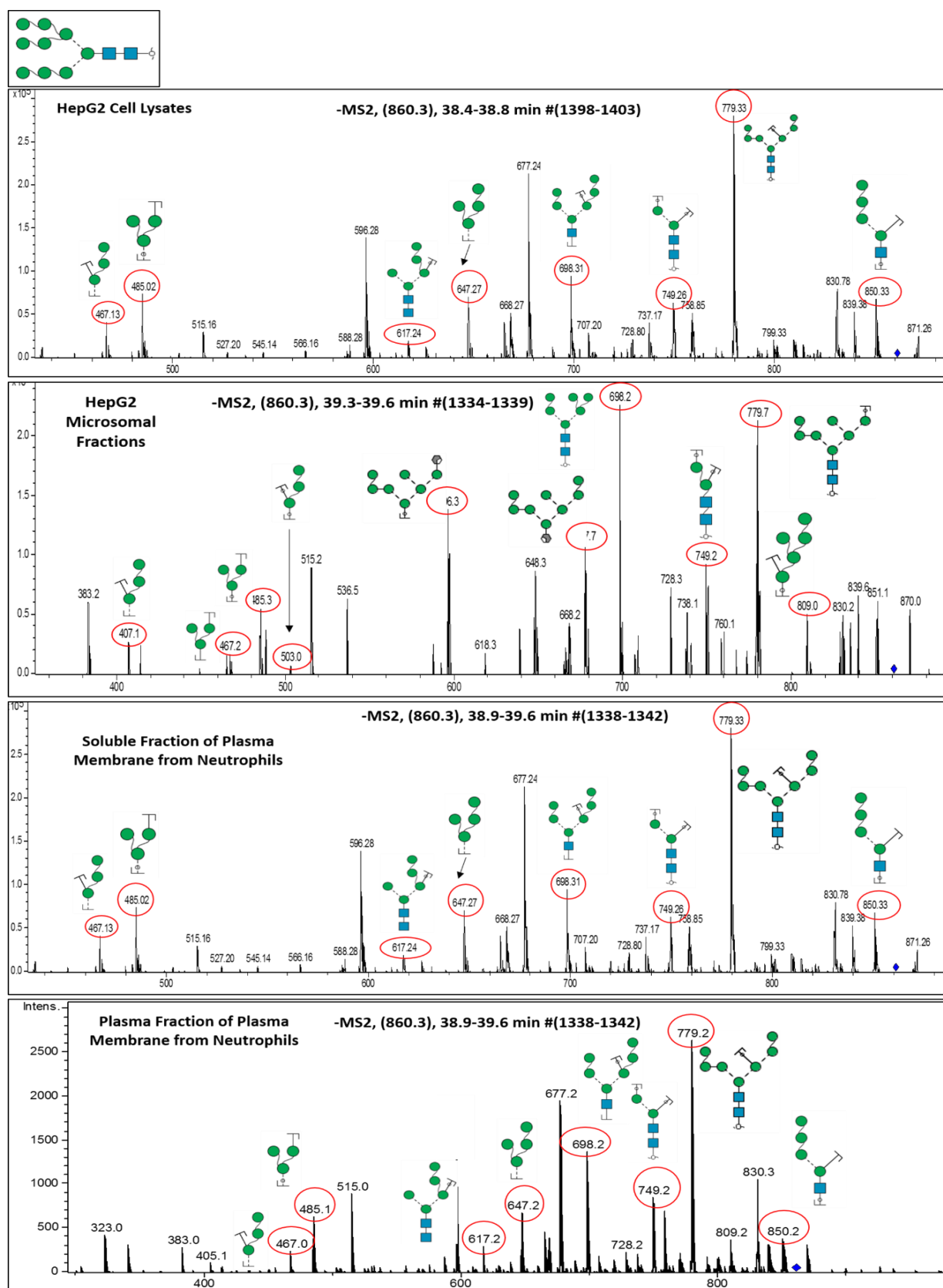


Figure S6: Manually annotated MS/MS of the common glycan #4, 860.3²⁻ (Table S4) among the four different samples; cell lysates and microsomal fractions from HepG2 liver cancer cells, and soluble and membrane fractions of the enriched plasma membrane obtained from resting human neutrophils.