Identification of tumour-associated gycoproteins recognised by the lectin from Helix pomatia in breast cancer.

Neela Devis Sing Rambaruth

School of Life Sciences

IDENTIFICATION OF TUMOUR-ASSOCIATED GLYCOPROTEINS RECOGNISED BY THE LECTIN FROM *HELIX POMATIA* IN BREAST CANCER

NEELA DEVI SING RAMBARUTH

A thesis submitted in partial fulfilment of the requirements of the University of Westminster for the degree of Doctor of Philosophy

June 2011

Author's declaration

I declare that the present work was carried out in accordance with the Guidelines and Regulations of the University of Westminster. The work is original except where indicated by special reference in the text.

The submission as a whole or part is not substantially the same as any that I previously or am currently making, whether in published or unpublished form, for a degree, diploma or similar qualification at any university or similar institution.

Until the outcome of the current application to the University of Westminster is known, the work will not be submitted for any such qualification at another university or similar institution.

Any views expressed in this work are those of the author and in no way represent those of the University of Westminster.

Signed **Date: 29th June 2011**

Abstract

Helix pomatia agglutinin (HPA) is a carbohydrate binding protein isolated from the Roman snail. There has been considerable interest in understanding HPA binding partners in cancer, as the lectin has been shown to identify glycosylation changes in cancers arising from the epithelia, from patients with poor prognosis. Identifying the HPA binding epitopes associated with a malignant phenotype may be useful for prognostication and may also offer potential as targets for immunotherapy. Previous studies have shown that HPA recognises a multitude of proteins in colorectal cancer (CRC). This study aimed to establish whether HPA recognises the same glycoproteins in breast cancer.

An *in vitro* model of human breast cancer cell lines was used, ranging from HPA negative, non metastatic, to HPA positive and metastatic. Four human breast cell lines were chosen to represent phenotypes ranging from 'normal'/benign (HMT3522), primary cancer (BT474) to metastatic cancer (T47D, MCF-7). HPA binding was assessed using confocal microscopy. Membrane proteins were extracted by differential centrifugation and the proteins were analysed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with Western blotting. The cell surface glycoproteins recognised by HPA were characterised using 2-dimensional electrophoresis (2-DE), Western blotting and mass spectrometry.

HPA binding correlated with integrin α6 levels, this concurred with previous findings in CRC. HPA also bound transcription factors HnRNP H1, HnRNP D-like, HnRNP A2/B1 as well as Hsp27, GFAP and ENO1. The recognition of HnRNPs, Hsp 27 and ENO1 by HPA correlated with O-GlcNAcylation of these proteins. Interestingly, these HPA-binding glycoproteins were either absent or showed decreased levels in the non-metastatic breast cancer cell line BT474 and in 'normal' HMT3522. A comprehensive analysis of the breast cells proteome showed a number of proteins with elevated level in the metastatic breast cancer cell lines T47D and MCF-7, but this is the first report to show elevated levels of elongation factor Tu, Enoyl Coenzyme A hydratase 1 peroxisomal and macropain subunits.

This work was extended to analyse the gene expression for UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase (ppGalNAc T1,T2, T3 and T6) and alpha 2,6 sialyltransferases (ST6GalNAc I and II) in the breast cell lines, but no correlation between the expression levels of mRNA of these enzymes and HPA binding was found in this study.

The results from the present study show that, as in CRC cell lines, integrin α6 was the most abundant HPA binding glycoprotein extracted from the breast cancer cells with a metastatic phenotype. This is the first report in which HPA has been shown to bind O-GlcNAcylated transcription factors. This class of proteins represent a new means by which HPA differentiates cancer cells with an aggressive metastatic phenotype. New approaches aimed at targeting these changes might have broad application for the treatment of breast, colorectal and possibly other epithelial cancers.

Dedication

To my late Dad who always believed in me

Publications arising from the work presented in this thesis

Rambaruth NDS, Greenwell P, Dwek MV. (2010). Characterisation of *Helix pomatia agglutinin* binding proteins in epithelial cancer of the breast. *Glycobiology* 20(11): 1471 (published abstract).

Rambaruth NDS, Dwek MV. Cell surface glycan–lectin interactions in tumor metastasis. *Acta Histochemica* (2011), doi:10.1016/j.acthis.2011.03.001

Rambaruth NDS, Greenwell P, Dwek MV. The lectin *Helix pomatia* agglutinin recognises O-GlcNAc containing glycoproteins in human breast cancer. *British Journal of Cancer*. (submitted).

Acknowledgements

I am deeply indebted to my supervisors, Dr. Miriam Dwek and Dr. Pamela Greenwell for giving me this unique opportunity to carry out this exciting research. Miriam, I am most grateful for all your guidance throughout the four years spent under your supervision and your precious support during bad times. Pamela, I am extremely thankful for all your advice and supports during this period.

I am very grateful to Dr. Anatoliy Markiv for inspiring me as a scientist and for giving me technical advice and teaching me bioinformatics throughout my PhD.

I highly appreciate the support of Dr. Hannah Lomax-Browne for giving me useful suggestions about thesis writing.

I am also indebted to Rozinder Gill for the staining shown on page 72, Tejeshwar Teji and Ewa Blazcack for assistance with the mRNA analysis. I thank my colleagues Dr. Daniel Gyamfi, Dr. Babak Afrough, Dr. Diluka Peiris, Omeco White, Yanitsa Nedelcheva and Simon Fry for their help and friendship.

I am extremely grateful to all the technical staff in the school of Biosciences for their great effort in providing me with all the reliable equipment. In particular I thank Thakor Tandel for his generous support.

I highly appreciate the support from the Cavendish Research Scholarship for giving me this golden opportunity to embark on a PhD.

A very special thank to Sailesh Gaungoo, for supporting and encouraging me throughout this long and often hard journey and still being there at the end. Thank you very much for believing in me.

Finally, I will like to thank my mum and sisters in Mauritius for encouraging me to successfully finish this research.

Abbreviations

1-DE	one dimensional electrophoresis
2-DE	two dimensional electrophoresis
aFGF	acidic fibroblast growth factor
AJ	adherens junction
bFGF	basic fibroblast growth factor
BM	basement membrane
BSA	bovine serum albumin
CA15.3	cancer antigen 15.3
CAM	cell adhesion molecule
CEA	carcinoembryonic antigen
CHAPS	3-[(cholamidopropyl) dimethylammonium]-1-propanesulfonate
CID	collision induced dissociation
CLR	C-type lectin receptor
CRC	colorectal cancer
CV	coefficient of variation
DAB	diamino benzidine
DBA	*Dolichos biflorus* agglutinin
DC-SIGN	denditritic cell specific intercellular adhesion molecule-3-grabbing non integrin
DMEM	Dulbecco's modified Eagle's medium
DNA	deoxyribonucleic acid
DTT	dithiothreitol
ECA	*Erythrina cristagalli* agglutinin
ECM	extracellular matrix
EDTA	ethylenediamine tetraacetic acid
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
ER	estrogen receptor
FITC	fluorescein isothiocyanate
FN	fibronectin
FUCT	fucosyltransferases
GAG	glycosaminoglycan

Gal	galactose
GalNAc	N-acetyl-galactosamine
Glc	glucose
GlcNAc	N-acetyl glucosamine
GnT	N-acetylglucosaminyltransferase
GnT III	UDP-GlcNAc α-mannoside β1-6-N-acetylglucosaminyltransferase III
GnT V	UDP-GlcNAc α-mannoside β1-6-N-acetylglucosaminyltransferase III
Gp120	glycoprotein 120
GpI	glycosylphatidylinositol
GTases	glycosyltransferases
HA	hyaluronic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER 2	human epidermal growth factor receptor 2
HIV	human immunodefiency virus 1
HPA	*Helix pomatia* agglutinin
HRP	horseradish peroxidase
HS	heparan sulphate
Hyl	hydroxylysine
Hyp	hydroxyproline
IGF	insulin-like growth factor
IPG	immobilised pH gradient
KDa	kilo Dalton
LOH	loss of heterozygosity
L-PHA	L-phytohaemagglutinin
MALDI/TOF	matrix assisted laser-desorption ionisation/time of flight mass spectrometer
Man	mannose
MMPs	matrix metalloproteinases
MS	mass spectrometry
MUC 1	mucin 1
MW	molecular weight
N/A	not applicable
OGA	beta-N-acetylglucosaminidase
OGT	O-linked N-acetylglucosaminyl transferase

OST	oligosaccharyltransferase
PBS	phosphate buffered saline
PDGF	platelet-derived growth factor
pI	isoelectric point
PNA	peanut agglutinin
PSA	prostate specific antigen
Pro	proline
Rb	retinoblastoma
SA	sialic acid
SBA	soybean agglutinin
SCID	severe combined immunodeficient
SD	standard deviation (of the mean average)
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Ser	serin
SNA	*Sambucus nigra* agglutinnin
SLe^A	sialyl Lewis A
SLe^X	sialyl Lewis X
SRCL	scavenger receptor-C-type lectin (SRCL)
TBS	tris-buffered saline
TGFα/β	transforming growth factor α/β
Thr	threonine
Tyr	tyrosine
TRITC	tetranethylrhodamine isothiocyanate
TWEEN	polyoxyethylene(20) sorbitan monolaurate
UEA-1	*Ulex europeaus* agglutinin 1
uPA	urokinase plasminogen activator
VEGF	vascular endothelial growth factor
WGA	wheat germ agglutinin

Contents

List of figures

List of tables

Chapter 1

Introduction

1.0 Introduction

1.1 Breast cancer: the clinical problem

Breast cancer is the most prevalent form of cancer and is the leading cause of cancer related deaths in the Western world (McPherson *et al*., 2000). In the UK, around 46,000 new cases of breast cancer are diagnosed each year (Cancer Research UK, 2008). Breast cancer occurs mainly in women rather than in men who account for less than 1% of all the cases (Cancer Research UK, 2008). It is estimated that breast cancer causes more than 12,300 deaths each year in the UK (Cancer Research UK, 2008). Althrough breast cancer mortality rates have fallen dramatically since 1989, largely as a result of widespread use of hormonal and biological treatments, most patients succumb to the disease due to metastasis formation (Chambers *et al*. 2002; Steeg 2006). The ability of breast cancer cells to detach mechanically from the primary tumour mass and to migrate to distant organs remains a major clinical problem. To date a congruent strategy for metastasis prevention and treatment has remained elusive.

1.2 Female breast anatomy

The development of the mammary gland is unique as it mostly occurs postnatally, reaching full development during pregnancy and correlating with the main function of the breast to produce and deliver milk to the newborn (Lanigan *et al*., 2007). In humans, the development of the breast begins prenatally during the 4^{th} week of gestation (Dawson, 1934). During embryogenesis, a poorly branched primitive duct system develops (Howard and Gusterso, 2000), here maternal hormones cross the placenta and induce the development of the ductal system. In the foetus the major development however occurs just preceding and during puberty (Russo *et al*., 1987; Monaghan *et al*., 1990). There is no difference in the development of the male and female breast between birth and just before puberty, the major ductal system develops in the early pubertal transition approximately 2-3 years prior to menarche (Juul *et al*., 2006). Oestrogen, progesterone and prolactin together with local growth factors stimulate the development of the primitive ductal scaffold into

a network of ducts and secretory lobules this occurs along with marked stromal expansion (Hennighausen and Robinson, 2005; Hens and Wysolmerski, 2005; Howard and Gusterso, 2000; Monaghan *et al*., 1990).

In its development, the mammary gland resembles a modified sweat gland functioning to synthesise milk for the newborn (Lanigan *et al*., 2007; Tortora & Grabowski, 1993). The mature human breast is composed of a branching duct system of 15 to 20 lobes separated by adipose tissue (Bannister, 1995). There is a preponderance of glandular tissue in the upper outer section of the breast resulting in tenderness in many women prior to their menstrual cycle. Each lobe comprises smaller sections called lobules usually 1-2 millimetres in diameter and embedded in the surrounding stroma and fat (Park, 1959; Tortora & Grabowski, 1993). The lobule is a complex system of ducts called ductules which differentiate into secretory units known as alveoli or acini (Park, 1959). The structure of an adult human female breast is shown in figure 1.1.

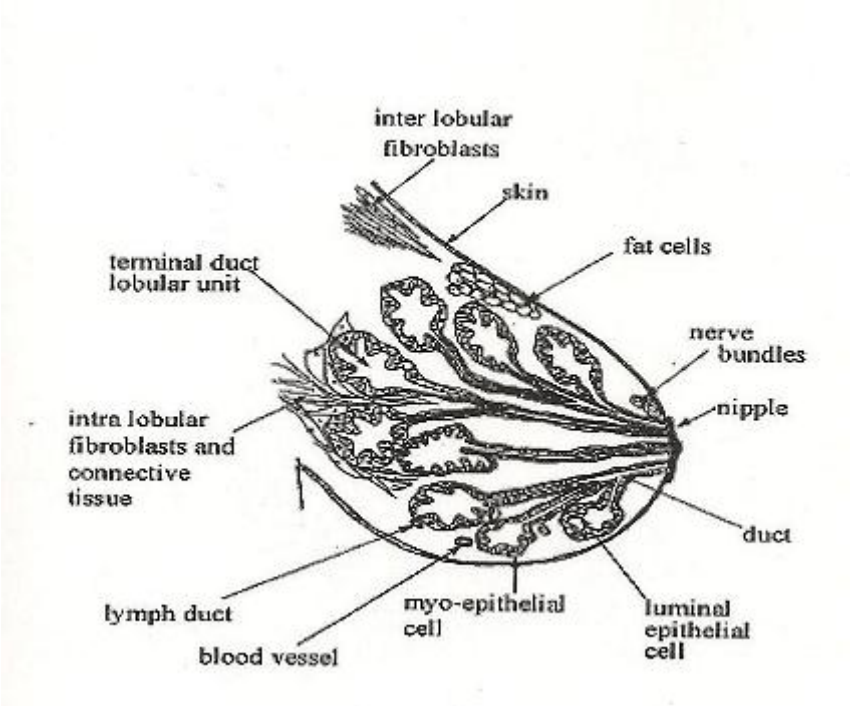


Figure 1.1: Structure of an adult human breast. Adapted from Dwek (1998).

The alveoli drain into small ducts that merge to form larger ducts and function to convey milk from the ductules. These larger ducts converge forming one milk duct for each lobe, and as the main duct approaches the nipple it enlarges to form a lactiferous sinus (Bannister *et al*. 1995; Vorherr, 1974), the sinus narrows into the lactiferous duct which terminates in the nipple and exits via the secretory pores. The nipple is a small hyperpigmented projection on the anterior surface which is innervated by the fourth intercostal nerves. The pigmented area surrounding the nipple is the areolar (Tortora & Grabowski, 1993). Breast cancers can arise in any part of the organ, althrough the majority of tumours have been reported in the upper outer quadrant of the breast (Douek *et al*, 1999).

The main blood supply to the breast is from the internal mammary artery and the lateral thoracic artery (Cunningham, 1977). The posterior intercostal arteries and the pectoral branch of the thoracoacromial artery are smaller sources of blood supply to the breast (Bannister, 1995). The venous system of the breast is composed of deep and superficial systems linked by short connecting veins (Cunningham, 1977). The main venous drainage of the breast is via the axillary vein, while the internal thoracic vein and the intercostal veins also contribute to some venous drainage. The lymphatic system of the breast has been extensively documented with particular emphasis on its important role in the metastasis of breast cancer cells. The axillary nodes and the internal mammary nodes are the two main systems for lymphatic drainage of the breast. The axillary nodes drain the lymph from both the medial and the lateral portions of the breast whilst the lymph drainage from the deep areas of the breast is accomplished by the internal mammary nodes. Other small networks of the lymphatic systems have also been documented (Bannister, 1995).

1.3 Causative factors in breast cancer

Breast cancer development is unpredictable, its evolution may take between 5 to 30 years, or it may appear suddenly and progress rapidly (Holt *et al*., 1993). A single change or transforming event in the genetic pathways (mainly involving susceptibility genes) in breast cells may initiate the formation of breast cancer (reviewed in Polyak, 2007). The subsequent progression of the tumour is accompanied by cellular immortalisation, clonal expansion and selection (reviewed in Baum & Schipper, 2005, reviewed in Polyak, 2007). Approximately 90% of breast cancers occur sporadically, in this case women have no family history and very little is known about the specific factors giving rise to these cancers (Polyak, 2001). Only approximately 10% of all breast cancers arise from inherited mutations in tumour susceptibility genes. Initially, mutation in the *p53* suppressor gene was throught to be important (Malkin *et al*., 1990), but since then, mutations in other genes such as *BRCA1* (King, 1992), *BRCA2* (Wooster *et al*., 1995) and *BRCA3* (Thompson *et al*., 2003) have been associated with familial breast cancer. Other genetic changes linked to breast cancer development are the over-expression of *myc* and (*HER2/neu*) oncogenes and suppression of *p53* and (*Rb)* tumour suppressor genes (Couldrey & Green, 2000).

Other recognised risk factors associated with the development of breast cancer include early menarche, late menopause, low parity, environmental conditions (for example exposure to radiation and carcinogens) and relatively late age at first full time pregnancy (MacMahon *et al*., 1973; Shapiro *et al*., 1989). Moreover, of the many candidates of endogenous hormones, oestrogen and progesterone have also been implicated in breast cancer (Key & Pike, 1988; Thomas *et al*., 1997; Toniolo *et al*., 1995). A number of population based studies have also reported that sources of exogenous hormones such as the prolonged use of the oral contraceptive pill correlate with an increased risk of developing breast cancer (reviewed in Harris & Hellman, 1996). It has also been demonstrated that certain formulations of hormone replacement therapy taken for prolonged periods of time increase the likelihood of developing the disease (Faiz & Fentiman, 1998).

Of the many types of breast diseases, atypical hyperplasia and recurrent or cystic disease have been correlated with increased risk of breast cancer (Bodian, 1993; Dupont *et al*., 1993). It has also been shown that patients with diabetes mellitus are at greater risk of developing breast and endometrial cancer (Weiderpass *et al*., 1997), this may be due to increased levels of insulin-like growth factor in the breast tissues of these patients (Stoll, 1997).

1.4 Breast cancer staging

The majority of breast cancers originate from the terminal duct lobular unit (Ronnov-Jessen *et al*., 1996). Benign (non-malignant) tumours mainly arise from the myoepithelia or fibroblast cell populations, whilst invasive carcinomas arise mostly from the ductal luminal or lobular cells (Taylor-Papadimitriou & Lane, 1987; Lakhani *et al*., 1999). Cancers arising from ducts are classified as ductal carcinomas and those arising from lobules are classified as lobular carcinomas.

There are two common types of breast carcinoma recognised in symptomatic individuals, these are; carcinoma *in situ* (lobular carcinoma *in situ* and ductal carcinoma *in situ*) and invasive carcinoma (invasive lobular carcinoma and invasive ductal carcinoma). The invasive ductal carcinoma can further be classified, depending on their histological subtypes, as mucinous, tubular, medullary and metaplastic invasive carcinomas (Millis, 1984; Gallager, 1984). Invasive ductal and lobular carcinomas are often grouped together as patient outcome is similar.

Initially, it was throught that breast cancer spread centrifugally from the primary tumour, in a local manner, first throughout the breast tissue, into the lymphatics, then to the regional axillary lymph nodes and later to distant organs. It was believed that surgical removal of the affected breast and axilla would result in a cure (Halstead, 1898; Haagensen, 1986). However, amongst others, Brinkley and Haybittle, (1975) reported that breast cancer patients die steadily from metastatic disease over 20 years following their initial surgery. From this it became apparent

that systems to ‘stage’ or ‘group’ the tumours according to their likely biological outcome were needed.

Human breast cancer tissues are histologically complex, consisting of a multitude of different cell types amongst the carcinoma cells (Ronnov-Jessen *et al*, 1990) and mammary gland tumourigenesis involves a sequence of marked clinical and pathological stages (Polyak, 2001). Atypical proliferation is generally throught to precede the formation of *in situ* carcinoma which then develops into invasive carcinoma, often resulting in metastatic disease (Beckmann *et al*, 1997). When a patient is diagnosed with breast cancer, they are classified by the extent to which the disease has spread and are grouped with patients who have similar outcomes.

This classification, or staging, is essential to determine the prognosis (probable cause and outcome of the disease) and the appropriate treatment for the patient (Brower *et al*., 1999; Baum & Schipper, 2005). In order to ensure accuracy, consistency and to allow for comparison of data between different centres a single system is used for breast cancer staging (Brower *et al*., 1999; Baum and Schipper., 2005). Table 1.1 summarises the TMN classification system. The TNM staging system was refined by the American Joint Committee on Cancer (AJCC) and the International Union against Cancer (UICC), in the 1980s, allowing further classification of tumours categorised by the TNM system, into one of the four stages (Brower *et al*., 1999; Weiss, 2000). T refers to primary tumour, N refers to regional lymph nodes, and M refers to distant metastasis (Weiss, 2000).

Table 1.1: TNM classification system. Adapted from Baum and Schipper (2005).

Tumour stage	**Description**
Tx	Tumours cannot be assessed
T0	No evidence of Tumour
Tis	Carcinoma *in situ*.
T1	Tumour is not larger than 20 mm in diameter
T1a	Tumour is more than 1 mm but not more than 5 mm in diameter
T1b	Tumour is more than 5 mm but is not larger than 10 mm
T1c	Tumour is more than 10 mm but is not more than 20 mm
T2	Tumour is more than 20 mm but is not more than 50 mm
T3	Tumour is more than 50 mm in diameter
T4a	Tumour of any size and the tumour has spread into the chest wall
T4b	Tumour has spread into the skin (Ulceration; Ipsilateral satellite nodules; Edema)
T4c	Both T4a and T4b (Tumour has spread to both skin and chest wall)
T4d	Inflammatory carcinoma
N Stage	Features
NX	Lymph node cannot be assessed (e.g if previously removed)
N0	No cancer cells found in lymph nodes
N1	Metastases in ipsilateral axillary nodes (no fixation)
N2	Metastases in ipsilateral axillary nodes with fixation
N3	Metastases detected in ipsilateral internal mammary lymph nodes
M stage	Features
M0	No distant metastases detected
M1	Distant detachable metastases detected

The staging systems seek to serve as a means of identification of the likely patient outcome prognosis. A number of systems, involving the identification of prognostic markers, were previously explored but the system most widely used is the TNM system described above (Bloom & Richardson, 1957). The Nottingham prognostic index which includes age, tumour size, tumour grade, lymph node stage, and estrogen receptor (ER) content, maybe useful with other biological factors (Galea *et al*., 1991). There is further need for improved prognostic markers as well as new targets for breast cancer therapy.

1.5 Prognostic indicators

Breast cancer staging system also incorporates the identification and use of several prognostic indicators such as nodal status, tumour size, histological grade, hormone receptor status, oncogene over-expression, cancer cell proliferation and cell surface glycoproteins. Another prognostic factor discussed in this thesis is the *Helix pomatia* binding status. These factors are discussed below.

1.5.1 Nodal status

Detection of metastases in the axillary lymph nodes is one of the most accurate prognostic factors in breast cancer. Nodal status is part of the TNM classification system. The most accurate approach used for assessing axillary lymph node involvement is to surgically remove the lymph node. This is an invasive procedure and may lead to physical impairment such as lymphoedema, nerve damage and paresthesias. A less invasive procedure commonly used is the sentinel lymph biopsy. This technique involves the excision and examination of the the first lymph node from the breast drains through, for the presence of metastases (Hammer *et al*., 2008; Krishnamurthy, 2005).

1.5.2 Tumour size

Tumour size, after lymph node status, is the next most accurate indicator of breast cancer. Tumour size is also part of the TNM classification system. Prognosis becomes poorer with increase in tumour size. Patients with tumours of size ≤ 1cm have a longer survival rate (about 79%) compared to patient with larger tumours of size 2-5 cm (about 66%) (Hutson & Osborne, 2005).

1.5.3 Histological grade

Histological grading provides a measure of tumour differentiation and correlates with prognosis and survival rates. The first histological grading system for breast cancer was introduced by Greenough in 1925 and carcinomas were classified based on six histological characterisitics, including, adenomatous arrangement, secretory activity, hyperchromatism, number of mitoses and variations in cell size and shape. This classification system was further simplified in 1957 by Bloom and Richardson. Bloom and Richardson classification system was based on histological characteristics such as tubule formation, variation in nuclear size and shape (pleomorphism) and hyperchromatism. This system is most widely used in the UK to effectively predict patient prognosis and is called the Nottingham combined histological grading system (Elston & Ellis, 1991).

1.5.4 Hormone receptor status

Oestrogen receptor (ER) and progesterone (PR) status are not included in the staging process, but are routinely used to allow a more tailored treatment plan for patients. Approximately 50-80% of breast tumours are ER and PR positive (Elledge & Osborne 1997). ER positive breast tumours predict a longer disease-free survival rate (66%) compared to ER negative tumours (56%). ER positive patient respond well to endocrine therapies such as oophrectomy and luteinising hormone-releasing hormone (LHRH) analogues. These approaches reduce the levels of circulating oestrogen in the blood. Tamoxifen is also commonly used to

block the interaction between circulating oestrogen and the ER receptors (Cheung, 2007).

1.5.5 Oncogene Over-expression

Over-expression of *HER2* gene is reported in approximately 30% of breast cancers (Salmon *et al*., 1987). This gene encodes for the HER2/neu protein which is a 185 kDa tyrosine kinase receptor (Coussens *et al*., 1985). Over-expression of this oncogene is associated with shorter disease free survival rate (Slamon *et al*., 1987). Determination of the HER2 status of a patient is also of importance in developing appropriate treatments plans, for instance, HER2 positive patients are treated with Herceptin, a monoclonal antibody which slows the growth of HER-2 positive tumours (Vogel *et al*., 2002).

1.5.6 Cell proliferation

Tumour cell proliferation rate is also a valuable prognostic factor commonly used in breast cancer diagnosis. Two markers of cellular proliferation commonly employed are detailed below.

1.5.6.1 S-phase Fraction (SPF)

S-phase fraction (SPF) is a measure of the number of cells in the S-phase i.e undergoing active DNA synthesis/replication. High SPF has been shown to correlate with early disease recurrence (Wenger *et al*., 1993).

1.5.6.1 Nuclear Antigen Ki-67

The nuclear protein Ki-67 is expressed during all phases of the cell cycle except G0 phase (Gerdes *et al.,* 1991). A positive correlation exists between increase Ki-67 levels and decreased survival rate (Sullivan *et al*., 1993).

1.5.7 *Helix pomatia* agglutinin (HPA) binding status

Helix pomatia agglutinin (HPA) binding is a significant prognostic indicator of breast and other cancers. This is further explored in section 1.8.

1.5.8 Cell surface glycoproteins

Cell surface proteins have been subject of much research interest in breast cancer (Jacobson *et al.,* 1996). The post-translational modification of proteins, glycosylation and phosphorylation, in particular, have been associated with the complex transformation of malignancy. Some of these changes have been demonstrated to provide a selective advantage for tumour cells to progress to a metastatic stage (Krueger & Srivastava, 2009) and many of these changes may also offer potential for monitoring disease, for instance, glycoproteins such as carcinoembryonic antigen (CEA) and CA125 are commonly used as tumour markers for monitoring breast cancer patients' response to treatment (reviewed in Duffy, 2006). Whilst, these markers are of some utility for monitoring patient response to treatment, they have limited relevance as diagnostic tools because of their low specificity and sensitivity, two factors that are important for tumour markers (Pannell & Kotasek., 1997). Identification of cell surface proteins, with altered post-translational modifications may be useful for cancer diagnosis, prediction and prognostication.

1.6 Glycosylation

Glycosylation is the most frequent post-translational modification of proteins in eukaryotes. It is a complex modification which involves the enzymatic addition of carbohydrates to proteins (Lis & Sharon, 1993, Varki *et al*, 1999). Glycosylation plays a key role in mediating cellular functions such as ensuring correct protein folding (Walsh *et al*., 1990), cell to cell adhesion (Lasky, 1992; Springer, 1990), cell to cell communication (Wassarman, 1990), protection of proteins from enzymatic degradation (Homans *et al*., 1987) and signal transduction (Haltiwanger & Stanley, 2002). In mammalian cells, glycans attached to proteins or lipids consist of either linear or branched oligosaccharide chains are formed by the enzymatic addition of a combination of seven different monosaccharides consisting of either glucose, mannose, galactose, fucose, N-acetylgalactosamine, N-acetylglucosamine and/or sialic acids (the seven monosaccharides are shown in figure 1.2), giving rise to glycoproteins, glycosaminoglycans, proteoglycans and glycolipids. The oligosaccharide chains link to proteins in two main types of ways: (a) via glycosidic bonds between N-acetylgalactosamine (GalNAc) to the hydroxyl group of theonine or serine on the polypeptide chain, giving rise to O-linked glycans which are predominantly membrane bound or secreted on mucins, (b) via glycosidic bonds between N-acetylglucosamine (GlcNAc) to asparagine occurring on the sequon Asn-X-Ser/Thr (where X can be any protein except proline) of the protein chain, (Opdenakker *et al.*, 1993). Proteoglycans, (such as heparan sulphate or chondroitin sulphate) and glycosaminoglycans (such as hyaluronan) exist as free glyconjugates and form the extended components of the ECM. Oligosaccharides glycosidically linked to lipids include glycosylphatidylinositol (GPI) anchored proteins which exist as membrane components (Dwek & Brooks, 2004; Brooks *et al.*, 2008).

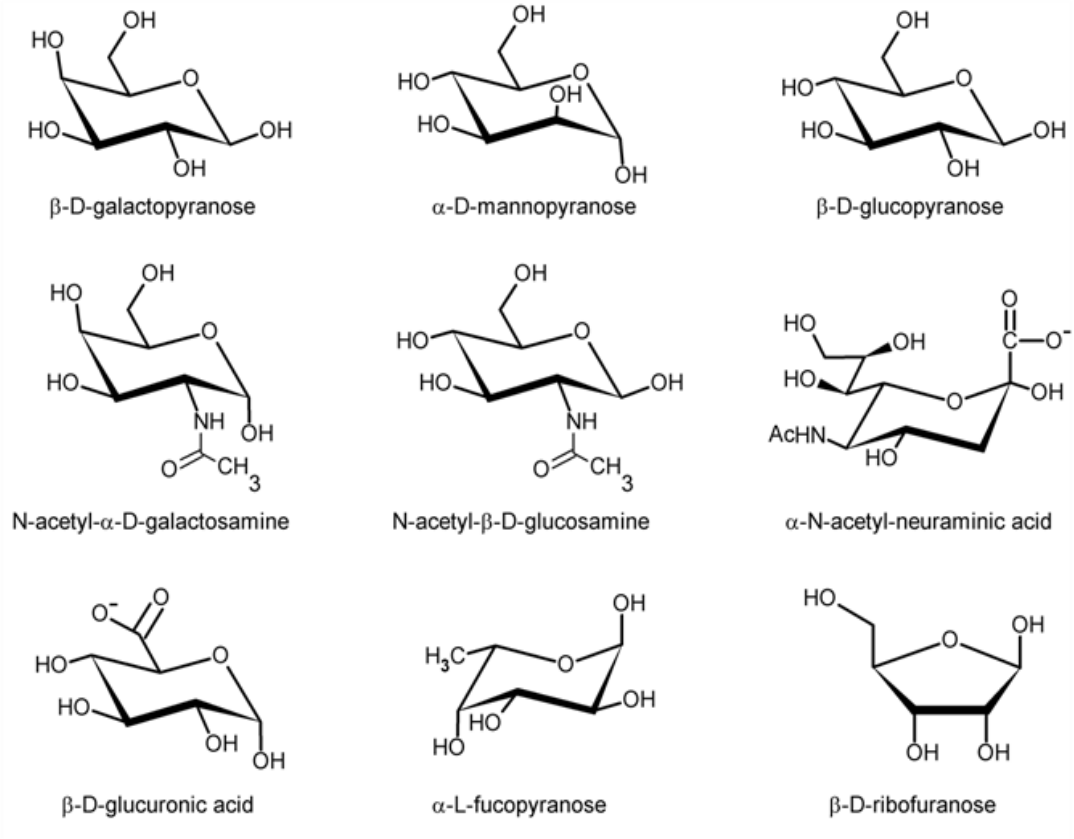


Figure 1.2: The seven monosaccharides associated with human glycoproteins and glycolipids. Adapted from Afrough (2009).

1.6.1 Enzymes of glycosylation

Glycans are not coded directly by the genome but are synthesised by enzymes which are themselves coded by genes. It has been speculated that one percent of the genome codes for the enzymes of glycosylation (Lowe & Marth, 2003) but it is also known that not all the enzymes involved in glycosylation reactions have been identified (De Graffenried & Bertozzi, 2004). Glycosylation is a complex process involving many enzymes located mainly in the endoplasmic reticulum (ER) and in the Golgi apparatus (De Grafenried & Bertozzi, 2004), and the glycosylation process is orchestrated by glycosidase and glycosyltransferase enzymes. Oligosaccharides or glycans consist of chains of monosaccharides linked by

glycosidic bonds and the synthesis of glycans involves the catalytic addition and removal of monosaccharide units to the growing chain by the action of glycosyltransferases and glycosidases respectively.

1.6.2 N-linked glycosylation

N–linked glycosylation is a co-translational process occurring mainly in eukaryotes, and rarely in bacteria (Langdon *et al*., 2009; Magidovich & Eichler, 2009; Spiro, 2002; Szymanski & Wren, 2005). N-linked protein glycosylation takes place in several distinct steps, the first of which is the covalent attachment of a pre-formed lipid-linked oligosaccharide chain to asparagine of the Asn-X-Ser/Thr sequence (where X can be any amino acid except proline) of a nascent polypeptide, this occurs in the lumen of the endoplasmic reticulum (Jones *et al*., 2005). In some cases the tripepetide sequon Asn-X-Cys may also be glycosylated with N-linked glycans (Imperiali & Hendrickson, 1995). It has been reported that due to protein folding and the accessibility of glycotransferases to the sequon, only about 70% of tripeptide Asn-X-Ser/Thr sequences in eukaryotic proteins are successfully glycosylated (Petrescu *et al*., 2004). The steps in *N*-linked glycosylation are illustrated in figure 1.3.

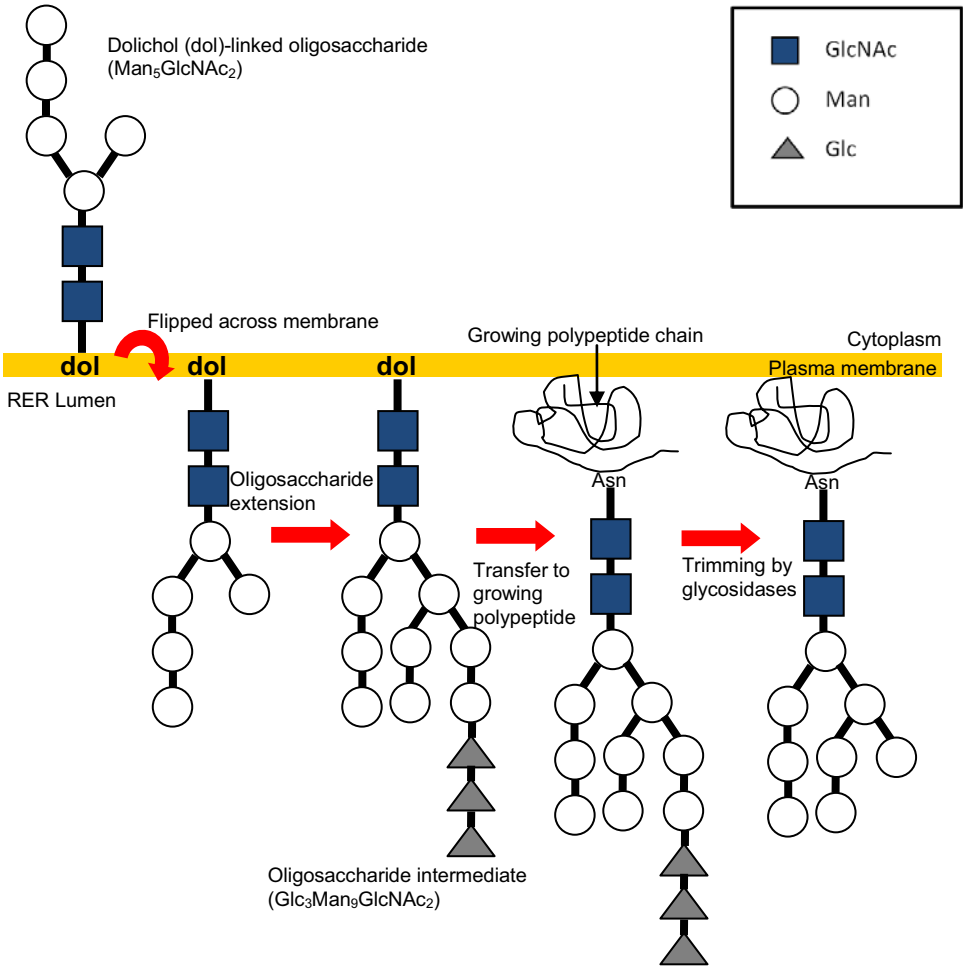


Figure 1.3 The steps involved in N-linked glycosylation. Adapted from Lomax-Browne (2009). N-linked oligosaccharides synthesis starts with the attachment of two GalNAc residues to a dolichol phosphate molecule. This is followed by the subsequent attachement of five mannose residue to form the $Man_5GlcNAc_2$ molecule. This molecule is further enlarged in the lumen of the ER to form a final dolichol-linked oligosaccharide intermediate ($Glc_3Man_9GlcNAc_2$).

N-linked biosynthesis involves the formation of the intermediate oligosaccharide structure on a dolichol-pyrophosphate molecule, this occurs by the sequential enzymatic activity of glycosyltransferases. There is initial attachment of a GlcNAc residue to the dolichol phosphate molecule resulting in the formation of Dol-PP-GlcNAc, this enzymatic process is mediated by Alg7, an N-acetylglucosamine-phosphate transferase (Kukuruzinska and Robbins, 1987). The Dol-PP-GlcNAc is further extended by the addition of another GlcNAc residue catalysed by an N-acetylglucosaminyltransferase such Alg 13/14 (Bickel *et al*., 2005; Chantret *et al*., 2005; Goa *et al.,* 2004). The resulting Dol-PP-GlcNAc-GlcNAc is further extended by the addition of five mannose residues in a process involving five mannosyltransferases to form the $Man_5GlcNAc_2$-PP-Dol structure (Brooks *et al*., 2002). The heptasaccharide is then flipped into the lumen of the ER by the action of Rft1p, a membrane-spanning protein (Helenius *et al*., 2002). The remaining step in the formation of the oligosaccharide occurs in the lumen of the ER (Weerapana and Imperiali, 2006). The completion of the dolichoyl-pyrophosphate decatetrasaccharide (Glu_3-$Man_9GlcNAc_2$-PP-Dol) involves the addition of four mannose and three glucose residues catalysed by a further four mannosyltransferases and three glucosyltransferases (Weerapana and Imperiali, 2006). Once the decatetrasaccharide assembly is completed, the enzyme complex oligosaccharyltransferase (OST) catalyses the transfer of the oligosaccharide from the lipid-linked oligosaccharide onto the nascent protein chain (Kelleher & Gilmore, 2006). The dolichol phosphate is regenerated and is re-exposed on the outer membrane of the ER (Rush *et al*., 2008; Schenk *et al*., 2001). After transfer of the oligosaccharide, protein folding is controlled by the calnexin/calreticulin cycle in eukaryotic cells, thus preventing the release of misfolded proteins (Bedard *et al*., 2005; Jaeken *et al*., 1993). If the protein is properly folded glucosidase I and II removes the three glucose residues. The fully folded protein is then cleaved by mannosidase resulting in a Man_8-$GlcNAc_2$ containing glycoprotein which is transported to the Golgi apparatus where the glycans are either elongated or trimmed by glycosidases and glycosyltransferases (Jaeken *et al.,* 1993).

The addition of monosaccharides including galactose, N-acetylglucosamine, fucose and sialic acid results in the elongation of the glycans and the final structures Man8GlcNAc2 can be further trimmed or extended to form a wide variety if N-linked glycans, all sharing common trimannosyl core (Man3GlcNAc2). Three main classes of N-linked glycans share this core: (1) high mannose type, which consists of a total of between 5 and 9 mannose residues attached to the inner GalNAc of the trimmannosyl core (2) complex type, which do not contain any mannose residues except from those already in the core structure and consists of repeated oligosaccharides branches in particular GlcNAc (β1-4)Gal branches. (3) hybrid type, which comprises features from both high mannose and complex type oligosaccharides (Brooks *et al*., 2002).

1.6.3 O-linked glycosylation of mucin-type glycoproteins

O-linked glycosylation is a complex post-translational event that is initiated in the Golgi apparatus (Kellokumpu *et al*., 2002). It involves the addition of sugar residues to amino acids carrying a hydroxyl functional group (Ser, Thr, Tyr, Hyp (hydroxyproline) and Hyl (hydrolysine) (Spiro, 2002). The O-linked glycans are branched structures comprising N-acetylgalactosamine, N-acetylglucosamine, galactose, fucose and sialic acid residues, the synthesis of which is controlled by a series of membrane bound glycotransferases, glycosidases (Van Den Steen *et al.,* 1998).O-linked glycosylation is initiated by the attachment of a GalNAc monosaccharide to a Ser or Thr on a polypeptide chain by the action of an N-acetylgalactosaminyltransferase. This structure is termed as the Tn antigen. The extension of the Tn antigen leads to the formation of eight basic core structures (as detailed in figure 1.4). There are seven different cores associated with O-linked glycans present in humans (Wopereis *et al*., 2006). The GalNAc-α-Ser/Thr linkage that occurs in mucins with N-acetylgalactosamine at the reducing end prevails in eukaryotes (Spiro, 2002). Many studies in eukaryotic cells have shown that O-linked glycan synthesis in eukaryotes involves at least fifteen GalNAc transferases (Clausen & Bennett, 1996, Ten Hagen *et al*., 2001). O-linked mucin core structures, core 1-6 and core 8, have been reported in humans (Brockausen,

2006). The core structures may be modified further resulting in several hundreds of different mucin-type O-linked glycan structures (Brockausen, 2006).

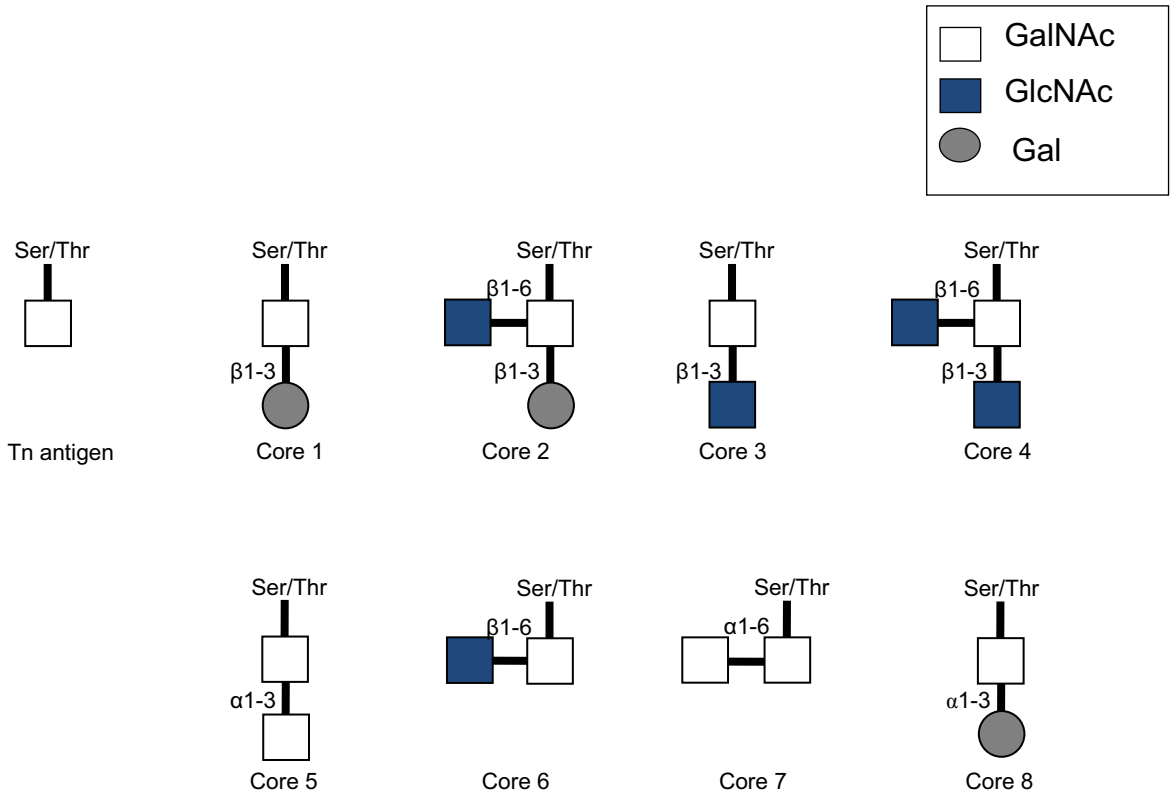


Figure 1.4: **The Tn antigen and the eight core structures that arise from Tn antigen extension**. Adapted from Lomax-Browne (2009).

1.6.4 O-GlcNAc glycosylation of cytoplasmic and nuclear proteins

O-GlcNAcylation involves the addition of a β-O-GlcNAc monosaccharide on a Ser and/or Thr amino acid on a polypeptide chain. This is a common form of glycosylation amongst cytoplasmic and nuclear proteins and has been shown to be an alternative pathway to phosphorylation, functioning in intracellular signalling processes (Slawson *et al*., 2008). Aberrant O-GlcNAcylation is a notable feature associated with chonic diseases such as diabetes (Copeland *et al*., 2008; Akimoto *et al*., 2005), cardiovascular disease (Laczy *et al*., 2009; Jones *et al*, 2008), neurodegenerative disorders (Lazarus *et al*; 2009; Lefebvre *et al*., 2005) and cancer (Chou *et al*., 2001).

1.6.5 Changes in glycosylation in cancer

The implications of aberrant glycosylation in the development of a malignant phenotype and tumour progression have been well documented in several studies (reviewed in Miyamoto, 2006). Alterations in glycosylation have been reported in cancer (Varki *et al*., 1999); (1) increased branching of N-linked oligosaccharides, (2) exposure of Lewis antigens, (3) synthesis of truncated O-glycans and (4) alteration in the sialylation of glycans (Dwek & Brooks, 2004). It has been shown that aberrant glycosylation is associated with tumour dissemination during the process of metastastasis (Brockhausen, 1999).

1.6.5.1 Increased in N-linked glycan branching

Increased β1-6 branched N-linked glycans on cell surface proteins have been established in numerous studies particularly in breast and colon cancer (Fernandes, 1991; Dennis *et al*., 1987; Korczak *et al*., 1994). Increased β1-6 GlcNAc branching occurs as a result of an increased expression of GlcNAc transferase V (GNT-V). The transcription of the GlcNAc transferase V gene (MGAT5) is upregulated in cancerous cells (Chen *et al*., 2006) and cells with an enhanced expression of GNT-V exhibit a metastatic phenotype, at least in animal

models of metastasis (Korczak *et al*., 2000). This is further described in section 1.7.1.1.

1.6.5.2 Alteration in Lewis antigens

Aberrant synthesis of Lewis antigens is a common feature of epithelial malignancies (Yu *et al*., 2005). Lewis antigens are formed by the catalysis of a fucose residue on core 1 or core 2 chains on O-linked glycans and are also found on N-linked glycans. This catalysis is mediated by the action of fucosyltransferases (FUCT I and VII). Increase synthesis of Lewis x (Le^x) and Lewis a (Le^a) as well as their respective sialylated structures, sialyl Lewis x (SLe^x) and sialyl Lewis a (SLe^a) have been shown in cancer cells (Varki *et al*., 1999) and their levels correlate with tumour progression, metastatic potential and poor prognosis in humans (Varki *et al*., 1999). This is futher discussed in section 1.7.4.1. It has been reported that α1,2 fucosyltransferase is elevated in colorectal cancer (Xiong *et al*., 2003) and the increase in this enzyme activity results in the synthesis of $Lewis^b$, Lewis ^{y} and formation of the ABH antigen (LaRue *et al*., 1997). Another mechanism that may lead to increased Lewis antigen synthesis on epithelial mucosa is through the downregulation of the ABH blood group transferases that result in blood group H or $Lewis^b$, $Lewis^y$ carbohydrate moieties with a GalNAc or Gal, thereby masking the substrate which would give rise to difucosylated $Lewis^e$, $Lewis^y$ (Orntoft *et al*., 1991). Loss of normal ABH structures and increases in Lewis glycans have been shown to correlate with poor prognosis in bladder and lung cancers (Marquez *et al*., 2004).

1.6.5.3 Truncated O-glycans

Mucins are high molecular weight glycoproteins which are normally present on the apical surface of cells (Burchell *et al*., 1983; Ellis *et al*., 1987; Kim *et al*., 1991). These glycoproteins carry glycan chains attached to the Ser/Thr residues in tandem repeat regions (Varki *et al*., 1999; Brooks *et al*., 2002). In contrast to normal tissue architecture, in cancer, mucins are expressed over the entire cell-surface and are released to the extracellular space and secreted into the body fluid

(Gold & Freeman., 1956). Changes in mucin glycosylation have been observed in numerous cancers including breast, pancreatic, ovarian and bile duct cancers (Yonezawa, 1994). The availability of monoclonal antibodies has facilitated the detection of altered levels of mucin in the blood of cancer patients (Berry *et al*., 1985, Ellis *et al*., 1985). The aberrant glycosylation of mucin CA125 has been observed in ovarian cancer patient serum (Hogdall, 2008). Incomplete glycosylation of O-linked mucins is a key feature that accompanies the development of a metastatic phenotype and is characterised by an increase in T-antigen, sialyl-Tn and Tn antigen (Springer, 1997). Increase levels of sialyl-Tn antigen in tumour cells due to increased levels of ST6GalNAc I transferase has been observed in several cancers (Julien *et al.,* 2006; Marcos *et al*., 2004; Senda *et al*., 2007; Sewell *et al*., 2006; Vazquez-Martin *et al*., 2004.,). The main changes that occur in O-linked glycosylation in cancer are summarised in table 1.2.

Table 1.2: Changes in O-glycans associated with cancer. Adapted from Brockhausen, (2006a)

O-type glycans	Structure	Increase (+) / decrease (-) in cancer
Core 1, T antigen	Galβ1-3GalNAcα-Ser/Thr	+
Core 2	GlcNAcβ1-6(Galβ1-3)GalNAcα-Ser/Thr	+/-
Core 3	GlcNAcβ1-3GalNAcα-Ser/Thr	+
Core 4	GlcNAcβ1-6(GlcNAcβ1-3)GalNAcα-Ser/Thr	+
Tn antigen	GalNAcα-Ser/Thr	+
Type 1 chain	[GlcNAcβ1-3 Galβ1-3]*n*	-
Type 2 chain	[GlcNAcβ1-3Galβ1-4]*n*poly-*N*-acetylactosamines	+
Sialyl-Tn antigen	Sialylα2-6GalNAcα-Ser/Thr	+
Sialyl Lewis^a	Sialylα2-3Galβ1-3(Fucα1-4)GlcNAcβ1-Gal-	+
Sialyl Lewis^x	Sialylα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-Gal-	+
Sialyl-dimeric Lewis^x	Sialylα2-3Galβ1-4(Fucα1-3)GlcNAcβ1-3 Galβ1-4(Fucα1-3) GlcNAcβ1-3Gal-	+

1.6.5.4 Alteration in sialylation

Many studies have reported that an increase in sialylation is a common feature of cancer (Yu *et al*., 2005). Increases in sialylation often occur as a result of an elevation in α2-6-linked sialic acid binding to N-acetylactosamine or to GalNAc-α1-O-Ser/Thr units on O-linked glycans (Varki *et al*., 1999). The elevation in cell-surface sialylation in cancer may reduce the attachment of malignant cells to the matrix and protect malignant cells by evading the host immune recognition via the alternative complement pathway (Pilatte *et al*., 1993). Significant changes in sialic

acid containing structures have been reported in cancer for example the ganglioside epitope 9-O-acetylated GD3 is elevated in human melanoma cells (Cheresh *et al*., 1984). Apart from an increase in sialylation in tumour cells, reduced sialylation has also been implicated in cancers such as colon carcinomas (Varki *et al*., 1999). This is further discussed in section 1.7.1.2.

1.7 Breast cancer metastasis and glycobiology

Tumour progression and metastasis may largely be attributed to alterations in the genetic makeup of the cancer cells; such genetic changes include mutations in proto-oncogenes and tumour suppressor genes which subsequently give rise to subclones of tumour cells with different behavioural characteristics compared with their normal counterparts (Hanahan and Weinberg, 2000). The acquisition of altered genetic traits promotes tumour cell dissociation, degradation of the basement membrane (BM) and invasion into the surrounding extracellular matrix (ECM) (Cavallaro & Chistofori, 2004; Condeelis & Pollard, 2006). After migration through the ECM, successful dissemination through the haematogenous or lymphatic circulatory system requires that the tumour cell survives immunological attack (Nash *et al.*, 2002). Finally, the tumour cell needs to adhere to the vasculature, extravasate into the organ of metastasis and proliferate at the secondary site (Kaplan *et al*., 2006). These steps are shown in figure 1.5.

A key feature of the metastatic cascade is the role of the cell adhesion molecules, many of which act via protein-glycan interactions. Oncogenic transformation is often accompanied by altered glycosylation patterns of proteins and lipids and these appear to have functional implications in potentiating the metastatic spread of tumours (Hakomori, 1996). While many genetic and physiological processes involved in the metastatic process have been documented, the patho-physiological role of altered glycosylation remains less well understood. Identifying glycosylation changes on glycoconjugates at an early stage of cancer development may offer the potential for earlier diagnosis, for monitoring disease progression, and/or as targets for biological tumour therapies. The role of glycans on proteins function in

the metastatic process is described below, has been published as a review (Rambaruth & Dwek, 2011).

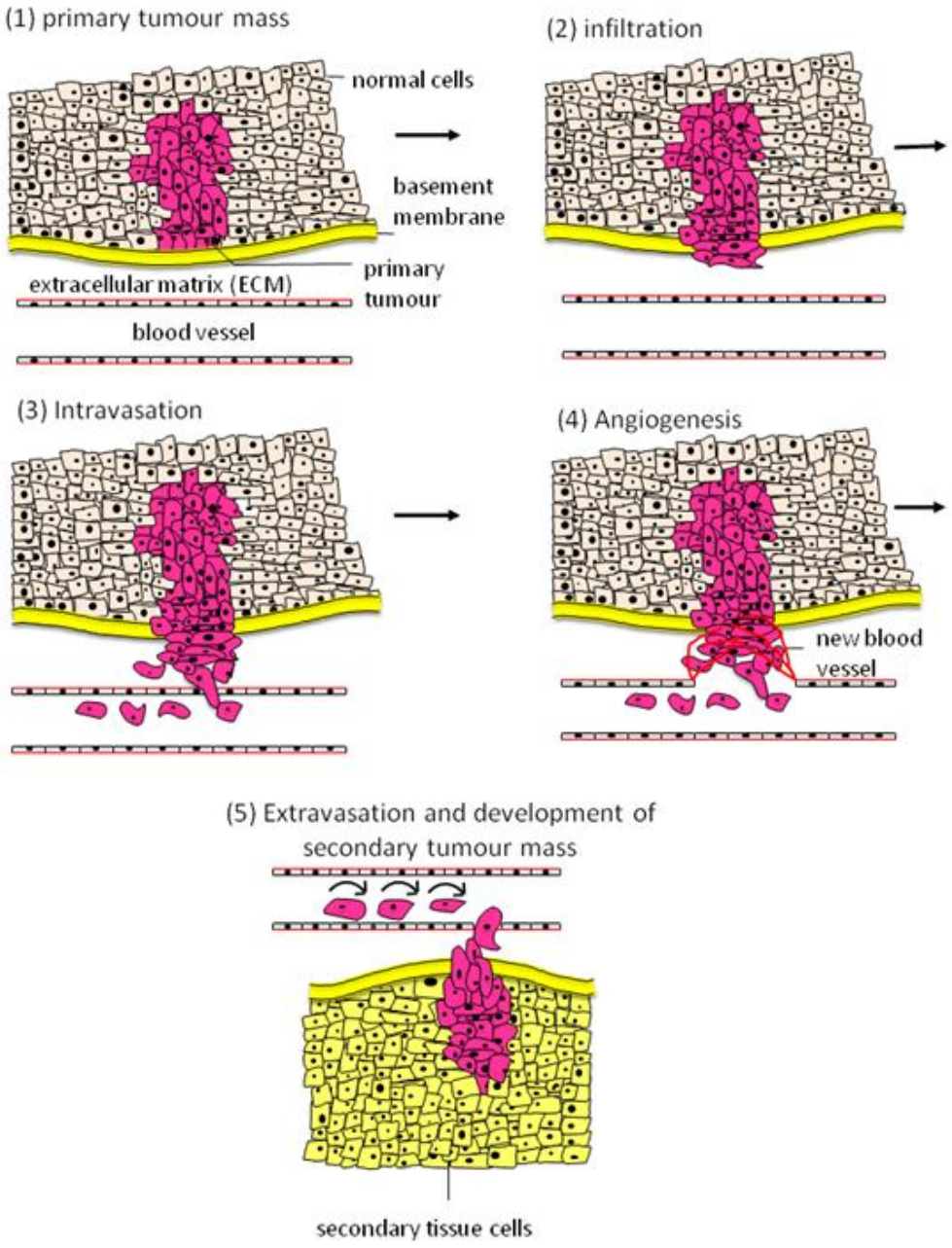


Figure 1.5: The main steps in the metastatic cascade. The following takes place; (1) formation of primary the tumour mass, (2) infiltration of primary tumour cells through the basement membrane (BM) and extracellular matrix (ECM), (3) intravasation of tumour cells into the circulatory system, (4) angiogenesis, (5) extravasation of tumour cells to finally establish a secondary tumour in organ the of metastasis.

1.7.1 Tumour invasion (loss of cell–cell homotypic adhesion)

Tumour invasion is dependent on a loss of intercellular adhesion and the transmigration of cells through the basement membrane (BM) as well as through the surrounding extracellular matrix (ECM). A key part of this process is the detachment of cancer cells from the primary tumour mass and acquisition of a more motile and invasive phenotype. A growing body of evidence shows that in addition to defective architectural remodelling and de-organisation of BM and ECM components (such as laminin-5 and collagen IV), changes to the cytoskeleton and degrading properties of proteases, the initial process of dissemination is facilitated by altered cell surface glycans structures which effect the adhesive properties of neoplastic cells (Hakomori, 1984; Dennis *et al*., 1999; Nakano *et al*., 2000; Lohi, 2001; Brockhausen, 2006; Abiatari *et al*., 2010).

1.7.1.1 E-cadherins

The adhesive interactions of epithelial cells are in part mediated by E-cadherin, a calcium dependant transmembrane glycoprotein receptor (Takeichi, 1990). E-cadherin has five extracellular domains that dimerise with the E-cadherin of adjacent cells to form homotypic cell–cell interactions (Takeichi, 1990). The cytoplasmic tail of E-cadherin interacts with the actin cytoskeleton through beta catenin and this function in the recruitment of multiprotein complexes at the plasma membrane known as adherens junctions (AJs) (Brieher *et al.*, 1996; Gumbiner, 2000; Wheelock & Johnson, 2003). The molecular organisation and stability of AJ formation is influenced by the presence of altered glycans on the cell surface and increased β1-6 branched N-glycans on E-cadherin, a phenotypic change commonly observed in malignancy (Jamal *et al.*, 2009). Elevated levels of N-acetylglycosaminyltransferase V (as described in section 1.6.5.1), required for the biosynthesis of β 1-6 N-glycans, has been correlated with tumour cell migration and invasion in a murine model of cancer metastasis and expression of *Mgat5*, the gene that codes for N-acetylglycosaminlytransferase V is regulated by Ras-Raf-MAPK, a signal transduction pathway commonly activated in tumours (Saito *et al.*, 1995). In normal cells, *N*-acetylglucosaminyltransferase III (GnT-III) catalyses the

formation of a bisecting GlcNAc structure in N-linkages and has been associated with tumour suppression of B16-hm murine melanoma cells (Yoshimura *et al.*, 1995). GnT-V knock-down with *si*RNA revealed enhanced cell detachment of cells from the ECM and gene expression analysis has also shown that there is bidirectional crosstalk between GnT-III/GnTV expression and E-cadherin which consequently can modify E-cadherin N-glycosylation status (Pinho *et al.*, 2009). Such studies helped to determine the involvement of glycosylation changes in the metastatic process. Figure 1.6 shows the influence of GnTIII and GnT V on AJ formation.

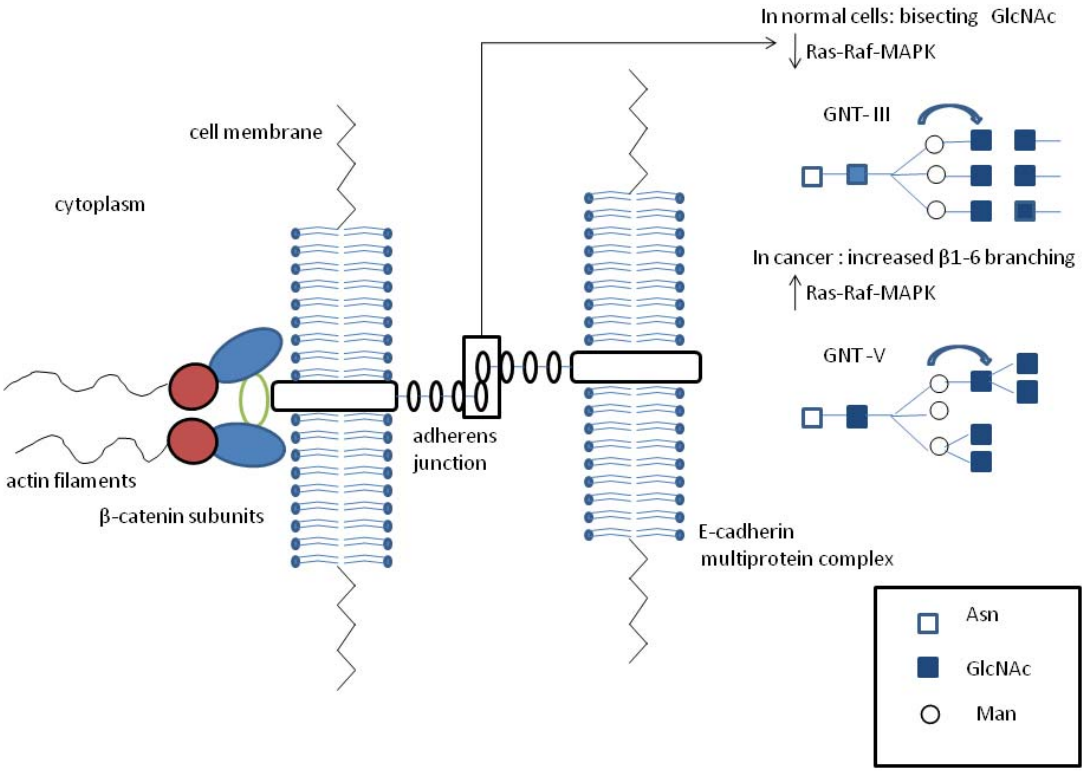


Figure 1.6: Adherens junction (AJ) formation between two adjacent epithelial cells. E-cadherins consist of an outer domain and a cytoplasmic tail. The cytoplasmic tail interacts with actin cytoskeleton through beta catenin and these functions in the recruitment of multiprotein complexes at the plasma membrane to form adherens junctions. GnT-III catalyses the formation of bisecting GlcNAc N-linked structures in normal cells whilst GnT-V catalyses increased levels of β1,6 GlcNAc containg N-glycans in cancer.

1.7.1.2 Siglecs and sialylation

The functional role of endogenous carbohydrate binding lectins in the mediation of homotypic cell-cell adhesion is not fully understood. Sialic-acid-binding immunoglobulin-like lectins (Siglecs) play key roles in mediating cell-cell interactions in the immune system via glycan recognition on pathogenic organisms (Crocker *et al.*, 2007). Sialylated epitopes recognised by Siglecs are common to many glycoproteins and glycolipids suggesting that Siglec-glycoconjuate interactions may have a functional role in physiopathological processes (as described in section 1.6.5.4). The binding of the Siglec, sialoadhesion, a macrophage-specific cell surface receptor is regulated by the sialylation status of the receptor (Barnes *et al.*, 1999). Sialoadhesin recognises ligands containing α2-3 linked sialic acid (mostly of SLe^x) on glycoconjugates of leukocytes. Sialylation of soluble sialoadhesin inhibits its binding to Jurkat cells (immortalized T lymphocyte cell) ligands, potentiating the loss of cell-cell interactions (Barnes *et al.*, 1999). Tumour cells tend to produce increased levels of sialic acid containing glycoconjugates, a phenotype which has been associated with invasion and malignancy (Bogenrieder & Herlyn, 2003; Seidenfaden *et al.*, 2003; Suzuki *et al.*, 2005). Increased levels of α2-6 linked sialylated oligosaccharides have been observed in colon cancer (Sata *et al*., 1991). Aberrant expression of sialic acid containing glyconjugates in cancer might inhibit interaction between Siglecs and their binding partners through electrochemical repulsion (due to the negative charge of the sialylated epitopes), and thereby resulting in reduced homotypic cell-cell adhesion. However additional studies are needed to explore this hypothesis.

1.7.2 Motility of cancer cells and glycosylation

1.7.2.1 Integrins

Integrins are an important group of cancer cell-surface receptors that mediate cell-ECM interactions by binding matrix proteins such as fibronectin (FN) and laminin. Integrins are involved in the regulation of cell adhesion, migration, differentiation, and apoptosis and remodelling of the cytoskeleton through their ability to transduce multiple intercellular signals (reviewed in Tarone *et al.*, 2000). Integrins consist of α and β subunits. The Arg-Gly-Asp (RGD) sequence of α5β1 integrin, interacts mainly with fibronectin, while α3β1 binds preferentially to basement membrane laminins (Ruoslahti, 1996; Frisch & Ruoslahti, 1997; Kreidberg, 2000; Gu *et al.*, 2001, Gu & Taniguchi, 2004) as shown in figure 1.7. Aberrant levels of integrin proteins have been reported in breast, colon and prostate cancer, with many of the RGD-binding integrins implicated in metastatic cancer progression (Natali *et al.*, 1992; Dedhar *et al.*, 1993; Hardan *et al.*, 1993; Zutter *et al.*, 1993; Oku *et al.*, 1996; Pouliot *et al.*, 2001). Amino acid sequencing studies have revealed fourteen and twelve potential N–linked glycosylation sites on the α5 and β1subunits, respectively, suggesting that cell surface integrins may act as major carriers of N-glycan structures (Nakagawa *et al.*, 1996; Gu & Taniguchi, 2004). In epithelial cells, changes in the N-glycosylation status of integrins (such as an increase β1-6 branched N-glycans) may affect the cell–cell and cell-matrix interactions and hence promote cell motility and invasiveness (Dennis *et al.*, 1987; Hakomori, 1996; Asada *et al.*, 1997). Experimental studies with NIH3T3 fibroblasts, transformed with the *Ras* gene, have shown that adhesion to FN was enhanced with increased in β1,6 GlcNAc branched structures on α5β1 integrin subunit (Asada *et al.*, 1997). Removal of α2, 8-linked sialic acids residues from the α5 integrin subunit of G361 melanoma cells resulted in an inhibition of cell adhesion to FN (Nadanaka *et al.*, 2001). Recently Saint-Guirons *et al* (2007) demonstrated that *Helix pomatia* agglutinin (HPA) bound to the integrin αv/α6 subunit in an *in vitro* model of metastatic colorectal cancer. HPA bound strongly to metastatic HT29 cells but not to non-metastatic SW480 cells. Taken together,

these results suggest that the glycosylation status of integrin, which is altered in cancer, may have direct impact in cancer cell dissemination.

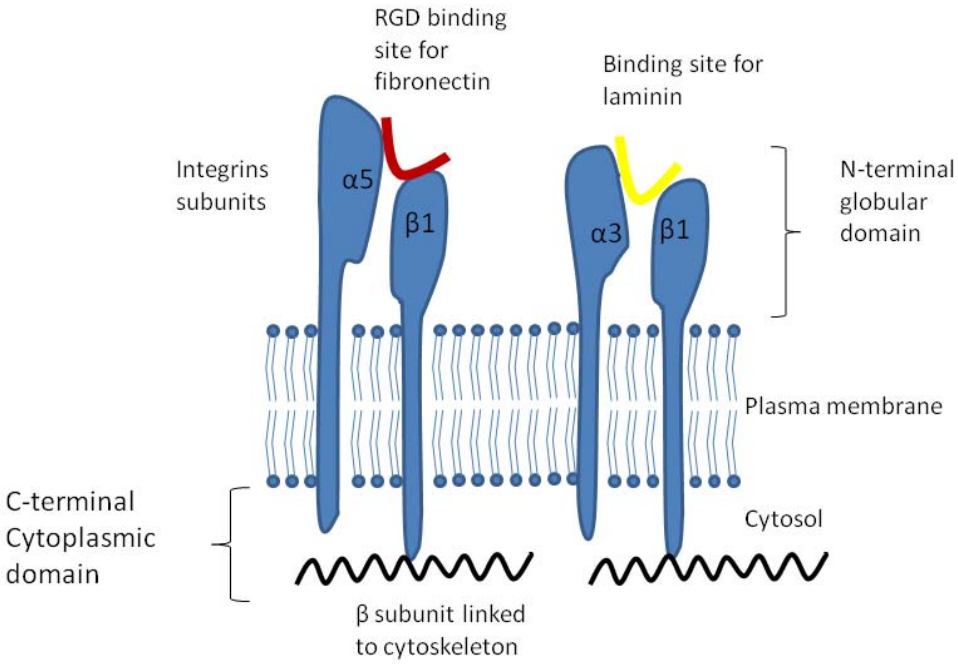


Figure 1.7: Integrin binding sites. Integrin consists of an α and β subunit, each dimeric unit has a large extracellular N-terminal globular domain which forms the RGD/ laminin binding site and a short C-terminal cytoplasmic domain. The β subunit of the C- terminal domain is linked with the intracellular cytoskeleton.

1.7.2.2 Laminin

Laminin is a basement membrane protein that associates with integrin to promote cell motility in a range of cellular processes such as those of wound healing and cancer cell invasion (Colognato & Yurchenco, 2000). Introduction of β1,6 branched N-glycans of MKN45 cells (a human gastric carcinoma cell line), resulted in an increase in cell migration (in Lm332-null keratinocyte plated wells compared to MKN45 cells with bisecting GlcNAc) on laminin and resulted in reduced cell migration (Kariya *et al.*, 2008). Galectin-3 cross-links laminin with α3β1integrin subunits and epidermal growth factor receptor (EGFR) resulting in a supramolecular complex. Galectins have emerged as an important class of

carbohydrate binding protein in cancer and are further discussed in section 1.7.4.2. It appears that N-glycans modulate the interaction between laminin and α3β1 integrin subunits and hence indirectly promote cell migration.

1.7.2.3 CD44 and hyaluronate

CD44 is a cell surface glycosaminoglycan (GAG) receptor which binds extracellular hyaluronan found on diverse cell types, including progenitor cells (Avigdor *et al.*, 2004), epithelial cells (Cichy *et al.*, 2002), fibroblasts (Henke *et al.*, 1996) and leukocytes (Xu *et al.*, 2002). CD44 has been shown to bind fibronectin (Jalkanen and Jalkanen, 1992), collagen (Ehnis *et al.*, 1996), growth factors (Wolff *et al.*, 1999) and matrix metalloprotease-9 (MMP-9) (Yu & Stamenkovic, 1999), by virtue of its differentially spliced isoforms. CD44 is involved in key cellular processes such as cellular motility (Pure and Cuff, 2001), cell trafficking (DeGrendele *et al.*, 1997), cell-cell and cell-ECM adhesion (Prosper & Verfaillie, 2001). CD44 also binds cytokines, growth factors, chemokines and enzymes which are sequestered and then available to other cells or to the surrounding tissue (Jones *et al.*, 2000). Altered CD44 spliced variants have been reported in pathological conditions such as autoimmune, chonic inflammatory diseases (Mikecz *et al.*, 1995, Blass *et al.*, 2001) and in neoplasms (Salmi *et al.*, 1993, Tanabe *et al.*, 1993). In cancer the splice variants of CD44 show altered binding to hyaluronan and certain splice variants are associated with promotion of metastatic spread. CD44 undergoes extensive modification of the N-and O-linked glycosylation during carcinogenesis (Carter & Wayner, 1988, Stamenkovic *et al.*, 1989). In colon cancer, CD44 receptors have been shown to be modified with O-linked glycans. Inhibition of the O-linked glycosylation step enhances CD44-mediated adhesion to hyaluronate whereas inhibition of N-linked glycosylation had no effect on CD44 adhesion, hence it has been suggested that O-linked glycosylation may be as important as alternative splicing in regulating CD44 function in tumour dissemination. Transfection studies have also demonstrated than O-linked glycosylation modulates interaction between the B loop domains of CD44 and hyaluronate (Dasgupta *et al.*, 1996). On the other hand, Bartolazzi *et al*(1996) showed that inhibition of N-linked sugars on CD44 by treatment with tunicamycin on different

human cells was associated with the loss of CD44-mediated cell adhesion to hyaluronan, althrough treatment of the cells with deoxymannojiramycin, which blocks the synthesis of high mannose type Asn-linked glycans, had no effect on adhesion suggesting that N-linked glycans are necessary for CD44 function (Bartolazzi *et al.*, 1996). The two studies described above highlight the need for further investigation into the role of glycans on CD44 function.

1.7.3 Angiogenesis: the role of heparan sulphate (HS)

Angiogenesis is a complex process which involves multiple cellular events that leads to the formation of new vasculature from pre-existing blood vessels. During this process urokinase-plasminogen activator (uPA) and other proteins such as the MMPs degrade the basement membrane and the surrounding stroma. This is followed by migration and proliferation of endothelial cells at localised regions in the surrounding tissues and the formation of new vasculature (Iozzo & San Antonio, 2001).

Angiogenesis is essential for normal development and for physiological processes such as wound healing, formation of the uterine lining during the female monthly reproductive cycle (Fraser & Lunn, 2000) and development of blood vessels which link the mother to the fetus during pregnancy (Reynolds & Redmer, 2001). Aberrant angiogenesis is associated with pathological conditions, including tumour growth and metastasis and in order for a tumour to progress to a metastatic stage it must be supplied with an intratumoural blood vasculature which nourishes the cancer cells with oxygen and nutrients. This is crucial to support tumour growth beyond 2 mm diameter (Folkman, 2006). Angiogenesis is modulated by a delicately controlled balance of growth factors such as pro- and anti-angiogenic factors which affect endothelial cell growth and differentiation. The main pro-angiogenic growth factors include fibroblast growth factors (aFGF and bFGF) and vascular endothelial growth factors (VEGF). VEGF levels are controlled and stimulated by hypoxia and by activated oncogenes. VEGF plays a critical role in the regulation of protein expression, cellular migration, division and apoptosis of endothelial cells (Ferrara, 2001, Robinson & Stringer, 2001). VEGF-A is a pro-

angiogenic factor, the level of which has been correlated with the development of blood vessels at the sites of the neoplasm (Chen *et al.*, 2004). Heparan sulphate (HS) containing proteoglycans, or the structurally related heparin, is frequently bound to growth factors at sites of angiogenesis. HS proteoglycans can occur as ECM components (for example perlecan), or as membrane, GPI anchored, molecules (for example syndecans and glypicans). These glycans also facilitate the binding of receptor tyrosine kinases to a variety of growth factors. Gene knock-down studies of N-deacetylase/N-sulphotransferase to block acetylation and sulphation of nascent heparan sulphate chains inhibited tumour angiogenesis (Zhou *et al.*, 2004). Absence of the HS proteoglycan perlecan has also been shown to block carcinoma growth in several *in vivo* models (Zhou *et al.*, 2004). Moreover a 10-15 times increase in the level of HS has been observed in the neovasculature acquired during angiogenesis (Marcum & Rosenberg, 1985). In this aspect of the metastatic process, proteoglycans support the process of tumour dissemination. Figure 1.8 illustrates the interaction of heparan sulphate with VEGF on endothelial cells receptors during the process of angiogenesis.

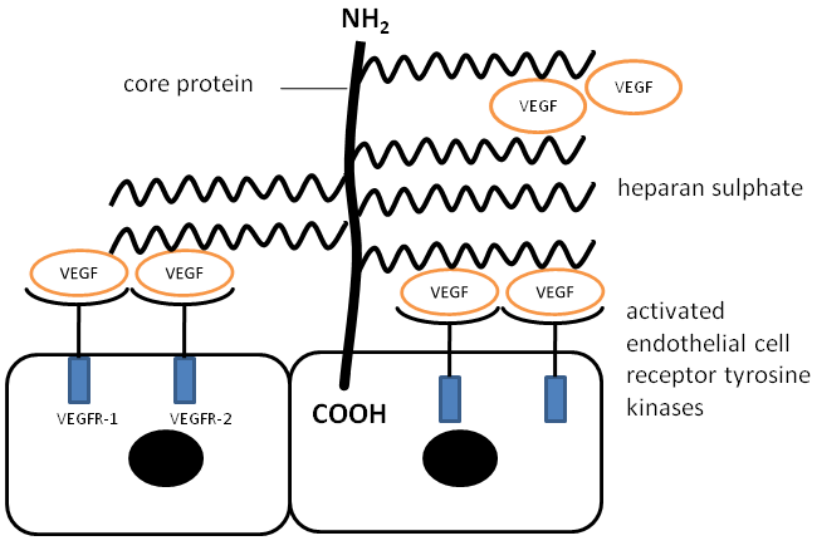


Figure 1.8: Heparan sulphate (HS) interaction with VEGF. HS proteoglycans enhances the interaction between VEGF and receptor tyrosine kinase of activated endothelial cells at the site of angiogenesis.

1.7.4 Tumour cell interactions with the microvasculature and evasion of immune recognition

The release of cancer cells from a primary tumour mass and their invasion through the BM and surrounding ECM is mediated by glycoconjugates and enzymes of the ECM. Haematogenous dissemination of tumour cells is also governed by the adhesive interactions between the tumour cell and the microvasculature within the tumour (Fidler *et al*., 1970). Tumour cells either adhere at the gaps between the endothelial cells of blood vessels and escape in the blood circulatory system (Kramer, 1982). This is an important process for intravasation (Weidner, 2002). Once in the blood-stream a fraction of tumour cells travel as emboli with platelets rather than as individual entities (Morimoto *et al.*, 2008). Cells at the centre of the tumour mass are physically shielded from immune recognition (Fidler & Bucana 1977). The tumours then adhere to the endothelium prior to extravasation at a distant site to form secondary cancer foci (Nicolson, 1985).

1.7.4.1 C-type lectin receptors (CLR) including selectins

The adhesion of circulating tumour cells to the endothelium and subsequent movement to the organ of metastasis involves similar mechanisms to those used by leukocytes when homing to sites of inflammation (Lasky, 1995, Springer, 1995). Selectins are adhesion molecules mediating the interaction of the leukocytes and endothelial cells with other cells including platelets in the vascular endothelium. This area of biology has been extensively reviewed by Gonzalez-Amaro and Sanchez-Madrid (1999). Selectins are a family of transmembrane proteins of the C-type lectin family which include L-selectin, P-selectin and E-selectin. These selectins recognise SLe^x and SLe^a, sialic acid containing ligands which are predominantly present at the tips of *O*-linked glycans on the surface of leukocytes (Varki, 1994). Tumour derived mucins, from carcinomas such as breast, colon, pancreas and stomach, frequently show increased level of Lewis glycans, and hence the interaction between tumour cells and the vasculature were throught and subsequently shown to be mediated by selectins on the endothelial cells (Kim *et al.*, 1998; Varki, 1997). *In vitro* studies with the CRC cell line HT29 showed that adhesion to human umbilical cord endothelial cells (HUVEC) was inhibited by a

monoclonal antibody directed against E-selectin ligands (Srinivas *et al.*, 1996). Introduction of HT29 colorectal cancer cells into E-and P- selectin knock-out severe combined immunodeficient (SCID) mice resulted in a decrease in metastasis formation compared to wild type mice, here, significant metastasis formation occurred in the lungs (Kohler *et al*., 2009). Flow experiments indicated that tumour cells roll and tether on an E- and P-selectin matrix in a manner similar to leukocytes but firm adhesion occurred only on E-selectin. The interactions of selectins with tumour cells and leukocytes are illustrated in figure 1.9.

Other transmembrane CLRs of importance in cancer are the dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin (DC-SIGN) and scavenger receptor C-type lectin (SRCL). DC-SIGN play a key role in the adhesion of tumour cells to endothelial cells, again by virtue of their ability to recognise Le^x containing glycans present on the tumour cells (Guo *et al.*, 2004). Recent studies have demonstrated that DC-SIGN recognises modified glycans of CEA or MUC1 in colon tissues (Berinstein, 2002; Denda-Nagai & Irimura, 2000). Another CLR that may be involved in tumour cell adhesion is the SRCL which recognises Le^x and Le^a containing oligosaccharides. The precise role of SRCL in glycan mediated cell adhesion has yet to be elucidated. However it has been demonstrated that it binds to glycoproteins on the cell surface of breast tumour cells and hence could have a role similar to that proposed for the selectins in the metastatic spread of cancer (Elola *et al.*, 2007).

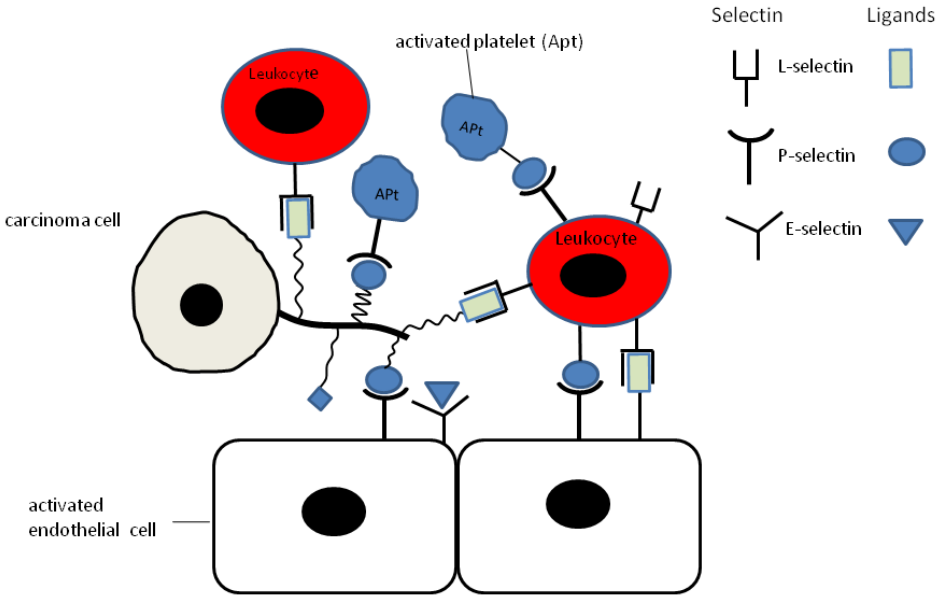


Figure 1.9: P-, E- and L- Selectin of activated endothelial cells. Tumour cells adhere to activated endothelial cells through the SLe^X or SLe^a and P-, E- L- selectin mediated interactions

1.7.4.2 Galectins

Galectins belong to a family of 14 structurally related proteins which share common homologous carbohydrate recognition domains (CRDs). Galectins have affinity for beta-galactoside-containing epitopes (Barondes *et al.*, 1994) and are involved in cellular events including cell-matrix adhesion, pre-mRNA splicing (Hughes, 2001), migration (Jung *et al.*, 2008), cell growth, apoptosis and differentiation (Yang & Liu, 2003). As galectins have the ability to form cross-linked structures between adjacent cells, they have been shown important for maintaining the stability of cellular adhesions. Galectin associated cross-linking is associated with receptor activation and induction of intracellular signalling which may subsequently induce apoptosis (Hernandez & Baum, 2002). Galectins are altered in pathological conditions including diabetes (Perone *et al.*, 2009), autoimmune disease, inflammation, atherosclerosis (Papaspyridonos *et al.*, 2008; Rabinovich *et al.*, 2007) and cancer cell metastasis (Hsu *et al.*, 1999; Irimura *et al.*, 1991, Lotan *et al.*, 1994). The loss of adhesive properties that galectins confer is implicated in a number of pathological states including tumour progression. Whilst most galectins

are intracellular proteins, some interact with glycoprotein receptors at the cell surface. The exposure of galactose residues at the terminal end of glycoconjugates (as observed in cancer associated antigens), for example the presence of the T antigen (Gal β1-3 GalNAc) and the Le^x trisaccharide (Galβ1-4(Fucα1-3) GlcNAc) may facilitate the binding of circulating tumour cells to galectins of the vascular endothelium during adhesion processes and may also affect the rolling events inside the vasculature (Glinsky *et al.*, 2000; Takenaka *et al.*, 2004). Figure 1.10 illustrates the involvement of galectins in the formation of cross-linked structures and how this is implicated in maintaining structural stability at the cell surface.

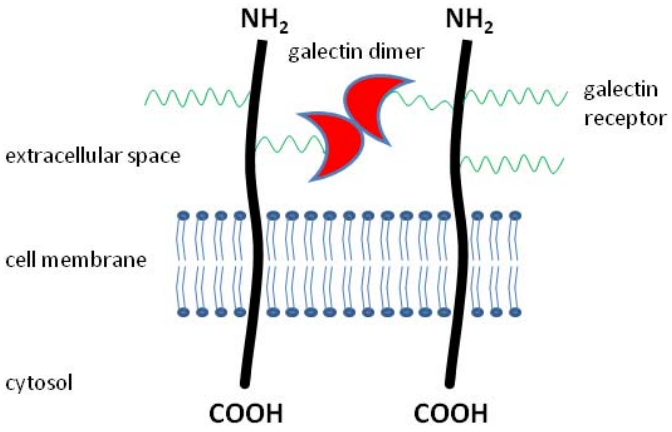


Figure 1.10: Galectin dimer formation. Galectin enhances the interaction between cell surface proteins by forming cross-linking structures between adjacent receptors.

1.7.5 Glycans in 'immune protection'

The survival of tumour cells and evasion of immune surveillance system is not fully understood. It is throught that sialylated glycans at the cell surface play a functional role in shielding them from immune attack, by processes similar to those involved in protection of parasites such as trypanasomes, bacteria and pathogenic fungi from phagocytosis by the cells of host organisms (Pilatte *et al.*, 1993). In such a model host immune system fails to discriminate the tumour mass as 'non-self' giving rise to secondary tumours and metastasis formation. Desialylation of pathogenic fungi results in an increased likeliness of phagocytosis by human macrophages, implying that sialic acid residues found on those organisms may have functional role in the immune protection (Pilatte *et al.*, 1993; Wasylnka *et al.*, 2001). On the other hand, the presence of branching N-linked oligosaccharides on gp120 on the envelope of the human immunodeficiency virus 1 (HIV-1), facilitates viral escape from the host immune system (Zhang *et al.*, 2004). Studies have revealed that gp120 carries an average of 25 potential N-glycosylation sites and that these N-glycans are critical for the correct folding, hence for proper structure and function of the virus envelope (Fischer *et al.*, 1996; Li *et al.*, 1993; Morikawa *et al.*, 1990; Zhang *et al.*, 2004) The cluster of N-glycans promotes the escape of the virus from the host immune system by hindering antigen presentation to neutralizing antibodies (Ly and Stamatatos, 2000) and cytotoxic T- lymphocytes (CTL) (Duenas-Decamp *et al.*, 2008; Li *et al.*, 2008; Kwong *et al.*, 1998). If parallels are drawn between tumour and infectious disease pathogenesis, the glycan structures present on the surface of tumour cells may serve as a 'glycan shield' facilitating cancer cell escape from the host immune system.

1.8 Lectins and cancer

1.8.1 History of lectins

Lectins are agglutinating proteins of non-immune origin that were discovered almost 150 years ago. Weir Mitchell first observed an agglutinating substance in snake venom in 1860 and then in 1888, Peter Herman Stillmark described a protein extract from the seeds of the castor tree (*Ricinus communis*) which had the ability to agglutinate erythocytes (Mitchell, 1860; Stillmark, 1888). It was not until 1954 that it was first shown that agglutination by lectin occurs selectively via recognition of specific carbohydrate epitopes (Boyd & Shapleigh, 1954). Lectins are now known to be present throughout nature from bacteria to plants, primitive animals to humans (Bies *et al*., 2004).

Whilst the major function of lectins in viruses and bacteria is often key to their pathogenicity (Lehmann *et al*., 2006), in plants they are believed to protect against pathogenic bacteria through the binding on the cell surface of the microbial cells (Rudiger & Gabius, 2001). Lectins are involved in the recognition of molecules in the immune system in invertebrates and animals. Invertebrates' lectins are throught to act as primitive immune molecules directed against pathogens by recognising extracellular oligosaccharides on the microorganisms. Several lectins in humans such as ficolins, mannan-binding lectin and the membrane bound macrophage mannose receptor mediate elimination of pathogens (reviewed in Kilpatrick, 2002). Several cell adhesion molecules with lectin properties, including selectins and CD44, are involved in cell recognition and cell trafficking. In animals, galectins are known to have immune-regulatory properties (Levi *et al*., 1983; Offner *et al.,* 1990) and cytokines play a role in immune-regulation via lectin-like interactions (Fukushima & Yamashita, 2001).

A growing body of evidence has shown that some lectins bind preferentially to malignant cells and these have attracted interest in the area of cancer research. Lectins can be used to investigate the presence of carbohydrate structures in or on cancer cells, in much the same way as antibodies are used to probe cells and tissues for the presence of specific antigens (Ikeda *et al*., 1994; Schumacher *et al*.,

1994). Aub *et al* (1965) first reported that the lectin wheat germ agglutinin (WGA) recognised cells with a malignant phenotype. Since that time, many lectins were shown to detect alterations in glycosylation that accompany malignant transformation, for instance the lectin from *Erythina cristagalli* (ECA) showed differential binding between normal and cancerous colonic epithelia (Baldus *et al*., 1996). Other lectins such as *Dolichos biflorus* agglutinin (DBA) and *Ulex europaeus* agglutinin-1 (UEA-1) and WGA displayed differential binding pattern between normal and cancerous colorectal cells. Another lectin that has been the focus of several studies is the lectin from *Helix pomatia* (the Roman snail) as its binding to cancer tissue samples has been associated with poor patient prognosis (as described in section 1.8.2 and 3.1). *Helix pomatia* agglutinin (HPA) has been suggested to be a useful tool for identifying aggressive epithelial cancer and the epitopes recognised by this lectin are the focus of the work in this thesis.

1.8.2 *Helix pomatia* agglutinin (HPA)

HPA is a 79 kDa hexameric molecule composed of identical monomers present in at least 12 glycoforms (Hammarstorm & Kabat, 1969; Hammarstorm *et al*., 1972; Vretbald *et al.,* 1979). Each HPA monomer consists of six anti parallel beta sheets connected by short loops, which form a sandwich structure stabilised by a disulphide bridge between cysteine 9 and 80. The dimers associate with each other to form a trimer by the formation of three disulphide bonds between cysteine 42 on each monomer. Each HPA monomer consists of one GalNAc and Zn^{2+} binding site with the sugar binding site formed by the hairpin-like loops that connect the strands at the extremities of each monomer. The lectin specificity for GalNAc results from the particular network of hydrogen bonds where a histidine residue makes hydrophobic contact with the aglycon, rationalizing the preference for GalNAc bearing an additional sugar or amino acid in the α position (Sanchez *et al*., 2006; Lescar *et al*., 2007). These lectin structures were solved and this is shown in figure 1.11 and provides the molecular basis for the use of the lectin HPA in cancer research.

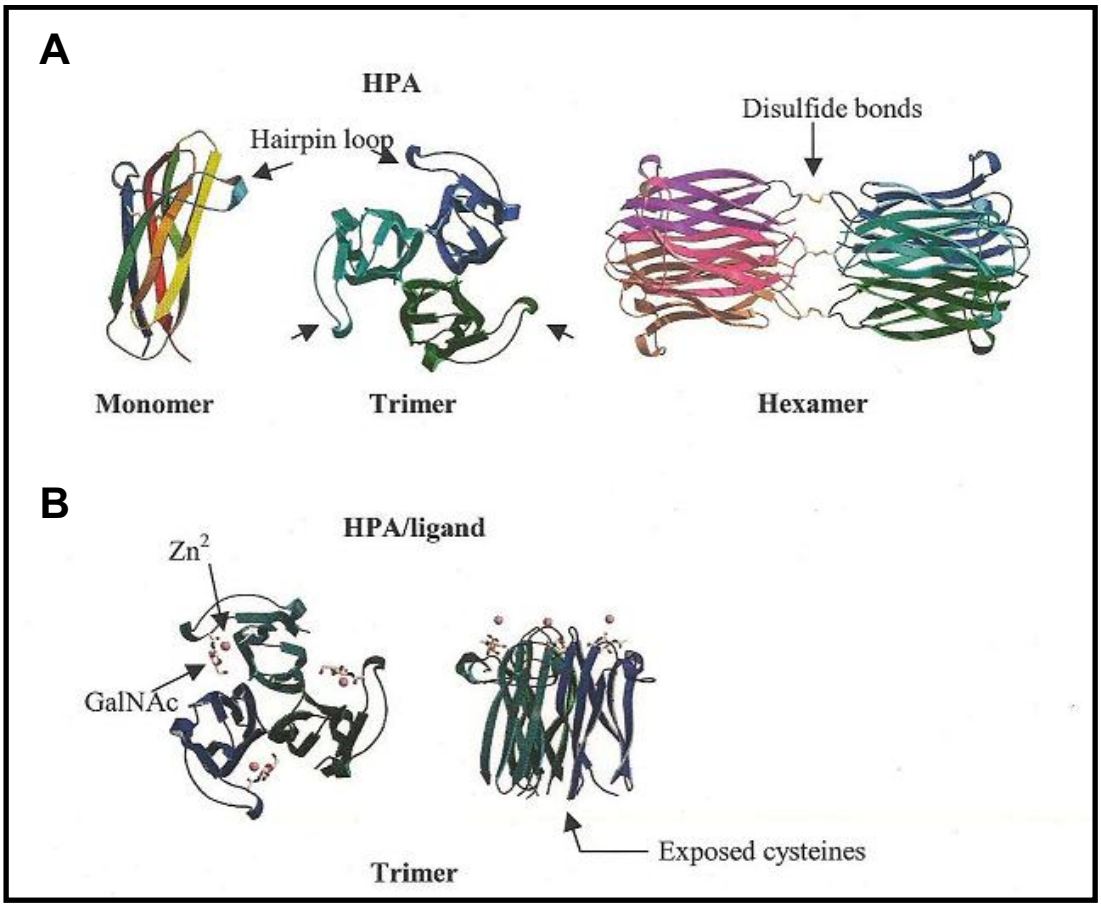


Figure 1.11: HPA structure. Panel A shows the structure of HPA alone obtained with a resolution of 2.5 A$^{o.}$ Panel B shows the HPA structure in presence of GalNAc and Zinc molecule (Adapted from Sanchez *et al*., 2006).

The lectin HPA is a blood group A specific lectin and played a significant role in experiments to determine the sugar composition of antigens associated with the ABO blood group system (Prokop *et al*., 1965; Uhlenbruck, 1966). HPA is extracted from the albumen gland of the Roman snail where it appears to be involved in the innate immune system of the snail by conferring protection for fertilised eggs (Prokop *et al*., 1965). The lectin has has the ability to aggregate bacterial pathogens and herpes virus (Kholer *et al.,* 1973; Patchett *et al*., 1991; Slifkin and Cumbie, 1989). HPA has binding specificity for the Forssman antigen (αGalNAc1-3GalNAc) (Baker *et al*., 1983), blood group A antigen (GalNAcα*1*-3Galβ1-4GlcNAcβ1-Fucα1-2) (Anderson & Haas, 1984; Mourali *et al*., 1980), Tn antigen (αGalNAc1 -O- Ser/Thr) (Piller *et al*., 1990; Springer, 1989), terminal α-

GalNAc and α-GlcNAc (Hammerstorm & Kabat, 1969) and also sialic acid (Dwek *et al*., 2001).

It is now known that in addition to agglutinating human blood-group A erythrocytes, HPA also binds cancer cells associated with poor clinical outcome suggesting that HPA recognises altered glycans on the surface of cancer cells. The predictive value of HPA was first shown over two decades ago when metastatic breast cancers of patients with poor prognosis were found to bind the lectin (Leathem & Brooks, 1987). In 1991, Leathem and Brooks conducted a study of 373 primary cancers from patients which had been followed-up for a period of up to 15 years and demonstrated that HPA staining of the primary tumours correlated with poor clinical outcome of the patients and lymph node status, althrough no correlation were made with tumour size, histological grade or age at diagnosis. The prognostic value of HPA has also been observed in other cancer types such as oesophageal (Yoshida *et al*., 1994; Takashi *et al*., 1994), gastric (Kakeji *et al*., 1994), prostatic (Shirashi *et al*., 1992), lung (Kawai *et al*., 1991) and colorectal cancer (Ikeda *et al*., 1994; Schumacher *et al*., 1992).

HPA has the ability to bind to metastatic cancer cells by recognition of epitopes on the surface of cancer cells and these may be useful as a prognostic marker in a range of solid tumours. Approximately 80% of metastatic tumours contain HPA binding epitopes (Brooks and Leathem 1998). The HPA binding characteristics of cancer cells has been correlated with *in vivo* models of metastasis and migration of HPA positive breast and colon cancer cells to secondary sites (Schumacher & Adam, 1997, Kohler *et al*., 2010). Other groups have, however failed to establish a relationship between HPA binding and poor prognosis cancer (Galea *et al*., 1991; Gusterson *et al*., 1993). It has been since been shown that the reasons for the differences in the results acoss the different teams were methodological, and that an indirect method for the detection for lectin binding histochemistry based studies is preferred (Brooks *et al*., 1996). Despite the evidence that HPA has the potential for prognostication, the lectin does not yet have a role in clinical decision making for any tumour type, and the HPA binding glycoproteins of most cancer cells types has yet to be determined. This is explored in more detail in section 3.1.

Recently, HPA binding to CRC cell lines was investigated by Saint-Guirons *et al* (2007). HPA binding to metastatic HT29 and non-metastatic SW480 CRC cell lines was investigated. Membrane glycoproteins from these CRC cell lines were isolated by affinity chromatography and was analysed by 2-DE coupled with MS. HPA labelling using confocal microscopy revealed intense binding in the HT29 cell lines consistent with the phenotype of the cell line. Proteomic studies showed HPA bound to several membrane glycoproteins involved in cell adhesion/migration (integrins and annexins), re-modelling (tubulin, cytokeratins, actin) and anti-apoptotic pathways (Hsp-70, Hsp-90, Hsp-96 and TNFR-1) in this cell line. The authors' showed that these proteins bound HPA only the metastatic HT29 but not in the non-metastatic SW480. This work was the key reference to the work carried out in this thesis.

1.9 Aims of the project:

1. To establish an *in vitro* model for HPA binding using breast cancer cell lines (chapter 3).

2. To determine whether HPA recognises the same glycoproteins in breast cancer as previously reported in CRC (chapter 4).

3. To identify whether HPA recognises glycoproteins via binding to blood group substance, aberrant O-linked structures or O-GlcNAcylation (chapter 5).

Chapter 2

Materials and Methods

2.0 Materials and Methods

2.1 Cell lines

Human breast cancer cell lines HMT3522, BT474, MCF-7, T47D, and were selected as they have known HPA binding properties and their behaviour in terms of metastasis formation differ when implanted in SCID mouse (Brooks *et al*., 2001; Schumacher & Adam, 1997). The details of the four cell lines used in this study are shown in table 2.1. The cells were kindly donated by Dr. Brooks (Oxford Brookes University, UK), grown in Dulbecco's Modified Eagle Medium (DMEM) (Lonza, UK) supplemented with 10% v/v foetal calf serum (FCS) (Biosera, UK), penicillin and streptomycin (1% w/v, Sigma) in 5% v/v CO_2. Cells were grown in 25 cm^2, 75 cm^2 or 175 cm^2 flasks at 37°C (Hera Cell Incubator 240). The culture media was changed every 2-3 days and the cells were passaged by trypsinisation (0.5 g/l trypsin and 0.2 g/l EDTA, Sigma, UK) when the cells reached approximately 70–80% confluence (Donohue *et al*., 2006). Cells used for protein or mRNA extraction were then grown to near confluence, washed 3 times in phosphate buffered saline (PBS, Sigma, UK), mechanically detached using a sterile plastic cell scraper in 25 ml of PBS and centrifuged at 500 x*g* in a Centaur 1 MSE centrifuge for 10 min. Before using the cells for experimental assays, a cell count was performed using a haemocytometer. Approximately 1.5 million cells were used for protein extraction experiments and 10,000 cells were used for mRNA extraction. All the cells used in this study, were kept within 10 cellular passages. Cell pellets were stored at -80^{o}C until protein extraction was carried out.

Table 2.1: Characteristics of the breast cell lines used in this study. ATCC number, cell origin and HPA binding status of HMT3522, BT474, MCF-7 and T47D. N/A: not applicable.

Cell line	ATCC number	Derived from	Reference (s)	Published HPA status
HMT3522	N/A	Benign fibrocystic breast tissue used to represent 'normal' breast cell.	Briand *et al*., 1987	Brooks *et al*. 2001
BT474	HTB-20	Primary breast cancer	Lasfargues *et al*., 1978	Brooks *et al*., 2001
MCF-7	HTB-22	Malignant pleural effusion from primary infiltrating ductal cancer	Soule *et al*., 1973	Brooks *et al*. 2001; Schumacher, 2004
T47D	HTB-133	Malignant pleural effusion from primary infiltrating ductal cancer	Keydar *et al*., 1979	Schumacher & Adam, 2004

2.2 Light microscopy

To assess the lectin binding properties of the cells, methods developed by Brooks & Hall, (2002) were adapted as follows: Cells were grown in 6 well plates until near confluence for approximately 24 hours after passage and fixed for 20-30 min in 10% v/v formalin in PBS (pH 7.4). After washing away the formalin with PBS, the fixed cells were blocked in 5% w/v bovine serum albumin (BSA) for 30 min. The cells were incubated with biotinylated HPA (Sigma, UK) prepared in PBS and used at 10 µg/ml for 2 h. The cells were then washed 3 times for 5 min with PBS and incubated with horseradish (HRP) conjugated streptavidin (Sigma, UK) prepared in PBS and used at 10 µg/ml for 1 h at room temperature. Detection was performed using the chromogenic substrate diaminobenzidine (DAB) for 5 min. DAB was prepared in PBS/ H_2O_2 (6 mg of DAB prepared in 9 ml of TBS and 60 µl of 30 vol H_2O_2).

2.3 Confocal microscopy

2.3.1 Lectin staining

To assess the lectin binding properties of the cells methods developed by Brooks & Hall, (2002) and Saint-Guirons *et al* (2007) were adapted as follows: Cells were grown in 6 well plates until near confluence for approximately 24 hours after passage and fixed for 20-30 min in 10% v/v formalin in PBS (pH 7.4). After washing away the formalin with PBS, the cells were incubated for 30 min at 37°C with 1 mg/ml trypsin (type II from porcine pancreas; Sigma, UK). Trypsinisation was used for antigen retrieval. Prior to fluorescent lectin staining, cells were fixed and washed as above and blocked in 5% w/v BSA for 30 min. Fluororescein isothiocyanide (FITC) and tetramethylrhodamine isothiocyanide (TRITC) conjugated HPA, SNA and PNA encompassing different sugar binding properties were used to stain the cells (table 2.2).

Cells were incubated in the dark, at room temperature, after washing 3 times for 5 min in PBS, with each lectin (10 µg/ml in PBS) for 1 h. All lectins, (except HPA which was obtained from Sigma, UK), were purchased from Vector Laboratories, Burlingame, CA. A DNAse free ribonuclease A stock solution was prepared by boiling for 10 min in 10 mM sodium acetate buffer, pH 5.2. The cells were then treated with 100 µg/ml ribonuclease A (Sigma, UK) for 20 min at 37°C in PBS and the nuclei were counter-stained using To-Pro-3 (Molecular Probes, Eugene, USA) at 1 µM in PBS for 20 min. As a negative control for lectin staining, cells were incubated with buffer alone.

Table 2.2: The lectins used in this study and their nominal binding sugars

Lectin	Abbreviation	Nominal binding sugar	Fluorophore
***Helix pomatia* agglutinin**	HPA	GalNAc/GlcNAc	TRITC
Sambucus nigra	SNA	α2,6 linked-sialic	FITC
agglutinin	PNA	Galβ1-3-GalNAc	FITC
Arachis hypogaea			
agglutinin			

2.3.2 Antibody labelling

The cells were grown as above and fixed for 20-30 min in methanol diluted with PBS (1:1 volume). After washing away the methanol with PBS, the fixed cells were trypsinised and blocked as above. The nuclei were counter-stained using To-Pro-3 prior to antibody incubation. The monoclonal mouse IgM anti-human-O-GlcNAc antibody (sc-81483, Santa-Cruz Biotech, UK) was used at 10 µg/ml for 1 h followed by 20 min incubation with a monoclonal goat anti-mouse IgG-FITC at 5 µg/ml (sc-2010, Santa Cruz Biotech, UK). A control was included in which the anti-O-GlcNAc step was omitted. For co-localisation studies, lectin staining was performed prior to antibody labelling.

2.3.3 Localisation of the Golgi apparatus

The cells were grown as above and fixed for 20-30 min in 10% v/v formalin in PBS, pH 7.4. After washing away the methanol with PBS, the fixed cells were trypsinised and blocked as above. Nitrobenzoxadiazole (NBD) labelled C_6 ceramide dye (NBD-C_6-ceramide, B-34400, Invitrogen, UK) has been shown to locate the Golgi apparatus (Ktistakis *et al.,* 1995). The NBDC6-ceramide was applied to the cells at 10 μM for 1 h at 4^{o}C. The ceramide was prepared in 4-(2-

hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer (Invitrogen, UK). For co-localisation studies, lectin staining was performed prior to labelling the Golgi apparatus.

2.3.4 Specificity of HPA binding

The specificity of HPA binding was assessed in T47D cells. Cells were grown in 6 well plates, fixed and blocked with 5% w/v BSA as described in section 2.2.1 The specificity of the lectin binding was evaluated by assessing the intensity of HPA staining after pre-incubating the lectin (10 μg/ml) for 30 min with increasing concentrations of GalNAc or GlcNAc (25 mM, 50 mM and 100 mM) (Sigma, UK). Mannose (50mM) (Sigma, UK) was included in the experiment as it has not been shown to bind HPA. The cells were incubated in the dark with the HPA/sugar mixture for 1 h and counter-stained with To-Pro-3, as described earlier. The intensity of the HPA staining was evaluated in 10 high power fields using a scoring system for each condition.

2.3.5 Image capture

Images were collected using a Leica TCS SP2 confocal microscope (Leica Microsystems, Milton Keynes, UK) with a X63 ceramic dipping objective and by sequential scanning. Images were acquired at a scanning speed of 400 Hz, 1024 x 1024 pixel resolution and with a line average of 4. A 466 nm laser was used for the excitation of NBD C_6-ceramide (intensity 35%, emission bandwidth 500-536 nm), 488 nm for IgG-FITC (intensity 40%, emission bandwidth 500-550 nm), 543 nm for HPA-TRITC (intensity 40%, emission bandwidth 550-630 nm) and 633 nm laser for the To-Pro-3 (intensity 35%, emission bandwidth 650-720 nm). These parameters are summarised in table 2.3. The background was compensated by adjusting the gain and offset commands. For 3-D images, Z-stacks were scanned at 1 μm increments.

Table 2.3: Parameters used for confocal microscopy work with fluorophores

Fluorophore	Excitation wavelength (nm)	Emission Bandpass (nm)	Laser intensity (%)
NBD	466	500-236	35
FITC	488	500-550	40
TRITC	543	550-630	40
To-Pro-3	633	650-720	35

2.3.6 2D models and 3D reconstructions of confocal images

3D reconstructions were produced using the Imaris® 7.1.0 software from Bitplane® -AG to combine the Z-stacks obtained by confocal microscopy. 2D models were prepared using the Surpass tool from Imaris® 7.1.0, after baseline correction.

2.4 Cell lysate preparations

The breast cancer cells were grown to near confluence in 175 cm^2 tissue culture flasks (Flacon) and harvested using a sterile cell scraper (Falcon) into 25 ml of PBS as described in section 2.1. Cells were centrifuged for 5 - 7 min at 500 x*g* in a bench top Centaur I MSE centrifuge before being stored as a dried pellet at - 80°C.

2.4.1 Preparation of the cytoplasmic and membrane enriched fractions

Frozen cell pellets were resuspended in 150 mM KCl and disrupted by sonication using an ultrasonic probe (MS73 Status 200) at 40% power, 10 times for 10 sec with intermittent cooling on ice cold water. Cellular debris was removed by centrifuging the cell lysate in a Sorvall Super T21 centrifuge with SL50T rotor for 20 min at 11,000 x*g* before the protein concentration was determined.

Microsomes containing membrane enriched proteins were prepared from the disrupted cell pellets by ultracentrifugation using a modified version of the procedure described by Saint-Guirons *et al* (2007) and Lehner *et al*. (2003). The preparation process is summarised in figure 2.1.

Briefly, the cells were disrupted using an ultrasonic probe (MS73 Status 200) at 40% power, 10 times for 10 sec in 40 ml of 150 mM KCl. The resulting homogenate was centrifuged at 11,000 x*g* for 20 min at 4°C to pellet the cell debris and nucleic acids. The supernatant was subsequently centrifuged for 1 h at 170,000 x*g* at 4°C (Sorvall Discovery 90SE with rotor T-865). These steps separated soluble proteins in the supernatant from the pellet containing microsomes. The pellet was re-suspended in water, sonicated as before and centrifuged for a further 1 h at 170,000 x*g* at 4°C to pellet the microsomes. The membrane-enriched proteins were then solubilised in 1 ml of U1T1 (8 M urea, 4% w/v CHAPS, 1% w/v DTT, 2% w/v ampholytes and 3 M thiourea) or urea lysis buffer (7 M urea, 4% w/v CHAPS, 1% w/v DTT) described in section 2.5.

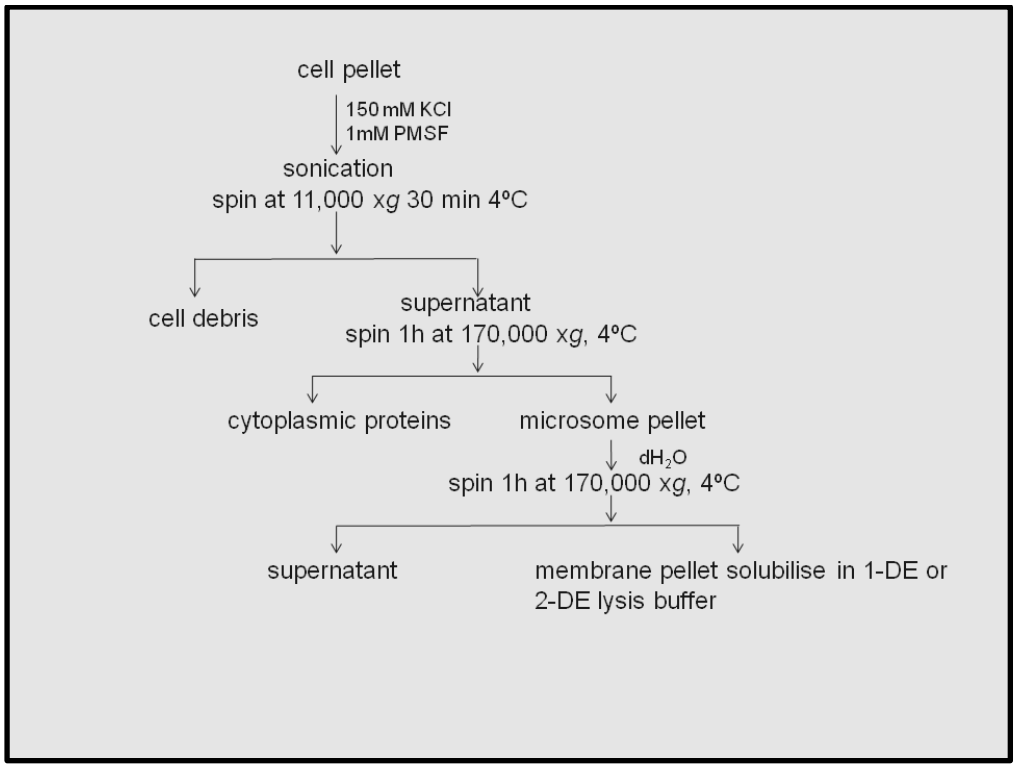


Figure 2.1: Schematic diagram showing the preparation of cytoplasmic and membrane proteins. Cells were disrupted by sonication (10 times for 10 sec) in 150 mM KCl. The cell lysate was then centrifuged at 11,000 x*g* in order to remove the cell debris and the nucleic acids from the cellular solution. The supernatant was centrifuged at high speed (170,000 x*g*) in order to pellet the cellular membranes prior to protein extraction.

2.5 Protein solubilisation

Optimisation of appropriate buffers for the solubilisation of proteins has been recognised as a key step prior to proteomic separation (Weiss & Gorg, 2008; Rabilloud, 1998).

2.5.1 Buffers for protein separation by one-dimensional electrophoresis (1-DE)

Prior to separation of membrane proteins by 1-DE the solublisation, properties of three buffers was compared. The buffer composition is shown in table 2.4.

Table 2.4: The composition of the solubilisation buffers used for 1-DE

Buffer	Composition
Chaps	1% w/v CHAPS
Thiourea	7 M urea, 4% w/v CHAPS, 1% w/v DTT, 2M thiourea
Urea	7 M urea, 4% w/v CHAPS, DTT 1% w/v

2.5.2 Buffers for protein separation by two-dimensional electrophoresis (2-DE)

A total of four solubilisation buffers were investigated for their suitability as 2-DE solubilisation buffers. The compositions of the complete 2-DE buffers are shown in table 2.5.

Table 2.5: The composition of the solubilisation buffers used for 2-DE in this study.

Buffer acronym	Composition
U	7 M urea, 4% w/v CHAPS, 1% w/v DTT, 2% v/v ampholytes and 2 M thiourea
U1	8 M urea, 4% w/v CHAPS, 1% w/v DTT, 2% v/v ampholytes and 2 M thiourea
T1	7 M urea, 4% w/v CHAPS, 1% w/v DTT, 2% v/v ampholytes and 3 M thiourea
U1T1	8 M urea, 4% w/v CHAPS, 1% w/v DTT, 2% v/v ampholytes and 3 M thiourea

2.6 Protein assay

Protein was quantitated using the Quant-iT™ Protein Assay Kit (Invitrogen, Molecular Probes). Briefly, a calibration curve of prediluted standards was set up and the Quant-iT™ reagent was diluted in Quant-iT™ buffer in a ratio of 1:199. A known volume of the sample was added to the working solution and a reading was taken using a Quant-iT™ fluorescence reader at excitation/ emission wavelengths of 470 nm and 570 nm respectively.

2.7 One-dimensional electrophoresis (1-DE)

Cytoplasmic or cell membrane proteins were prepared and assayed as described earlier and separated by SDS/PAGE according to the method of Laemmli (1970). Briefly, a suitable volume of cytoplasmic or cell membrane protein containing approximately 5-20 µg of protein in a maximum volume of 10 µl with an equal volume of 2x Laemmli reducing buffer: 125 mM Tris-HCl pH 6.8, 5% v/v glycerol, 4% w/v SDS, trace of bromophenol blue, 10% β-mercaptoethanol, was employed. The protein samples were boiled for 5 min to achieve complete denaturation and were allowed to cool to room temperature for 5 min before being loaded on an SDS-PAGE gel as detailed below.

2.8 Sodium dodecylsulphate–polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were loaded on to a 12% SDS-PAGE gel, along with protein markers (Bio-Rad or GE Healthcare) and were separated electrophoretically with running buffer; 25 mM Tris, 192 mM glycine, 0.1% w/v SDS pH 8.3 at 120 V for 1 h in the Mini Protean 3 gel system (Bio-Rad). All the buffers and recipes are summarised in Appendix 1.

2.9 Sample preparation for 2-DE

To enable a more in-depth analysis of the proteins in the samples, a combination of isoelectric focusing (IEF) followed by SDS-PAGE was utilised. This system, in a mini-gel format allows the separation of approximately 300-500 proteins in a single experiment (O'Farrell, 1975).

An appropriate volume of membrane proteins containing between 70 µg and 100 µg of protein was mixed with hydration buffer containing 7 M urea, 2 M thiourea, 4% w/v CHAPS, 1% w/v DTT, 2% v/v ampholytes (GE Healthcare, UK), to a final volume of 130 µl prior to in-gel rehydration of the immobilised pH gradient strip. DTT was added fresh to the rehydration buffer prior to use in order to prevent denaturation of the DTT.

2.9.1 In-gel rehydration

The protein samples were loaded onto 7 cm linear pH 3-10 Immobiline Drystrips (GE-Healthcare). Briefly, 130 µl of sample containing 70 µg to 100 µg of proteins prepared in rehydration buffer was pipetted into the groove of the reswelling tray. A dry IPG strip was inserted into the groove, gel face down, to cover the sample. Silicone oil (Sigma, UK) was layered on top of the strip to prevent dehydration of the strip. Rehydration was performed overnight at room temperature.

2.9.2 Isoelectric focussing (IEF)

Prior to IEF, the rehydrated strips were rinsed gently with distilled water. The strips were placed on the isoelectric focussing unit (Multiphor, GE Healthcare), gel facing up with a dampened filter paper between the gel and the electrodes to trap fast moving ions that may have caused an increase in current intensity and unwanted heating of the strip.

IEF was performed under the conditions summarised in table 2.6. Focussing was performed in two stages; in the first stage low voltage focussing was performed (300 V and 600V) to allow the removal of fast moving ions which are eventually trapped in the filter paper placed at the electrode, in the second stage focussing was performed at 3,500 V and during this step the proteins move toward the anode/cathode depending on their charge. The proteins eventually come to rest at the pH zone equal to their pKa (O'Farrell, 1975).

Table 2.6: The conditions of IEF set-up of the proteins loaded onto 7 cm IPG strip.

Temperature: 20°C
Current max: 2 mA
Voltage max: 3500 V

IEF	Volts	Time
Initial steps	300 V	30 min
	600 V	30 min
Focussing step	3,500 V	2 h 45 min

2.9.3 Equilibration of IPG strip

Focussed strips were maintained at -80^{o}C for no longer than 2 days. Prior to separation of proteins in the second dimension, the strips were equilibrated for 15 min in equilibration buffer: 50 mM Tris-HCl, pH 8.8; 6 M urea, 30% v/v glycerol, 2% w/v SDS, 1% w/v DTT, to reduce disulphide bonds and unfold the proteins to enable their exit from the IPG strips. This step was followed by another 15 min step in a second equilibration buffer: 50 mM Tris-HCl, pH 8.8; 6 M urea, 30% v/v glycerol, 2% w/v SDS, 1% DTT, 2.5 % w/v iodoacetamide to stabilise the unfolded proteins and prevent protein re-folding.

2.9.4 Second dimension: SDS-PAGE

Equilibrated strips were then gently placed on top of a vertical 1 mm thick, 8 x 7 cm, 10% SDS-PAGE gel. Molecular weight standards were loaded on a small piece of filter paper and inserted next to the positive end of the IPG strip. A sealing buffer: 0.5% w/v agarose, 25 mM Tris pH 8.3, 192 mM glycine, 0.5% w/v SDS, trace of bromophenol blue, was poured over the strip to seal the the IPG strip and prevent movement during protein separation. The gels were electrophoresed for 2 h at 120 V prior to protein staining or Western blotting.

2.10 Protein staining using Coomassie brilliant blue (CBB)

The protein gels were stained with Coomassie brilliant blue (CBB). The fixative, stain and destain solutions were freshly prepared and are shown in table 2.7. The gels were fixed for 30 min in methanol/ acetic acid; stained with CBB for 2 h and destained for 2 h. Complete destaining of gels was achieved by allowing the gels to stand overnight in water until the protein bands were stained deep blue against a transparent background. All these steps were performed under gentle agitation on a rocking tray at room temperature.

Table 2.7: Coomassie brilliant blue staining solutions used for fixing, staining and destaining.

Solutions	Composition	Time
Fixing solution	50% v/v methanol, 10% v/v acetic acid, dH_2O	30 min
Staining solution	10% v/v acetic acid, 0.025% w/v Coomassie brilliant blue, dH_2O	2 h
Destaining solution	10% v/v acetic acid, dH_2O	2 h

2.11 Protein transfer by Western blotting

The proteins separated by 1-DE or 2-DE were transferred onto nitrocellulose membranes (GE Healthcare) by wet transfer in a Mini Trans Blot transfer cell, Bio-Rad, using transfer buffer: 25 mM Tris, 192 mM glycine and 20% v/v methanol at 200 V for 2 h. The system was cooled throughout the experiment.

2.12 Ponceau S staining of Western blots

Prior to lectin or antibody probing, the membranes were stained with the protein dye, Ponceau S (Sigma, UK) to check the efficiency of the protein transfer and to localise the molecular weight markers. 0.1% w/v Ponceau S in 5% v/v acetic acid was used for 5 min for staining and was subsequently removed by rinsing the blot in water.

2.13 Lectin and antibody detection on Western blots

Prior to probing the blot with lectin or antibody, the nitrocellulose membrane was blocked with 5% w/v BSA in phosphate buffered saline/0.05% v/v Tween, pH 7.6 (PBS/T) for 2 h. After the blocking step, the membrane was washed 3 times for five min in PBS/T. Lectin/antibody steps and washing steps were carried out in PBS/T and were performed at room temperature with gentle rocking.

2.13.1 Lectin blotting with HPA

After blocking, the blots were incubated with biotinylated lectin (Sigma, UK) and streptavidin-HRP (Thermofisher Scientific, UK). Initially, the chromogenic substrate diaminobenzidine (DAB) was used for detection of lectin binding. Biotinylated lectin and streptavidin-HRP were prepared in PBS/T and used at 5 µg/ml for 2 h and 2 µg/ml for 1 h respectively. The blots were washed 5 times for 5 min in PBS/T between each step. DAB was prepared in PBS/ H_2O_2 (6 mg of DAB prepared in 9 ml of TBS and 60 µl of 30 vol H_2O_2). The blots were incubated with DAB for 5 min. The reaction was stopped by addition of water.

Due to the relatively high cost of HPA, a more sensitive method using less lectin was developed, this was based on using enhanced chemiluminsesce (ECL). The conditions used were optimised by using varying concentrations (between 0.125-1 µg/ml) of biotinylated lectin and streptavidin-HRP. After the optimisation steps had been completed the final protocol used was as follows: the protein blots were incubated with 0.5 µg/ml biotinylated lectin for 2 h and 0.125 µg/ml of streptavidin-

HRP for 1 h (Pierce, UK), the blots were washed 5 times for 5 min in PBS/T between each step and incubated for 3 min with Super Signal West Pico (Thermo Fisher Scientific, UK) substrate according to the manufacturer's instructions. Detection was performed by exposure of an X-ray film for between 30 sec to 1 min and subsequent development of the film for 30 sec.

2.13.2 Inhibition of HPA binding to T47D membrane proteins

T47D membrane proteins separated by 1-DE and transferred to nitrocellulose membranes were blocked with BSA (as above) and incubated for 2 h with 5 µg/ml biotinylated HPA in PBS/T which has been preincubated for 30 min with freshly prepared 100 mM GalNAc, GlcNAc, Gal and Man monosaccharide (Sigma, UK). The membrane was washed in PBS/T and incubated with streptavidin-HRP before detection with DAB as before. 2-DE separated T47D membrane proteins were transferred to nitrocellulose and blocked as above, the blots were incubated with HPA prepared with 100 mM GlcNAc as for the 1-DE and detection was performed using the Super Signal West Pico reagent.

2.13.3 Probing Western blots with anti-integrin α6 (anti-CD49f) antibody

Proteins were separated by SDS-PAGE, transferred to nitrocellulose and blocked in BSA (as before) and were probed with a mouse monoclonal anti-integrin α6 antibody (sc-59971, Santa-Cruz Biotech, UK) (Horton *et al*, 1985; Hynes, 1992; Levy *et al*., 2000), prepared in PBS/T at 20 µg/ml for 2 h. The membrane was washed 5 times for 5 min in PBS/T, this was then followed by 1 h incubation with 10 µg/ml goat anti-mouse IgG-HRP (Santa-Cruz Biotech, UK). Detection was performed using the Super signal West Pico chemiluminescent reagent as before.

2.13.4 Probing Western blots with anti-blood group A antibody

Proteins were separated by SDS-PAGE, transferred to nitrocellulose and blocked in BSA (as before) and probed with a murine monoclonal antibody directed against the blood group A antigen (Ortho-Diagnostics, Johnson and Johnson, kindly provided by Dr P Greenwell, University of Westminster) prepared in PBS/T at 20 µg/ml for 2 h, the blot was then washed 5 times for 5 min in PBS/T, followed by 1h incubation with 10 µg/ml of goat anti-mouse IgM-HRP (Sigma, UK). Detection was performed using the Super Signal West Pico chemiluminescent reagent as before.

2.13.5 Probing Western blots with anti-O-GlcNAc antibody

Proteins were separated by SDS-PAGE, transferred to nitrocellulose and blocked in BSA (as before) and were probed with a murine monoclonal antibody directed against the O-GlcNAc epitope (sc-81483, Santa-Cruz Biotech, UK) prepared in PBS/T at 20 µg/ml for 2 h, the blot was then washed 5 times for 5 min in PBS/T (Akimoto *et al*., 2003; Haltiwanger *et al*., 1992; Shafi *et al*., 2000), followed by 1h incubation with 10 µg/ml of goat anti-mouse IgG-HRP (Santa-Cruz, UK). Detection was performed using the Super signal West Pico chemiluminescent reagent as before.

2.14 Data analysis of 1-DE and 2-DE

2.14.1 Digital image processing

Digital images of the 1-DE, 2-DE separations and Western blots were obtained using a BioRad GS-800 densitometer. Data was analysed using the BioRad Quantity One computer software package. A lane profile and intensity histogram was obtained with relative intensity values for each of the bands detected. Images in TIFF format were imported into the Progenesis Same Spots system, Version 3.0 (Non–Linear Dynamics, UK) and processed by background subtraction, spot detection, landmarking and overlay analysis.

2.14.2 Reproducibility of the 2-DE system

The robustness of the 2-DE system was assessed by analysing gel to gel variation in analytical (protein gels from same passage of the cell lines) and biological repeats (protein gels from different passage of the cell lines). An Anova test was performed using 6 replicate gels. An Anova test and fold difference values were generated by the Progenesis Same Spots Software system Version 3.0. Coefficient of variation (CV) was calculated as a ratio of the standard deviation (σ) to the mean (μ).

2.15 Spot picking of 2-DE separated HPA binding proteins

HPA binding membrane proteins of T47D (identified on the X-ray film exposed to the Super Signal reagent following the Western blotting step) were used as a template to locate the equivalent protein species on a CBB stained gel; run at the same time as the Western blot. HPA binding proteins of interest were manually excised from the CBB stained gel using a clean scapel and stored at -80^{o}C in 0.5 ml sterile Eppendorf tubes.

2.16 Protein identification by MALDI-TOF Mass Spectrometry

Protein identification was performed by commercial arrangement with Dr. Thomas, Department of Biology, University of York, using a MALDI-TOF/TOF Applied Biosystems, 4700 analyser.

Briefly, the proteins were reduced using DTT, S-carbamidomethylation and iodoacetamide prior to tryptic digestion. The excised protein gels pieces were washed thee times in 50% v/v acetonitrile/ 25 mM ammonium bicarbonate and air dried before rehydration in 10 ml of 20 μg/ml sequencing-grade, modified porcine trypsin (Promega, UK). Protein digestion was performed overnight at 37^{o}C. A 0.5 μl aliquot of each tryptic digest and 0.5 μl of a solution of α-cyano-4-hydroxycinnamic acid (CHCA) (Sigma, UK) in 50% v/v acetonitrile containing 0.1%

v/v trifluroacetic acid was applied to the MALDI target plate. Mass spectra were obtained in the reflection mode with an accelerating voltage of 20 kV. The peptide mass fingerprint (PMF) generated was compared to the masses of all theoretical tryptic peptides generated *in silico* by the MASCOT search program updated in 2010 (Pappin *et al*., 1993). Collision-induced dissociation MS/MS was also performed to corroborate the significant matches from the MALDI/MS.

2.17 Post-translational modification (PTM) prediction

Potential N-linked, O-GalNAc, O-GlcNAc and O-phosphate sites in the HPA binding proteins were predicted using the following web servers: http://www.cbs.dtu.dk/services/NetNGlyc/ , http://www.cbs.dtu.dk/services/NetOGlyc/, http://www.cbs.dtu.dk/services/YinOYang/ and http://www.cbs.dtu.dk/services/NetPhos/ respectively.

2.18 Genomic studies

Sterile plastic-ware and RNAse/DNAase free water and reagents were used throughout.

2.18.1 mRNA extraction

Cellas use for mRNA extraction were grown and collected as described in section 2.1.Approximately 10,000 cells were used from each cell line for the mRNA extraction steps. mRNA extraction was undertaken using the RNeasy Mini extraction kit (Qiagen, UK) according to the manufacturer's instructions. Briefly, the pelleted cells were loosened by vortexing and pipetting. A suitable volume (~400 µl) of lysis buffer (RLT) was added to the cell pellet. The cells were resuspended by vortexing for 1 min and pipetted directly into a QIAshedder spin column placed in a 2 ml collection tube. The lysate was centrifuged for 2 min at full speed. An equal volume of 70% v/v ethanol was added to the homogenised lysate and mixed by pipetting. The resulting lysate (~700 µl) was pipetted into an RNeasy spin column placed in a 2 ml collection tube and centrifuged at 8,000 x*g* for 15 sec. The flow-through was discarded. Approximately 700 µl of wash buffer (RW1) was then added to the RNeasy column and the column was centrifuged at 8,000 x*g* for 15 sec to wash the membrane. A further 500 µl of a second wash buffer (RPE) was added to the RNeasy column and centrifuged at 8,000 x*g* for 15 sec again to wash the membrane of the column. This step was repeated for 2 min. Lastly, the spin column was placed in a 1.5 ml collection tube and 50 µl of RNase-free water was added directly to the column and was centrifuged at 8,000 x*g* for 1 min to elute the RNA.

2.18.2 mRNA quantification and purity

The mRNA concentration was determined in 10 mM Tris-HCl pH 7.5, taking the optical density (OD) at a wavelength of 260 nm using a quartz cuvette in a Lambda 35 spectophotometer (PerkinElmer, UK). The RNA concentration was calculated using the following equation

$$\text{RNA concentration (µg/ml)} = \frac{\text{OD260 X 40 (dilution factor) X 50 µg/ml}}{1000}$$

To determine the extent of protein contamination the absorbance at 280 nm was measured. The RNA preparations used in this study had an OD_{260}/OD_{280} in the range of 1.9-2.1.

2.18.3 Reverse transcription

Reverse transcription of the purified mRNA was performed using the Quantiscript Reverse transcription kit (Qiagen, UK) and Oligo-dT (5'TTTTTTTTTTTTTTTTTTTTV3') primers, according to the manufacturer's instructions. Briefly, in order to eliminate genomic DNA, 2 µg of RNA was added to 12 µl of gDNA wipeout buffer and incubated at 42^{o}C for 2 min. This mixture containing the template RNA was then mixed with the reverse transcription master mix containing 1 µl of Quantiscript Reverse Transcriptase, 4 µl of Quantiscript RT buffer and 1 µl of RT primer mix. This mixture was incubated at 42^{o}C for 15 min to allow cDNA transcription followed by incubation at 95^{o}C for 3 min to inactivate Quantitect Reverse Transciptase.

2.18.4 cDNA quantification and purity

The DNA concentration was determined in TE buffer (10 Mm Tris-HCl, pH 7.5, 1mM EDTA) taking the optical density (OD) at a wavelength of 260 nm using a quartz cuvette and a lambda 35 spectophotometer (PerkinElmer, UK). The DNA concentration was calculated using the following equation

$$\text{DNA concentration (µg/ml)} = \frac{\text{OD260 X 50 (dilution factor) X 50 µg/ml}}{1000}$$

To determine the extent of protein contamination the absorbance at 280 nm was measured. The DNA preparations used in this study had an OD_{260}/OD_{280} in the range of 1.7-2.0.

2.18.5 Primer design

Primers for the *ppGalNAc T1, T2, T3* and T*6* genes used in this study were purchased from Qiagen, UK, therefore the sequences of primers are currently not available as these are proprietary information. Areas of the genes targeted in the PCR experiments are, however, shown in Appendix 7. Primers for *ST6GalNAc I/II* and *β-actin* genes was designed by using Basic local Alignment Search Tool (BLASTn) on http://www.ncbi.nlm.nih.gov/tools/primer-blast/ website and obtained from MWG-Biotech, UK. The primer sequences are shown in table 2.8.

Table 2.8: Primer sequences of for ST6GalNAc *I/II* and *beta-actin* genes

Genes	Entrez gene ID	Forward and Reverse primers
ST6GalNAc *I*	55808	F 5'-GCAAAGCGGCAACCACAGCC-3'
		R 3'CTGCTGGGGCACTGGAGGGA-5'
ST6GalNAc *II*	10610	F 5'- CACCTGGCCATTCAGCGGCA-3'
		R 3'- GTGGCACGGAGGTGAAGCCC-5'
beta actin	126867	F 5'CCAGACAGCACTGTGTTGGC3'
		R 5'GAGAAGCTGTGCTACGTCGC3'

2.18.6 Quantitative Real time, two step RT-PCR

After the RNA was reverse transcribed to cDNA using the Quantitect Reverse Transcriptase Kit (Qiagen, UK) the subsequent cDNA was used for quantitative PCR using the QuantiFast SYBR green PCR kit (Qiagen, UK). Briefly, 100 ng of cDNA was added to PCR vessels containing a mixuture of 12.5 µl 2X QuantiFast SYBR Green Master mix, up to 1 µM of forward and reverse primers. The Rotor Gene time cycler (Qiagen, UK) was used for this experiment. The Real-time cycler conditions were set as detailed in table 2.9.

Table 2.9: The Real-time PCR cycler conditions. N/A = not applicable

Step	Time	Temperature
PCR initial activation step	5 min	95^{o}C
Denaturation	10 s	95 ^{o}C
Combined annealing/extension	30 s	60 ^{o}C
Number of cycles	35-40	N/A

2.18.7 DNA separation by agarose gel electrophoresis

The PCR products were separated by agarose gel electrophoresis (Ogden & Adams, 1987) using a mini-Sub cell GT apparatus (Bio-Rad, UK). The composition of the buffers used is shown in Appendix 1. Briefly, to prepare a mini gel 2% w/v agarose was weighed into 250 ml conical flasks dissolved in 30 ml of TAE buffer: 40 mM Tris-HCL, pH 8.2, 1 mM EDTA and 20 mM acetic acid by microwaving for approximately 1 min (Stellwagen & Stellwagen, 2002) and was allowed to set for 30 min. The gel was then placed in a running tank with sample wells prepared at the cathode end and the tank was filled with 250 TAE buffer. Aprroximately, 10 µl of DNA was mixed with 2 µl of concentrated sample buffer: 0.4% w/v orange G, 0.03% w/v bromophenol blue, 0.03% w/v xylene cyanol FF, 15% w/v 400 in TAE buffer, pH 8.2 and loaded into the well of the gel. Molecular weight DNA ladders (Promega, UK) were run at the same time as the samples. Electrophoresis was performed at 120 V for 1 h. After electrophoresis, the gel was incubated for 30 min

in 1% w/v ethidium bromide solution. The DNA fragments were detected using a UV transiluminator (Ultra-Violet Products Ltd, UK) using a wavelength of 200-280 nm.

2.18.8 DNA purification from the agarose gels

For DNA purification, the area containing the desired DNA fragment was excised from the agarose gel with a clean scalpel and the DNA was purified using a QIAquick gel extraction kit (Qiagen, UK), according to the manufacturer's recommendations. Briefly, the DNA fragment was excised from the gel and minced with 300 µl of QG buffer pH 5.5, guanidine thiocyanate, 20mM Tris-HCl, pH6.6, 0.0025% cresol red (Sigma, UK), in an Eppendorf tube. The tube was incubated at 50^{o}C for 10 min followed by the addition of 100 µl of isopropanol to the sample. The sample was then pipetted into the QIAquick column and centrifuged for 1 min at 14,000 x*g*. The flow-through was discarded and the QIAquick column and 500 µL of QG buffer added to the column and centrifuged for 1min at 14,000 x*g*. To wash the column, 0.75 ml of PE buffer was added to the QIaQUICK column and centrifuged for 1 min at 14,000 x*g*. The flow-through was discarded and the QIAquick column was placed back into the same tube and centrifuged for an additional 1 min at 14,000 x*g*. To elute the DNA, the QIAquick column was placed in an Eppendorf tube, 30 µl if MilliQ water was added, incubated for 1 min at 20^{o}C and then centrifuged for further minute at 14,000 x*g*. The purified DNA was sent for sequencing at MGW, UK.

Chapter 3

Evaluation of HPA binding at cellular and protein level

3.0 Evaluation of HPA binding at cellular and protein level

3.1 Introduction

The glycosylation of cell surface proteins is a post-translational modification which plays significant roles in mediating cellular functions such as correct protein folding (Walsh *et al*., 1990), cell to cell adhesion (Lasky,1992; Springer, 1990), cell to cell communication (Wassarman, 1990), protection of proteins from enzymatic degradation (Homans *et al*., 1987) and signal transduction (Haltiwanger *et al*., 2002). It is widely accepted that altered glycosylation of cell surface glycoproteins accompanies malignant transformation (Brockhausen, 2006; Dennis *et al.*, 1999; Hakomori, 1984). Lectin based studies have enabled the identification of abrogated glycan structures accompanying the physiopathological development of cancer (Aub *et al*., 1965; Burger & Goldberg, 1967). A positive correlation between HPA staining of tissue sections and poor prognosis was established more than a decade ago by Leathem & Brooks (1987) and has since been reported by independent researchers in other cancers such as oesophageal (Takashi *et al*., 1994; Yoshida *et al*., 1994), gastric (Kakeji *et al*., 1994), prostate (Shirashi *et al*., 1992), lung (Kawai *et al*., 1991) and colorectal (Ikeda *et al*., 1994; Schumacher *et al*., 1992). Other lectins with nominal GalNAc binding properties do not bind poor prognosis cancer in the same way as HPA (Sharma & Surolia, 1997). The utility of HPA for detecting poor prognosis cancer appears to be come from its ability to simultaneously recognise a number of proteins which may directly be implicated in the metastatic process.

Saint-Guirons *et al* (2007) showed that HPA bound a multitude of proteins in metastatic HT29 colorectal cell lines. The authors used fluorescently labelled HPA and revealed cell surface localisation of HPA binding in colorectal HT29 cells. The work of Saint-Guirons *et al* (2007) in our laboratory used 2-DE coupled with MS and demonstrated that HPA binds simultaneously to many glycoproteins which has previously been associated with metastatic spread and poor prognosis cancer. The HPA binding proteins in CRC includes molecules involved in cell migration and adhesion (integrin αV/α6 and annexin A2/A4), anti-apoptotic pathways (Hsp-90,

Hsp-96 and TNFR-1) and cellular remodelling (αβ tubulin, actin, and cytokeratins 8 and 18). Integrins have been implicated in promoting tumour cell invasion through the ECM (Chao *et al.,* 1996). This study therefore assisted in the discovery of membrane glycoproteins with aberrant glycosylation in HT29 CRC cells. This is a key reference for the work carried in this thesis, as a similar approach was used in the current investigation to identify HPA binding partners in metastatic breast cancer.

Schumacher *et al* (1995) were the first to investigate the cell surface HPA binding glycoproteins in human breast cancer cell lines. Cell membrane glycoproteins were isolated and analysed by 1-DE SDS-PAGE with lectin blotting. The Western blot analysis revealed that HPA bound to several membrane glycoproteins, however, the identities of these proteins remained unknown. Mitchell *et al* (1995) reported seven HPA-binding bands ranging in molecular weight from 20–90 kDa in several breast cell lines and tentatively identified the band at 90 kDa, as the transferrin receptor. In 2001, Brooks *et al* (2001) characterised a range of human breast cell lines, normal and malignant, for the synthesis of HPA-binding ligands. The cell lines used in the study were HMT3522 (derived from fibrocystic disease); BT474 (derived from primary breast cancer); MDA MB435, MDA MB 468, MCF-7, ZR-751, DU4475 (all derived from metastatic breast cancers). Clinical breast tumour samples were also analysed to determine if tumour samples and cell lines produce the same HPA-binding glycoproteins. Light and confocal microscopy revealed that the cell lines exhibited varying degrees of HPA binding ranging from negligible in HMT3522 (‘normal’/benign disease), and weak in BT474 (primary cancer) to very intense in MCF-7 and T47D (metastatic cancer). Analysis of HPA-binding glycoproteins by Western blotting revealed eleven prominent HPA-binding glycoproteins in all of the metastatic cell lines and these matched those found in clinical tumour samples. These proteins were evaluated with the reference to their molecular weight but there remained an opportunity to identify the HPA binding glycoproteins of breast cancer.

The current study employed four cell lines (HMT3522, BT474, MCF-7 and T47D) to represent a range of phenotypes, from normal to highly metastatic. These human breast cell lines have defined HPA-binding properties and metastatic

characteristics when implanted into SCID mice (Schumacher & Adam, 1997), therefore were a suitable *in vitro* model for cytochemical analysis of HPA binding in breast cancer.

Brooks *et al* (2001) have developed a scoring system in an attempt to quantify the lectin histochemistry. The system uses a scale ranging from "–'' for negative HPA binding and "++++'' for strong HPA binding. The scoring for HPA labelling was given as follows: HMT3522 (-), BT474 (+) and MCF-7 (++++).

Schumacher and Adam (1997) were first to establish a correlation between HPA binding and an *in vivo* model of metastasis using human breast cells. The work found that HPA positive breast cancer cells, transplanted subcutaneously into SCID mice, metastasised spontaneously to the lungs of the mice, whereas HPA negative cell lines generally did not metastasise, emphasising the clinical relevance of this model and the utility of the breast cancer cell lines used in this investigation.

In this study, I initially sought to confirm the observations of Brooks *et al* (2001) for the HPA binding properties of the breast cell lines. Secondly, the proteins of the cells were separated and transferred to nitrocellulose to localise HPA binding glycoproteins in the cytoplasmic and membrane proteins fraction of the breast cancer cell lines. The specificity of HPA binding to T47D was also considered and was evaluated by competitive inhibition, by preincubation of the lectin with monosaccharides and using both confocal microscopy and protein separation.

3.2 HPA cytochemistry

The initial work was undertaken to assess HPA binding of the breast cancer cell lines using light and fluorescence microscopy. The breast cell lines were grown to near confluence overnight in 6-well plates and fixed in formalin to mimic tissue processing prior to lectin staining. The images from the microscopy experiments were processed using the Scope Photo and Imaris software package version 7.1.0). The methods used are described in section 2.2 and 2.3.

3.2.1 HPA binding to the breast cells

The breast cells HMT3522, BT474, MCF-7 and T47D were evaluated for HPA binding using a colorimetric method previously employed by Brooks and Hall (2002). Biotinylated HPA was used in conjunction with streptavidin-HRP. Diaminobenzinidine (DAB) / hydrogen peroxide (H_2O_2) was employed as the substrate.

In this experiment, marked quantitative differences in HPA staining was observed across the breast cell lines (figure 3.1). The cells of metastatic origin, T47D and MCF-7, exhibited intense and moderate HPA labelling respectively, whereas weak or 'negligible' binding was noted in 'normal'/benign (HMT3522) and non-metastatic (BT474), consistent with the phenotype of these cell lines. In experiments where the lectin incubation step was omitted, HPA staining was non detectable.

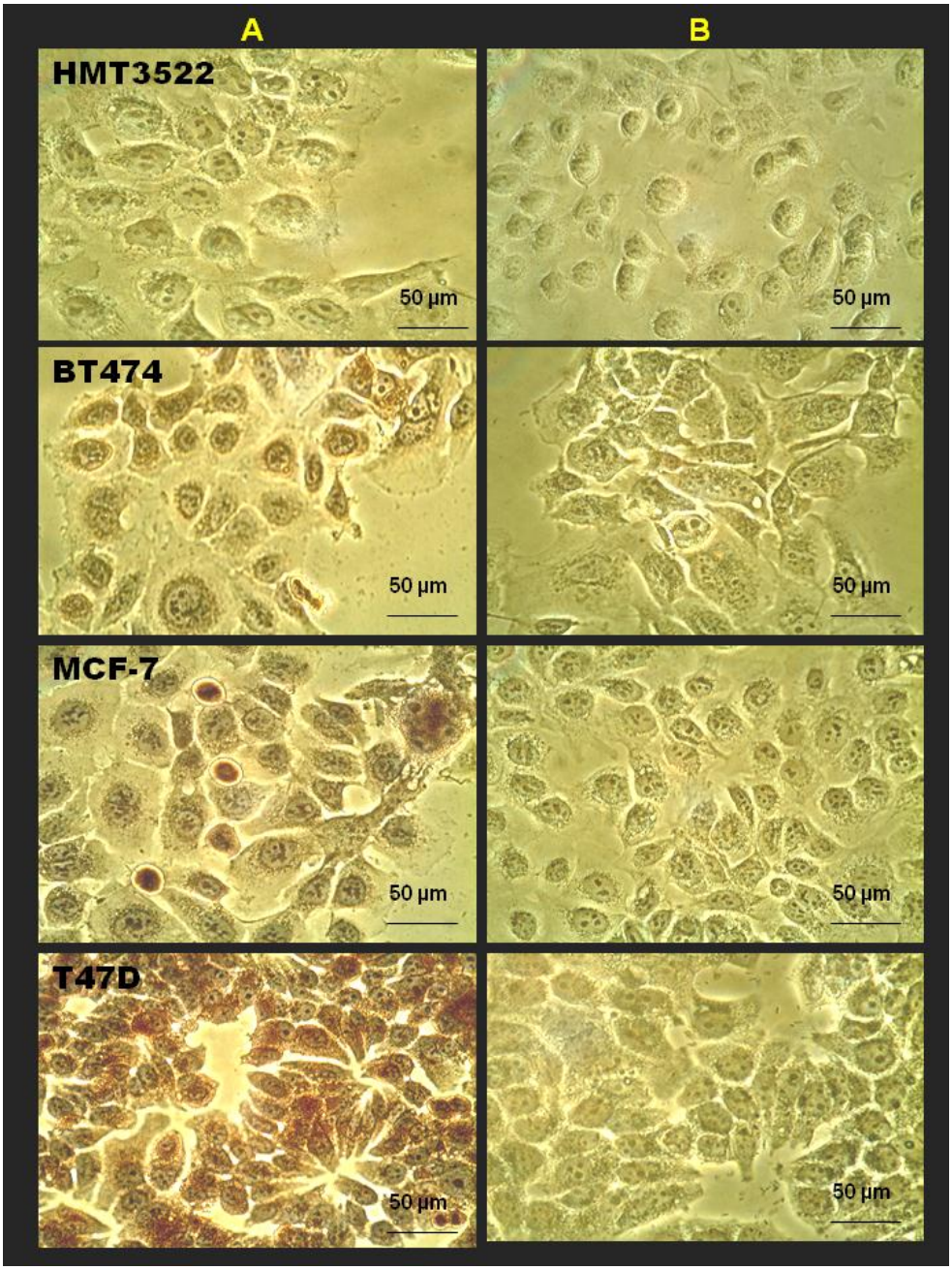


Figure 3.1: Light microscopy image of HPA binding to breast cancer cell lines. : Panel A: mages of breast cancer cells incubated with biotinylated HPA (10 µg/ml) and streptavidin-HRP (5 µg/ml). The brown colouration shows the peroxidase reaction with DAB/H_2O_2. Levels of labelling ranged from 'negligible' or weak in HMT3522, and BT474, moderate in MCF-7 to intense in T47D. Panel B: Images of breast cells incubated with streptavidin-HRP (5 µg/ml) alone showed no significant staining. Scale bars= 50 µm

To obtain a more detailed evaluation of HPA binding in the 'normal'/benign (HMT3522), non-metastatic (BT474) and metastatic (MCF-7 and T47D) cells, the cell preparations were incubated with TRITC labelled HPA and observed using confocal microscopy. The aim of this work was to confirm previous observations of cytoplasmic and cell surface labelling in the breast cells (Brooks *et al*., 2001).

This system allowed a detailed analysis of HPA binding to the cell lines (figure 3.2). HPA bound intensely to T47D at the cell surface and, in addition, granular and perinuclear staining was also observed. Moderate perinuclear staining were noted in the MCF-7 cells, whilst very low levels or 'negligible' HPA binding were detected in HMT3522 and BT474. These data were consistent with the observations made using light microscopy where intense HPA labelling was observed in T47D, moderate in MCF-7 and weak or negligible binding were noted for BT474 and HMT3522 experiment MCF-7.

In this experiment the use of HPA-TRITC to study lectin interaction with the breast cells was defined and the results obtained confirmed observations of HPA binding made by Brooks *et al* (2001) and Saint Guirons *et al* (2007) in breast and colorectal cancer cells respectively.

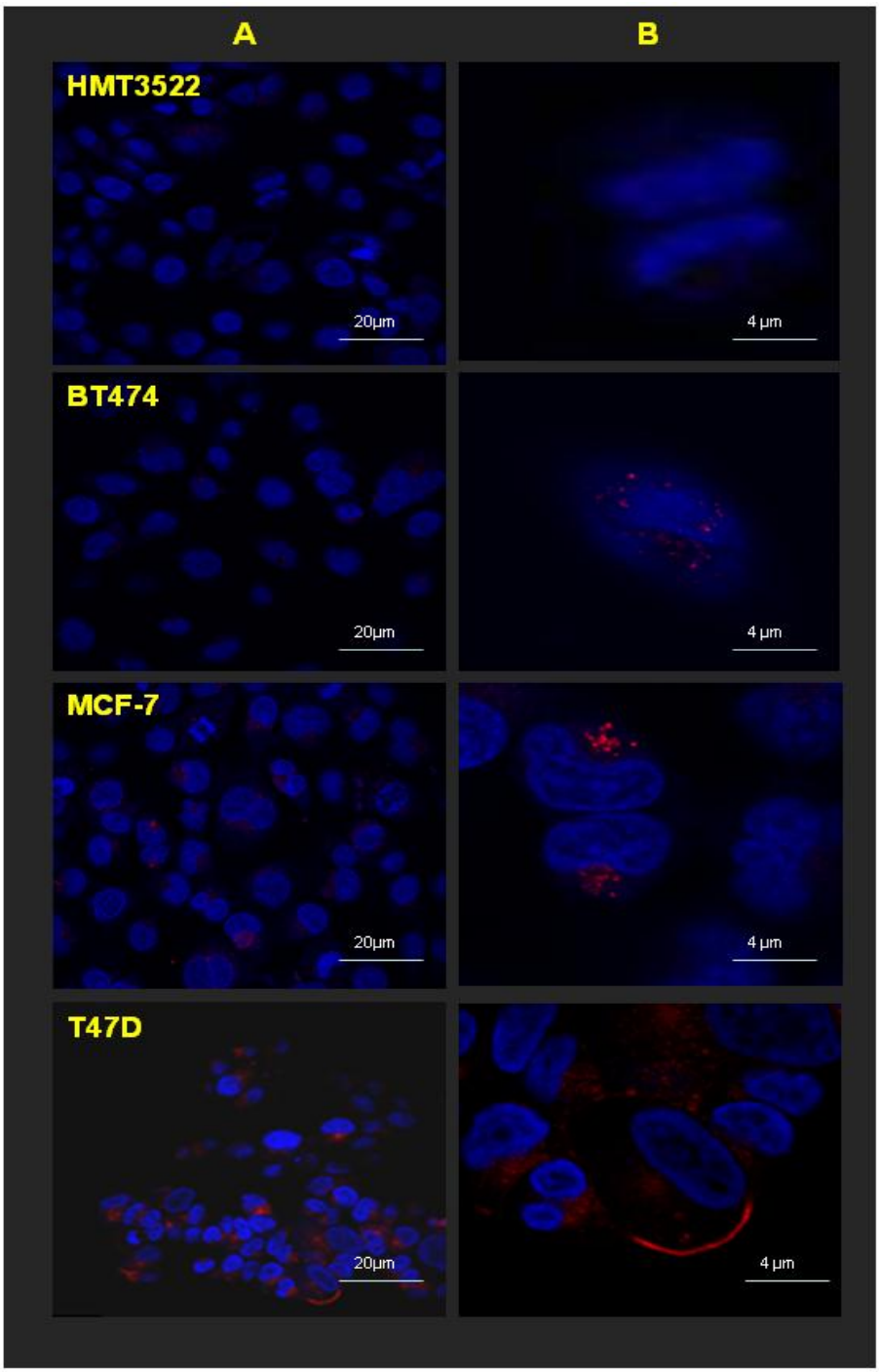


Figure 3.2: Confocal images of HPA labelling in breast cell lines. Images showing the binding of TRITC-labelled HPA (10µg/ml) (red) to breast cancer cell lines, counter-stained with the nuclear label To-Pro-3 (blue). Images showing HPA binding in HMT3522, BT474, MCF-7 and T47D, Panel A: scale bars = 20 µm and Panel B: scale bars= 4 µm. HMT3522 and BT474 showed occasional or 'negligible' granular staining, whereas T47D showed intense HPA binding on the cell surface with some granular intracellular staining. MCF-7 showed intense granular staining in the perinuclear region perhaps representing the binding of lectin to glycoproteins transiting through the Golgi apparatus.

3.2.2 HPA binding to T47D cells: 3D reconstruction of confocal images

The cellular distribution of HPA binding ligands in a subpopulation of T47D cells was assessed by generating 3D models of HPA binding in T47D cells. Z-stack sections of HPA stained T47D cells were taken at 4-5 μm increments, the sections were combined to generate a 3D representation of HPA binding using the Imaris Software package (version 7.1.0) (figure 3.3).

Whilst intense HPA labelling was observed in selected T47D cells, a somewhat heterogenous distribution of HPA binding ligands was observed. Some cells exhibited granular and perinuclear staining only and some showed cell surface, in addition to cytoplasmic labelling. Overall the 2D images showed intense cell surface and cytoplasmic staining in T47D (figure 3.3 A), however the 3D model revealed intense perinuclear staining in this cell line (figure 3.3 B). The 3D model offered a spatial view of the localisation of HPA binding within a cell. The white and yellow arrows in figure 3.3 B show perinuclear and cell surface labelling respectively. In this way HPA binding epitopes were observed in variable levels across a subpopulation of T47D cells.

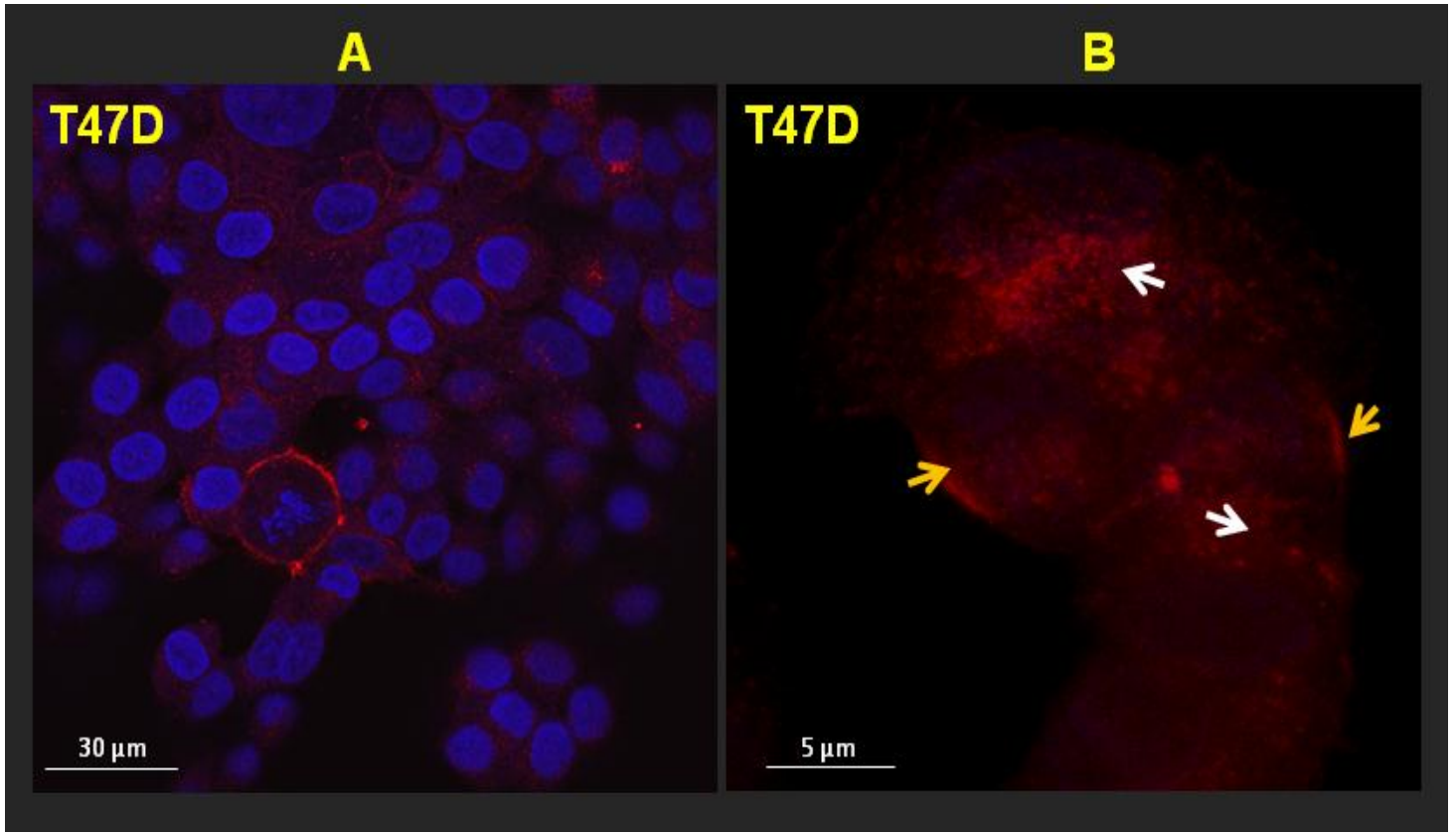


Figure 3.3: HPA binding in subpopulations of T47D cells. The Imaris software ''Surpass view'' shows a 2D and 3D view of cell surface and intracellular HPA binding in T47D. Panel A: 2D view of HPA in T47D. Panel B: 3D representation of HPA binding in six T47D cells. The imaris software ''easy 3D view'' showed cell surface (yellow arrow) and perinuclear (white arrow) in a subpopulation of cells. The 3D view offers a deeper insight into the cellular localisation of HPA labelling withing a small population of cells. Panel A: scale bars = 30 µm and Panel B: scale bars= 5 µm.

3.3 Establishing conditions for evaluating HPA binding glycoproteins in breast cells using SDS-PAGE and Western blotting

One of the main aims of this work was to identify HPA binding glycoproteins in breast cancer cell lines and compare them to those previously reported for CRC. Fractions containing cytoplasmic and membrane proteins were collected following ultracentrifugation (as described in section 2.4.1). In each experiment, 10 µg to 20 µg of proteins were loaded in the wells of a gel and separated by SDS-PAGE. The gels were stained with Coomassie brilliant blue (CBB). A replicate gel was transferred onto a nitrocellulose membrane and probed with biotinylated HPA followed by streptavidin-HRP for the detection of HPA binding glycoproteins (as described in section 2.13.1).

In the first part of this study, a method was established for the detection of HPA binding glycoproteins by probing Western blots with biotinylated HPA and streptavidin-HRP. A comparison of two detection systems was undertaken; using a colorimetric based method: DAB/H_2O_2 and a second system based on enhanced chemiluminescence (ECL). The optimal buffers for solubilisation and separation of membrane proteins prior to analysis by SDS-PAGE were assessed. Lastly, the HPA binding glycoproteins in the membrane fraction across the different breast cell lines was analysed.

3.3.1 Detection of HPA binding to cytoplasmic proteins with diaminobenzinidine (DAB) substrate and enhanced chemiluminescence (ECL) systems

In this first part of the work the DAB and ECL detection systems were compared. A key consideration was to minimise the use of lectin in Western blotting analysis and at the same time there was a need to develop a sensitive method for the detection of low quantities of HPA binding glycoproteins on the Western blots.

Figure 3.4 shows a CBB stained gel of the fraction of cytoplasmic proteins from the four breast cell lines. The protein levels were consistent and proteins were homogeneous across the cell lines. In contrast, the Western blots revealed heterogenous levels of HPA binding glycoproteins, with HMT3522 and BT474 having the lowest levels of HPA binding glycoproteins and the cell lines derived from metastatic phenotype showing a significantly higher number of intense bands that bound the lectin. The observation was consistent across both the colorimetric and chemiluminesent detection systems. The Western blot probed with HPA and detected using the DAB substrate showed six major glycoproteins species in the range of approximately 28-102 kDa that interacted with the lectin in T47D and five in MCF-7, all of which, except protein bands 1 and 2, were detectable in HMT3522 and BT474, althrough the intensity of the HPA binding varied across the cell lines, with the highest protein band intensity noted in T47D, moderate amounts in MCF-7 and lowest intensity in HMT3522 and BT474.

A similar HPA binding pattern was observed using the ECL detection system. All the bands observed in the DAB blot were also identified on the chemiluminesence X-ray film. However, two additional bands at a lower molecular weight (bands 7 and 8) were noted with the ECL detection system. Moreover, the HPA binding protein species appeared as intense bands. In the case of T47D, band 1 and 2 were not dectable as discrete bands, instead numerous bands bound HPA in this region, perhaps suggesting that HPA recognises a multitude of lower abundance proteins of similar molecular weight which were only detected by the ECL system. The improved sensitivity of the ECL system was confirmed by the dectection of protein band 1 which was observed in the ECL system but not the DAB system for BT474 and HMT3522. In experiment where the lectin step was omitted and streptavidin-HRP was incubated alone, no bands were detected (data not shown). The estimated molecular weight of the HPA binding glycoproteins is displayed in table 3.1.

In summary, this experiment confirmed that HPA glycoproteins are present in higher levels in the cells associated with a metastatic phenotype, (MCF-7 and T47D). This study also showed that the ECL method offers improved detection of

low abundance HPA binding glycoproteins and gives an insight into the HPA glycoproteins in the cancer cell lines.

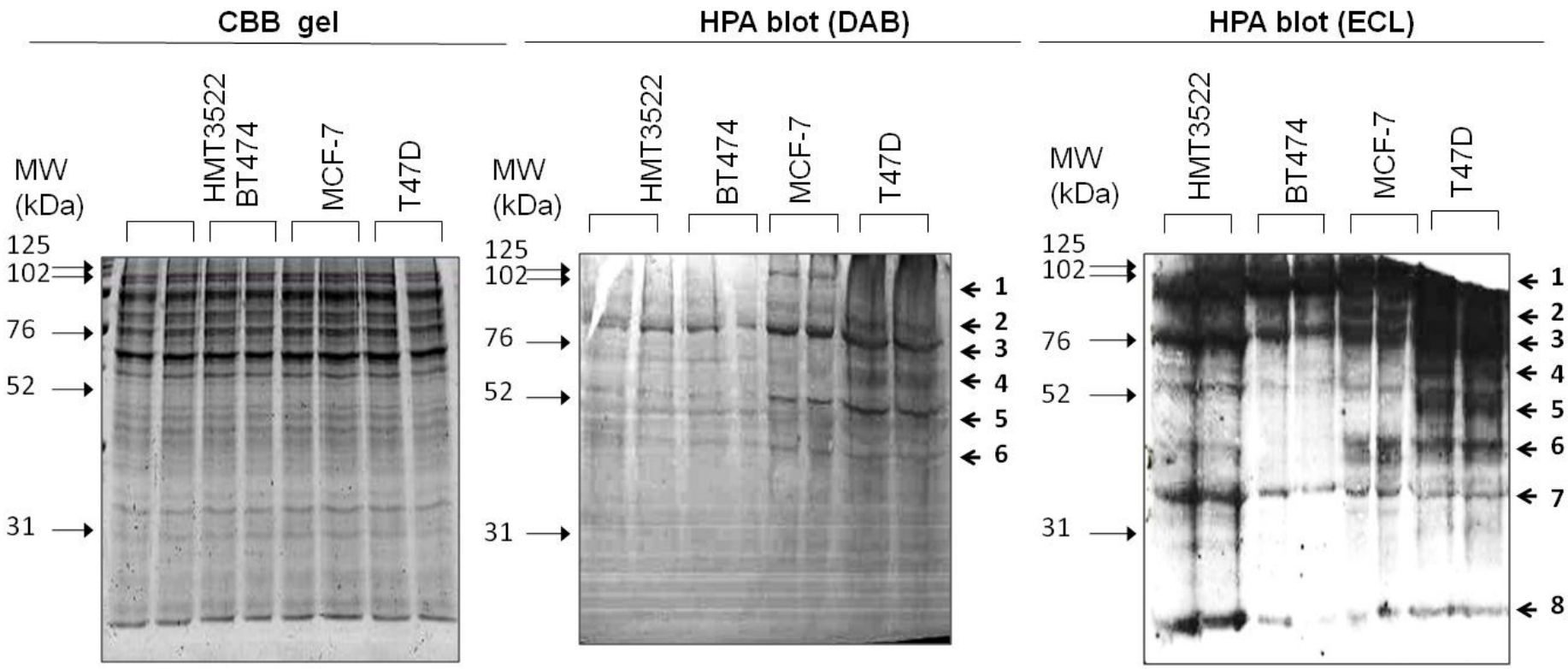


Figure 3.4: Representative CBB stained gel and lectin blots of cytoplasmic proteins. MW markers (kDa) are indicated on the left of the lanes. 20µg of proteins were loaded in each lane, separated by SDS-PAGE, replicate gels were transfered on nitrocellulose and blotted with HPA lectin. Detection was done with the diaminobenzinidine (DAB) and enhanced chemiluminescence (ECL). The CBB gel shows homogenous distribution of protein levels of the cytoplasmic proteins. The DAB blot shows a higher number of proteins bands detected in MCF-7 (5 bands detected) and T47D (6 bands) compared to HMT3522 and BT474 where only 4 weak bands were detected. The ECL blot shows 8 and 7 major glycoproteins in T47D and MCF-7, respectively, all of which, except band 1 and 2, were common to HMT3522 and BT474. The ECL method offers an improved detection of low abundance HPA binding glycoproteins. In experiment where the lectin step was omitted and streptavidin-HRP was incubated alone, no bands were detected (data not shown).

Table 3.1: Estimated MW of HPA binding cytoplasmic glycoroteins in the breast cell lines. N/D= not detectable.

Band Number	Estimated molecular Weight (kDa)			
	HMT3522	**BT474**	**MCF-7**	**T47D**
1	102	102	102	102
2	N/D	N/D	83	83
3	81	81	81	81
4	N/D	N/D	N/D	67
5	55	N/D	55	55
6	N/D	N/D	45	45
7	41	41	41	41
8	28	28	28	28

3.3.2 Solubilisation of membrane proteins for 1-DE SDS-PAGE

Prior to separation of membrane proteins for subsequent analysis by SDS-PAGE, the proteins are denatured, reduced, dissagregated and solubilised, the aim of this process was to achieve disruption of molecular interactions and to ensure that each band represents, ideally, an individual polypeptide chain.

In this study, preliminary work was undertaken to solubilise and separate membrane glycoproteins prepared from the breast cell line MCF-7. Different chaotropes such as urea and thiourea are commonly used in combination with detergents such as CHAPS and reducing agents such as DTT to unfold hydrophobic proteins and reduced disulphide bonds (O'Farrell, 1975). To evaluate an optimal buffer condition using these chemicals for the solubilisation of membrane proteins the following three buffers were evaluted, CHAPS buffer (1% w/v CHAPS), thiourea buffer (7 M urea, 4% w/v CHAPS, 1% w/v DTT and 2M thiourea), urea buffer (7M urea, 4% w/v CHAPS and 1% w/v DTT) were used in this experiment. Membrane proteins of MCF-7 were solublised in each buffer and approximately 15 µg of proteins were loaded with an equal volume of Laemilli buffer into the wells of the stacking gels (as described in section 2.5.1).

A number of protein bands were observed with each buffer (figure 3.5), however significant quantitative and qualitative difference in the protein banding patterns were observed with the lowest number of intense protein bands noted with the CHAPS buffer (panel A) , moderate with thiourea buffer (panel B) and the highest with urea buffer (panel C). Protein separation was very poor at the high MW ranging from 50-125 kDa with thiourea buffer, with some vertical streaking also noted across the gels, probably due to the incompatiblity of the thiourea chaotrope with the SDS-PAGE gel constituents. This experiment showed that the urea buffer resulted in the highest number of well resolved protein bands and this buffer was used for further separations of membrane proteins for all the 1-DE SDS-PAGE gel analysis.

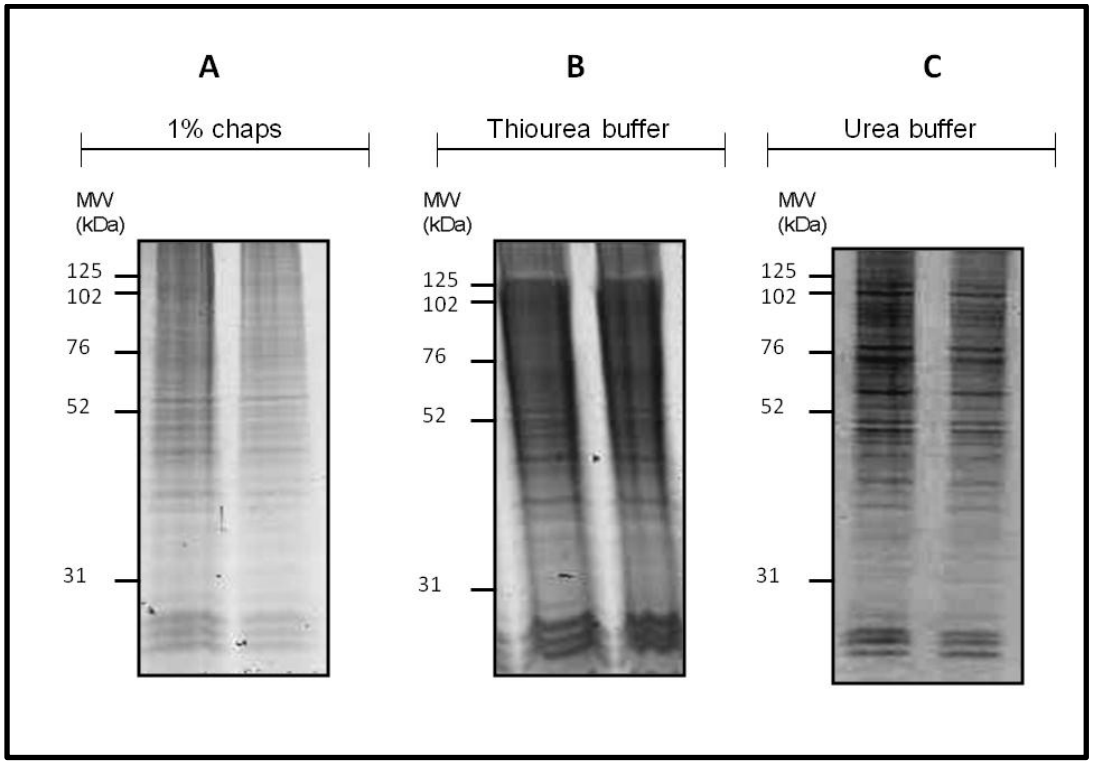


Figure 3.5: Representative CBB stained gels of MCF-7 membrane proteins separated by SDS-PAGE. 15 µg of proteins were loaded in each lane. Panel A: A number of weak intensity proteins bands were noted, Panel B: poor separation of proteins with some vertical streaking occured, Panel C: discrete, high intensity protein bands were noted.

3.3.3 HPA labelling of membrane glycoproteins

To analyse and identify HPA binding glycoproteins of the breast cell lines, the membrane fractions from HMT3522, BT474, MCF-7 and T47D were compared (figure 3.6). Similarities as well as marked qualitative and quantitative differences in the protein band intensities across the cell lines were noted; with T47D exhibiting the lowest number of detectable bands and HMT3522 the highest. Western blotting revealed the highest number of HPA binding glycoproteins in T47D and the lowest in BT474. Strikingly, HPA binding was most intense in the membrane fraction of proteins from T47D which exhibited the lowest protein level in the CBB stained gel (figure 3.6) and it is assumed that these proteins were heavily glycosylated species.

T47D exibited six major HPA binding glycoproteins, ranging in molecular weight from 29 kDa to 130 kDa, five HPA binding bands were observed in MCF-7, all of which were common to T47D. Protein band 1 was identified only in T47D whereas protein band 2 was common to both of the metastatic cell lines (T47D and MCF-7). Protein bands 3, 4, 5 and 6 were common to all the cell lines, althrough the intensity of HPA binding varied across the cell lines with the highest intensity observed in T47D, moderate in MCF-7 and were least intense in HMT3522 and BT474. The two major glycoproteins at approximately 130 kDa (GP130) and 80 (GP80) kDa attracted our interest as they were the most intense HPA binding species in the HPA positive T47D and MCF-7 cell lines and virtually non-detectable in HMT3522 and BT474. Moreover, recently, Saint-Guirons *et al* (2007) reported two major glycoproteins of similar molecular weight in the HT29 CRC cell line. The estimated molecular weight of the of the HPA binding membrane glycoproteins for main bands 1-6, for each cell line is shown in table 3.2.

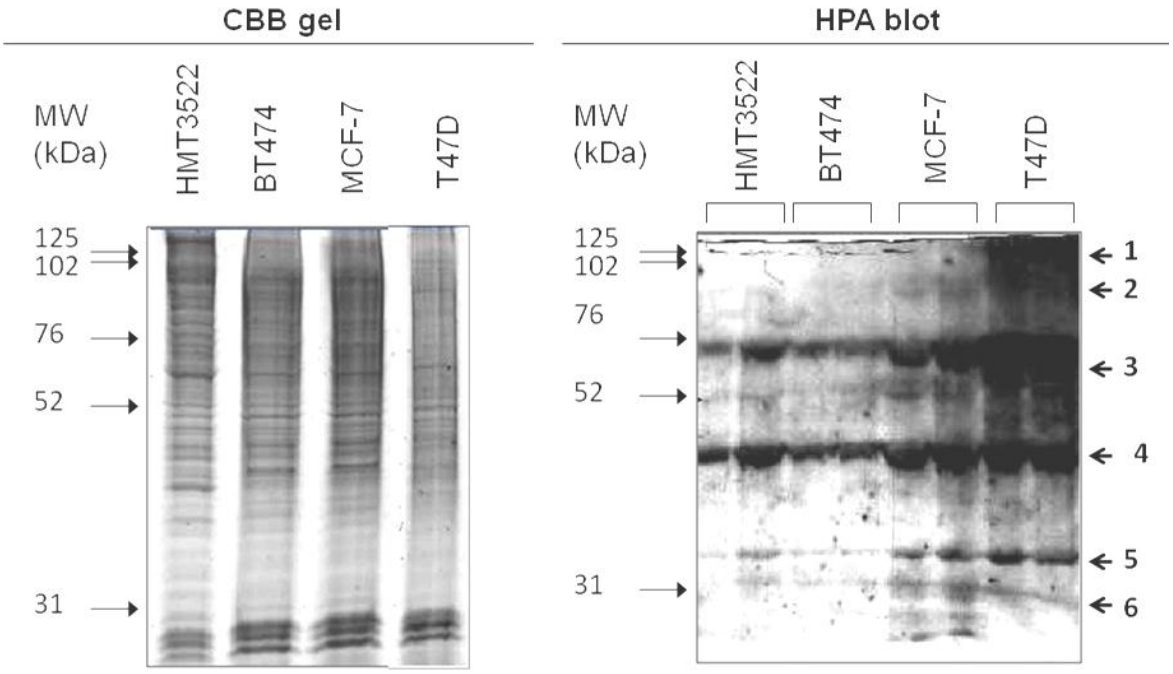


Figure 3.6: Representative CBB stained gel and lectin blot of membrane proteins. MW markers (kDa) are indicated on the left of the lanes. Panel A: 20µg of proteins were separated in each lane. Panel B: A replicate gel was transferred on nitrocellulose and blotted with HPA lectin, detection was performed with ECL reagent. Four bands (bands 3, 4, 5 and 6) were detected in all the four cell lines. Band 1 was found only in T47D while band 2 was detected in both metastatic cell lines (MCF-7 and T47D). In the negative control where the lectin step was omitted and streptavidin was incubated alone, no bands were detected (data not shown).

Table 3.2: Estimated MW of HPA binding membrane glycoroteins in the breast cell lines. N/D= not detectable.

Band Number	**Estimated molecular Weight (kDa)**			
	HMT3522	BT474	MCF-7	T47D
1	N/D	N/D	N/D	130
2	N/D	N/D	80	80
3	75	75	75	75
4	50	50	50	50
5	35	35	35	35
6	29	29	29	29

3.4 Evaluation of the specificity of HPA binding

HPA is known to have a nominal affinity for glycans bearing the GalNAc and GlcNAc residues (Hammarstrom & Kabat, 1971; Lescar *et al*., 2007; Vretblad *et al*., 1979; Wu & Sugii, 1991). In this study, the specificity of HPA binding to T47D membrane proteins was evaluated by competitive inhibition using GalNAc, GlcNAc and Man monosaccharides (as described in section 2.3.4). A scoring system was used to assess inhibition of HPA binding.

3.4.1 Evaluation of the specificity of HPA binding at cellular level

Pre-incubation of the lectin with different concentrations of GalNAc (25 mM to 100 mM revealed specific inhibition of HPA-binding to T47D cells, with a decrease in HPA-TRITC binding in proportion to the amount of competing sugar (figure 3.7). HPA staining was almost completely abolished when the lectin was preincubated with 50 mM GalNAc. When the cells were incubated with HPA lectin preincubated with 100 mM GlcNAc and 100 mM Man, a marked difference in HPA labelling of T47D cells was observed. Moderate labelling of T47D was noted when the lectin was preincubated with mannose, however, no significant binding occured when the lectin was preincubated with GlcNAc. A scale ranging from '-' for low HPA binding and '++++' for strong binding was used to score the HPA binding. The scoring for HPA labelling was as follows: HPA only (++++), 25 mM GalNAc (+), 50 mM GalNAc (-), 100 mM GalNAc (-), 100 mM GlcNAc (-) and 100 mM Man (+).

These observations suggest that HPA recognises the cancer cells by virtue of recognising GalNAc or GlcNAc containing epitopes. This finding supports the observations made by several other groups (Hammarstorm & Kabat, 1969; Hammarstorm & Kabat., 1971; Lescar *et al*., 2007; Sanchez *et al*., 2006).

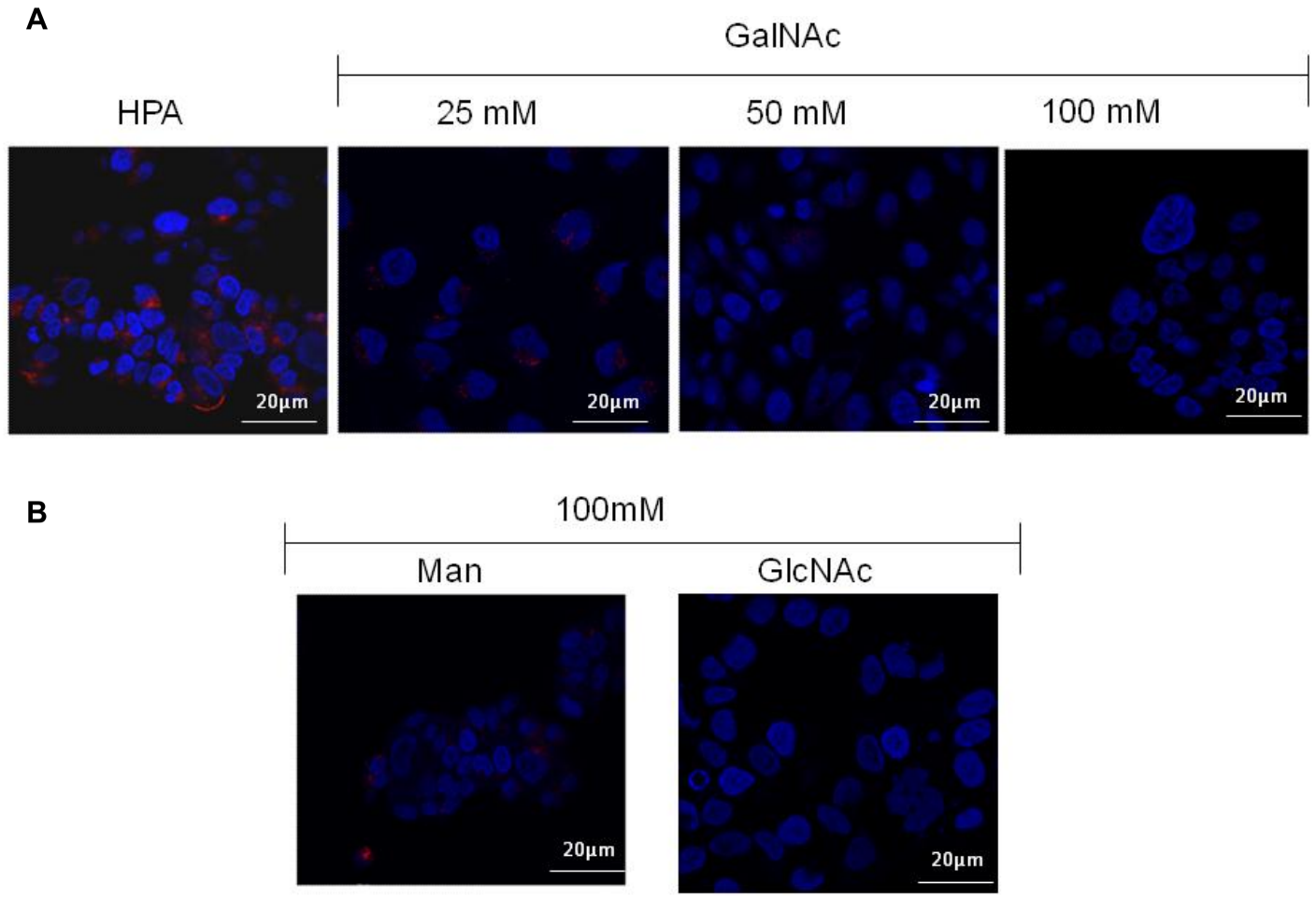


Figure 3.7: Specificity of HPA binding to T47D cells. Panel A: the cells were incubated with 10 µg/ml HPA-TRITC alone or with HPA-TRITC preincubated with increasing concentrations of GalNAc (25mM-100mM). The cells were counter-stained with To-Pro-3. Inhibition of HPA binding was observed with increasing concentration of GalNAc. Staining was almost completely abolished when the lectin was preincubated with 50 mM GalNAc. Panel B: Moderate inhibition was noted when the lectin was preincubated with 100 mM Man whilst labelling was completely abolished when HPA was preincubated with 100 mM GlcNAc. Scale bars = 20µm.

3.4.2 Evaluation of the specificity of HPA binding to cell membrane glycoproteins

To further evaluate the specificity of HPA binding, membrane proteins from T47D were separated by 1-DE and blotted on to nitrocellulose and probed with HPA again, following pre-incubation with different buffers containing 100 mM Gal, 100 Man, GalNAc and GlcNAc (as described in section 2.13.2). Weak inhibition of the HPA binding to the proteins was observed when the lectin was pre-incubated with Gal and Man with Gal being a slightly more effective inhibitor of HPA binding to lower molecular weight proteins (approximately 49kDa) than Man (figure 3.8). On the other hand, almost complete abrogation of HPA binding to T47D proteins were noted when HPA was preincubated with GalNAc and GlcNAc.

These observations indicated that GalNAc and GlcNac were both effective inhibitors of HPA binding to the higher molecular weight proteins, (for example, GP130 and GP80), supporting the view that these proteins may contain GalNAc or GlcNAc bearing epitopes (Markiv *et al*., 2011).

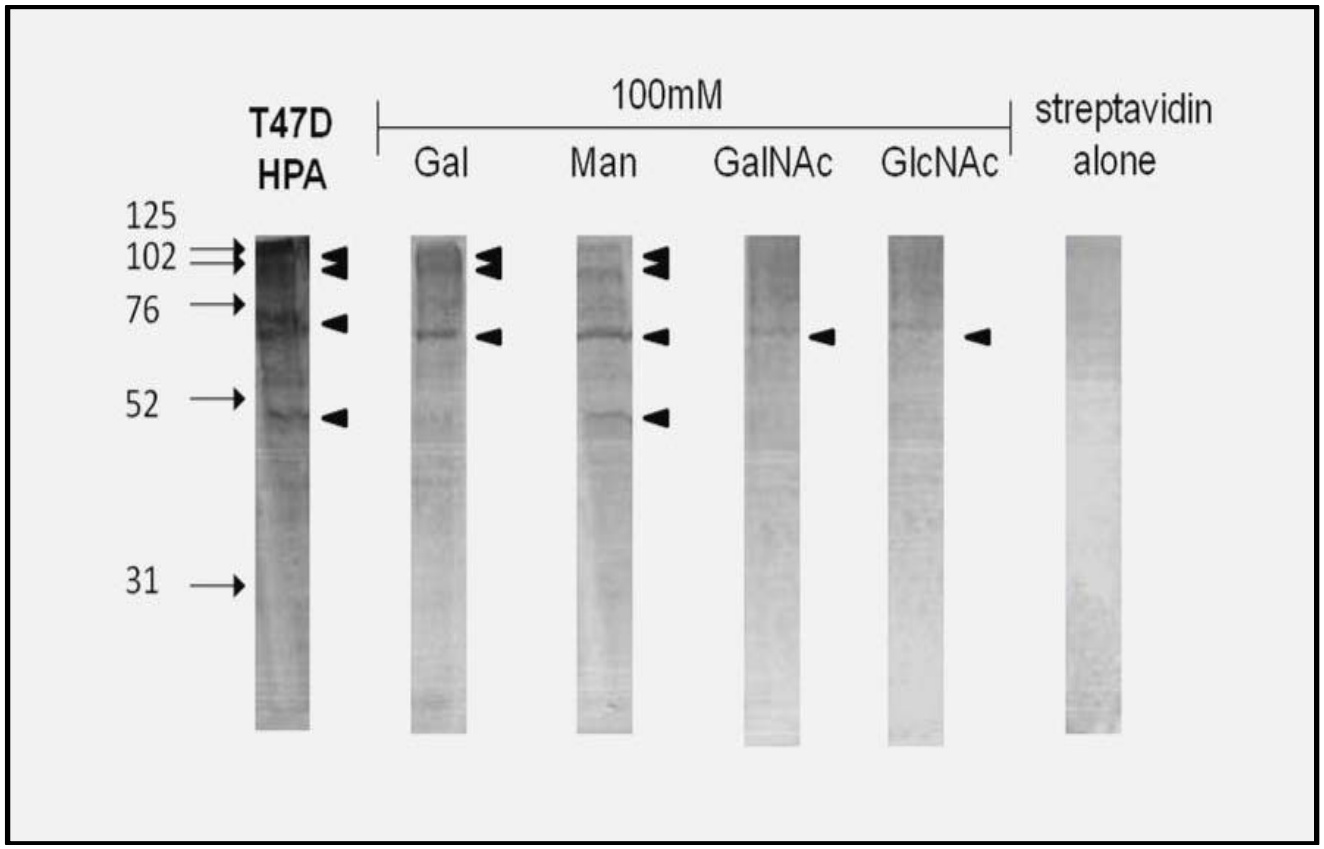


Figure 3.8: Inhibition of the HPA binding to T47D membrane proteins. T47D proteins were incubated with HPA-biotin, preincubated for 30 min with 100 mM Gal, Man, GalNAc or GlcNAc followed by streptavidin-HRP. Detection was performed using the DAB method Inhibition of HPA binding to GP130 and GP80 was more prominent when HPA had been preincubated with GalNAc and GlcNAc. The negative control where HPA was omitted showed a non-specific band of 75kDa present in every experiment.

3.5 Discussion

3.5.1 Microscopy

The initial findings revealed intense HPA labelling in T47D and MCF-7. Whilst light microscopy confirmed that T47D and MCF-7 were HPA positive and HTM3522 and BT474 were HPA negative, confocal microscopy greatly facilitated detailed localisation of HPA binding in these cells. T47D exhibited intense cell surface as well as cytoplasmic labelling whereas MCF-7 showed punctuate HPA binding in the perinuclear region. Both metastatic cell lines bound HPA in an intense manner in the perinuclear region, this may represent binding of HPA to glycoproteins in transit through the secretory pathway or it may be that HPA binds glycosyltransferse/glycosidase enzymes resident in the Golgi apparatus (Brooks *et al*., 2001; Laitinen *et al*., 1990; Mitchell *et al*., 1995; Roth, 1984). In 2001, Brooks *et al*reported weak or negligible HPA staining in BT474 and HMT3522 and intense in MCF-7, but with moderate cell surface labelling in BT474 and MCF-7. This study, however, failed to report cell surface labelling in both these cell lines. These inconsistencies in observations may be attributed to a number of factors including cell passage, cell growth parameters or the selection of a subclone of cells which do not express HPA binding ligands. Nevertheless, these observations correlates with the HPA binding patterns previously reported in colorectal cell lines studies with metastatic HT29 cells exhibiting intense cell surface/granular/perinuclear labelling and non-metastatic SW480 cells exhibiting weak or negligible staining (Saint-Guirons *et al*., 2007).

The reconstruction of 3D images of T47D (from the evaluation of T47D confocal microscopy), confirmed the observation made with the classical 2D models that HPA binds strongly to the metastatic cell line T47D on the cell surface as well as in the perinuclear region. The images from confocal microscopy offered a deeper insight into the distribution of HPA binding epitopes within a subpopulation of the cells and heterogenous expression of HPA binding epitopes observed, across the cell population reflects the heterogeneous nature of the cancer cell population and the pleomorphic nature of the nuclei (figure 3.3).

3.5.2 SDS PAGE and Western blotting

Comparative analysis of CBB stained SDS-PAGE gels and Western blots probed with HPA showed that there are significant differences in HPA binding protein species across the four breast cell lines. The T47D and MCF-7 cells exhibited the highest number of HPA binding proteins whilst HMT3522 and BT474 had the lowest. This result was entirely consistent with the phenotype of these cell lines and correlated with the results of both light and confocal microscopy experiments.

Lectin blotting of a fraction containing cytoplasmic proteins (from the breast cell lines) revealed intense HPA binding in MCF-7 and T47D. The HPA binding glycoproteins varied in relative molecular weight and were between 28-102 kDa. In the case of T47D, six main glycoprotein bands were detected with the DAB method and eight with the ECL detection method. However ten times lower concentration of lectin was used in the ECL system (0.5 µg/ml) than the DAB method (5 µg/ml), thus the ECL method of detection offered a much greater level of sensitivity than the DAB detection system. In the case of MCF-7 seven bands were identified with the ECL system, however, Brooks *et al* (2001) reported eleven glycoproteins of 20 kDa to 200 kDa using whole cell lysates from this cell line and using this method. This discrepancy between observations may be due to a loss of HPA binding glycoproteins during the differential centrifugation step that was adopted for the separation of cytoplasmic and membrane proteins. It is worth noting that Mitchell *et al*(1995) reported seven HPA-binding bands of 20–90 kDa in breast cell lines and tentatively identified a protein species at 90 kDa, as the transferrin receptor, the data correlating with the observations that have been reported here.

To enable the analysis of cell membrane glycoproteins, a range of buffers to solubilise the proteins were evaluated. The buffer containing urea resulted in the best separation of membrane proteins as assessed by SDS-PAGE. Lectin blots of membrane proteins revealed five and six major glycoproteins of 29-130 kDa which interact with HPA. Many of the proteins that were less than 80 kDa were

detectable in all four cell lines, although the bands were more intense for MCF-7 and T47D. Two major HPA binding partners GP80 (MW: 80 KDa) and GP130 (MW: 130 kDa) were identified in the membrane protein preparations of T47D cells. GP130 was specific to T47D whilst GP80 was common to both MCF-7 and T47D. These observations were consistent with the findings of Saint Guirons *et al*(2007) who previously reported two major glycoproteins of similar relative molecular weight in HT29 a CRC cell line, and identifed those proteins to be Hsp 90 (GP80) and integrin α6/αv (GP130).

3.5.3 Inhibition of lectin binding

Preincubation of lectins with monosaccharides is a technique that has widely been used to define nominal lectin-binding properties (Macedo *et al*., 2007; Sumida *et al*., 1997). Whilst it is rather a simplistic approach it does offer the potential to understand which interactions are glycan mediated and which are likely to be 'non specific', for example protein-protein interactions.

Inhibition studies showed that HPA binding could be inhibited using GalNAc and GlcNAc, confirming previous findings that HPA has affinity for these epitopes (Hammarstorm & Kabat, 1969; Hammarstorm & Kabat, 1971; Lescar *et al*., 2007; Sanchez *et al*., 2006).

3.6 Conclusion

Taking together the results obtained from confocal microscopy and lectin blotting, the present study confirmed that the cell lines used show differences in glycosylation as detected using HPA. Disruption in glycosylation pathways in metastatic cancer cells is mostly reflected by the presence of abrogated glycans structures on cell surface proteins. It is interesting to speculate as to whether these glycosylation changes contribute to the metastatic behaviour of the breast cells *in vivo*.

Chapter 4

Identification of HPA binding glycoproteins in breast cancer cells

4.0 Identification of HPA binding glycoproteins in breast cancer cells

4.1 Introduction

In the previous chapter, the experiments with the breast cancer cell lines confirmed the HPA binding properties of the cells (Brooks *et al.,* 2001; Schumacher & Adam, 1997). In addition the HPA binding ligands were localised mainly to the cell surface in the cell line T47D. In order to identify the HPA binding glycoproteins in the cell lines, a method based on the 2-DE separation of proteins was used. In a large gel format this method has the advantage of allowing the resolution of 1,000-1,500 proteins (based on their relative MW and pI) in a single experiment (O'Farrell, 1975). The mini-gel format has been shown to enable separation of between 300-500 proteins in a single experiment (Weiss & Gorg, 2008). Saint-Guirons *et al* (2007) previously identified several HPA binding glycoproteins from CRC using the 2-DE method in HT29 colorectal cells. However, the authors used a pre-fractionationation step, based on the separation of the proteins using HPA affinity chromatography. In the system adopted by Saint-Guirons *et al* (2007) HPA was observed to recognise proteins involved in cell adhesion and migration (integrin α6/αV, annexin A2/A4), remodelling (α and β tubulin, actin, cytokeratins) and anti-apoptic pathways (Hsp70, Hsp90, TRAP-1, and TNFR 1).

To extend this analysis to breast cancer, membrane proteins from the breast cancer cell lines were separated by 2-DE in a proteomic approach that was used in combination with lectin blotting. The aim of the work was to establish whether the same glycoproteins are recognised by HPA in breast cells as in CRC cells. First, a solubilisation buffer maximising the yield and separation of membrane glycoproteins was identified. Secondly, the robustness of the 2-DE system was assessed by evaluating gel reproducibility. The proteome profile across the four breast cell lines was compared and proteins in increased levels in the metastatic T47D cell line were identified using MALDI-MS and MS-MS, finally, the HPA binding glycoproteins of T47D were identified again by MALDI-MS and MS-MS.

4.2 Solubilisation of membrane proteins for separation by 2-DE

The solubilisation of membrane proteins remains a critical step for high-performance 2-DE. Transmembrane proteins comprise hydrophobic groups of amino acids and the use of detergents which form micellar structures, such as ionic detergents (for example, SDS), can be used to achieve complete solubilisation of these proteins. However, ionic detergents are not compatible with the IEF step, therefore only non-ionic detergents, including urea, thiourea and CHAPS, appropriate for the solubilisation of membrane proteins, are then used in a 2-DE based work-flow (Rabilloud, 1998).

In this section, a number of buffers were evaluated in terms of their utility for solubilisation and separation of membrane proteins from MCF-7 breast cells. As a minimum requirement, a solubilisation buffer should contain; chaotropes such as urea and/or thiourea, zwitterionic detergents such as CHAPS to unfold proteins; reducing agents such as DTT to break disulphide bonds within proteins and carrier ampholytes to facilitate solubilisation with the aim that each spot contains a single peptide (Molloy, 2000; O'Farrell, 1975; Santoni *et al.,* 2000). In an attempt to obtain the maximum protein yield and the clearest protein separation, four buffers was assessed, these were; U buffer (7 M urea, 4% w/v CHAPS, 1% w/v DTT, 2% v/v ampholytes and 2 M thiourea), U1 buffer (as U buffer but with 8 M urea), T1 (as U1 buffer but with 3 M thiourea) and U1T1 (as U1 buffer but with 8 M urea and 3 M thiourea). MCF-7 membrane protein samples were prepared in each of these buffers and separated by 2-DE (as described in section 2.9). The gels were stained with Coomassie brilliant blue.

The MCF-7 membrane proteins solubilised in U buffer, U1 buffer, T1 buffer and U1T1 buffer showed similar protein migration patterns, although there was sharp variation in the quality of separation due to the solubility achieved with each buffer (figure 4.1). The resolution of separated proteins was poor for the urea containing buffers (U and U1) whereas when thiourea was added (T1 and U1TI), a better separation was achieved. Good resolution of both high and low abundance

proteins was noted with the U1TI buffer. A region of relative molecular weight of 55-130 kDa and pI of 4.5-6.5 (framed in red), which was subject to variation in resolution depending on the buffer used to solubilise the protein, was selected and is shown in a zoom view in figure 4.2. The framed region contains a large selection of proteins and was compared using the 2-DE analysis Progenesis Same Spots software.

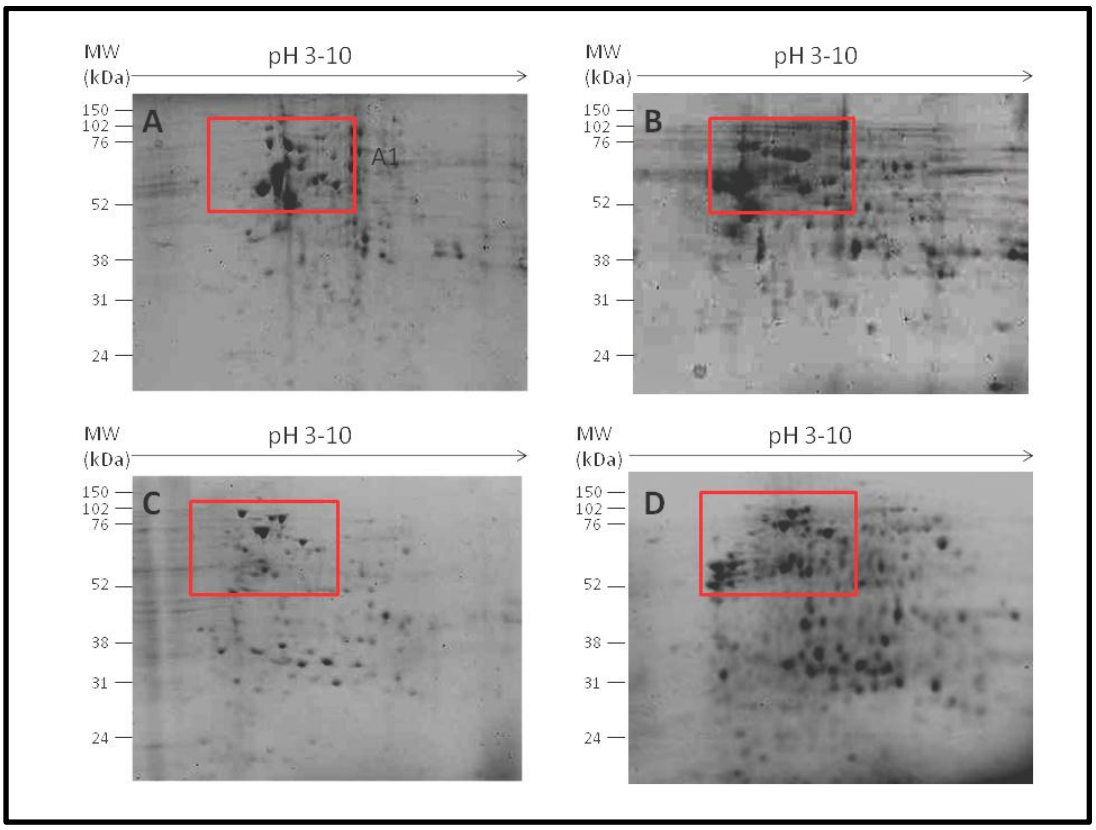


Figure 4.1: MCF-7 membrane proteins solubilised in four buffers and separated by 2-DE. MCF-7 membrane proteins were prepared and solubilised in U buffer (panel A), U1 (panel B), T1 buffer (panel C) and U1T1 buffer (panel D). In each case, 70 µg of proteins were loaded on a pH 3-10 IPG strip and subsequently separated on a 12% SDS-PAGE in order to achieve a bidimensional separation of the proteins. The proteins were visualised by staining with CBB. Similarities in the general protein spot pattern are observed in the four gels.

Poor resolution was observed in buffers which did not contain thiourea (panel A and B). However, an improved protein separation was observed with (T1 and U1T1) (figure 4.2 panel C and D) suggesting that thiourea is an important chaotrope for unfolding membrane proteins. In these two latter systems, the proteins were all well separated, although there was a marked variation in the quality of separation observed due to the solubilisation achieved with each buffer.

Two specific regions were further investigated with proteins solubilised in either T1 or U1T1 buffer (figure 4.2, A1 and A2), as they contained proteins that seemed to vary across the two gels. Zone A1, panel C contained five major well- separated protein spots. This detailed view showed that the T1 buffer offered good separation of proteins including those with similar isoelectric points. On the other hand, proteins solubilised in the U1T1 (panel D) buffer showed better resolution of proteins in terms of the number and intensity of proteins observed with 10 intense protein spots observed (A1 boxed area). The U1T1 buffer contains the same components as T1 buffer with the difference that urea was added in higher concentration and the results obtained were consistent with the hypothesis that thiourea and urea allows improved solubilisation of hydrophobic proteins (Rabilloud, 1998). The T1 buffer failed to resolve proteins, shown boxed in area A2, suggesting that the concentration of thiourea may be critical component for the solubility of these proteins. This area contained six major proteins observed with U1T1 buffer.

It can be summarised from this experiment that the addition of 8M urea and 3M thiourea in combination gave a better resolution than solubilisation in urea containing buffers (without the addition of thiourea) and U1T1 was the most suitable buffer of those tested for the solubilisation of the membrane proteins. The U1T1 buffer was therefore used in subsequent experiments.

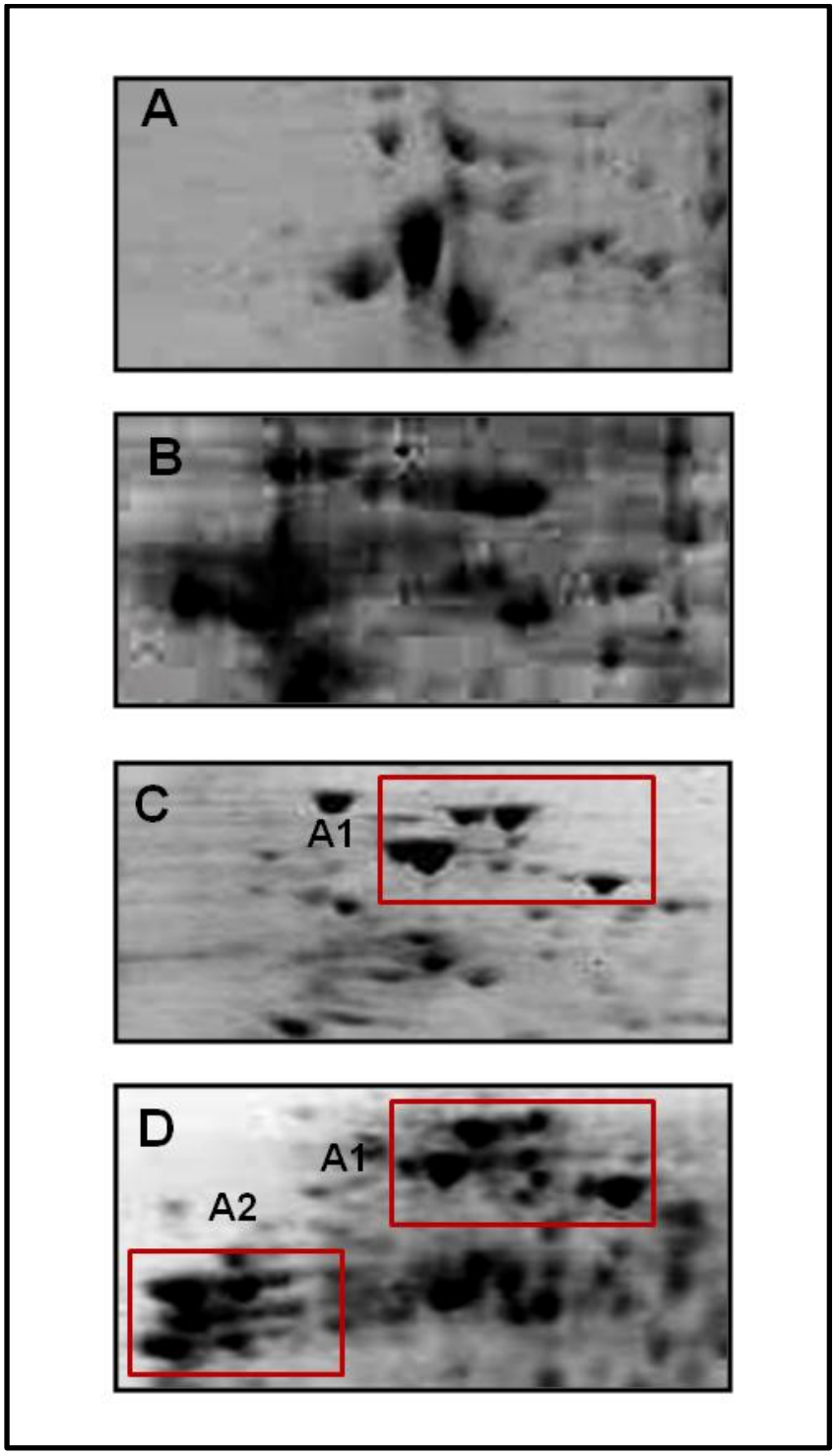


Figure 4.2: Evaluation of four solubilisation buffers. A zoomed view of the framed regions selected on gel A, B,C and D in figure 4.1. Evaluation of U buffer (panel A), U1 buffer (panel B), T1 buffer (panel C) and U1T1 (panel D) for MCF-7 membrane protein separation. Membrane proteins solubilised in U buffer and T1 buffer were poorly separated. Distinct protein species were observed with T1 and U1T1 buffer. Zone A1 shows 5 and 10 major protein spots in panel C (T1 buffer) and D (U1T1 buffer) respectively. Zone A2 was only observed with U1T1 buffer (panel D).

4.3 Reproducibility of the 2-DE system

The next important consideration was determining a suitable 2-DE system for the study of changes in protein levels in the metastatic and non-metastatic cancer cells. The main factor here was the biological/analytical variation between gels and this aspect of the 2-DE separation method was considered; (1) by assessing the analytical reproducibility of protein migration in two replicates gels of protein extracted from the same passage of T47D cells but collected on two different occasions (same passage of the breast cell lines but different protein preparation), (2) evaluating the biological variation between gels from two different batches of cells with different passage numbers (different passage number and different protein preparation).

4.3.1 Analytical reproducibility of the 2-DE system using T47D cells

Membrane proteins from two different membrane preparations, but from the same cellular passage, were run on two separate gels (in parallel) and the reproducibility was assessed by evaluating the reproducibility between selected regions of the gels and by estimating variation in spots detected in the analytical replicates. In each case 100 µg of protein was loaded on the IPG (pH 3-10) strips. An overview of the separation is shown in figure 4.3.

4.3.1.1 Analytical reproducibility: gel to gel comparison

To evaluate the reproducibility between two gels, three areas were chosen, these contained proteins ranging from low to high abundance and from low to high molecular mass, the three areas were numbered 1, 2 and 3 on the gels (figure 4.3) and are shown in zoom view (figure 4.4). Similar protein migration patterns were observed in the replicate gels (A and B) for both the high and low abundance proteins in both gels. The zoomed view of Area 1 represents proteins of 76-130 kDa and pI 5.5-5.8 and these showed a similar protein separation pattern in both gels (figure 4.4), with some slight differences observed due to gel to gel variation. This region contained seven high abundance proteins (labelled 1-7). Some

differences were observed in terms of intensity, for example protein ‘ 7’ was less intense in replicate B. Area 2 represented proteins with of 50-70 kDa, pI 6.5-6.8 and showed seven (labelled 1-7) high abundance protein species in replicate A and six in replicate B (labelled 1-6). Protein ‘7’ in replicate A was not observed in replicate B. This difference may be due to batch to batch variability in protein solubilisation. Area 3 represented proteins of 35-40 kDa, Pi 5.7-6.6 and revealed six high abundance proteins in both replicate A and B, however protein ‘5’ and ‘6’ were less intense in replicate B. Overall, the protein pattern of migration was reproducible on 2-DE gels across both replicates. Althrough variations in spot intensity were observed, analysis of the three characteristic regions of the two replicate protein samples from different protein preparations showed that a reproducible and robust methodology has been developed, enabling the separation of breast cancer cell membrane proteins by 2-DE.

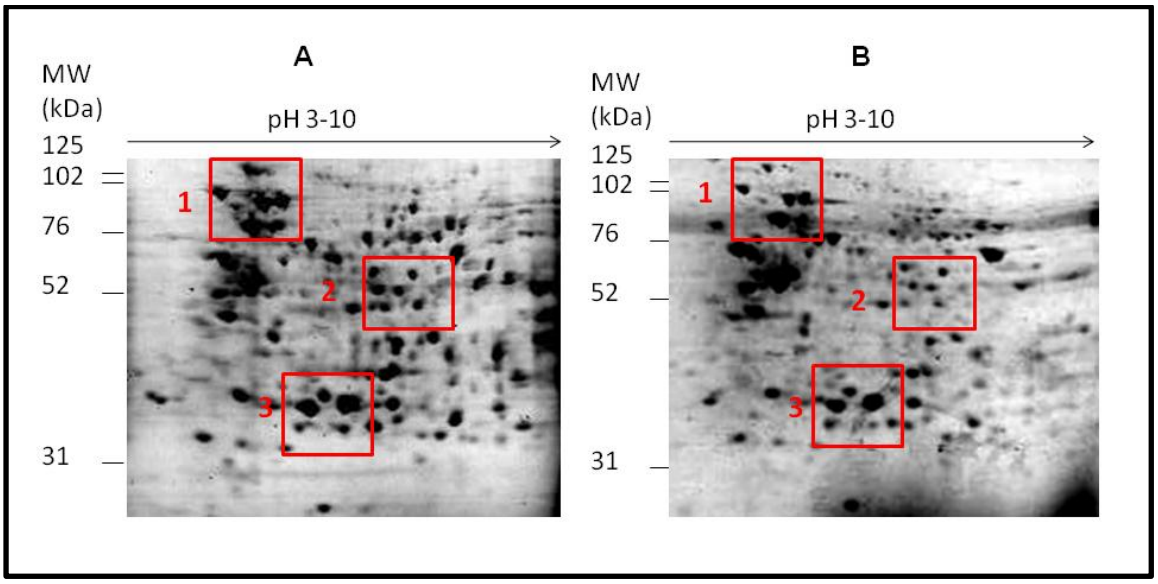


Figure 4.3: T47D membrane proteins separated by 2-DE on pH 3-10 strips. 100 µg of proteins were separated and the gel was stained with Coomassie brilliant blue. Three regions numbered 1,2,3 (red box) were analysed in two replicates (gel A and B) of different protein preparation.

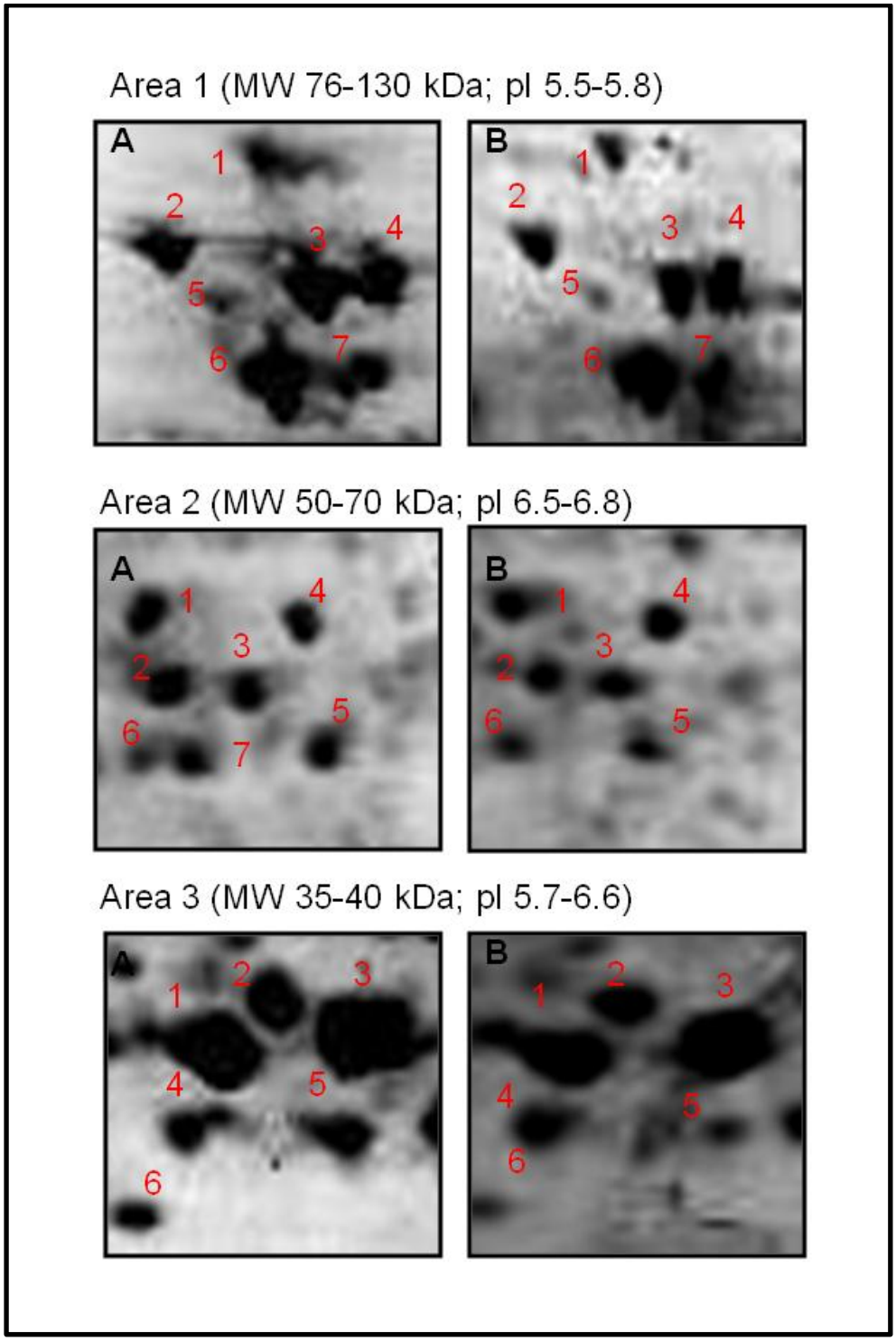


Figure 4.4: Zoom view of area 1,2,3 in replicates A and B of T47D membrane proteins. Area 1 (MW 76-130 kDa; pI 5.5-5.8) shows 7 major protein spots labelled 1-7. High abundance proteins were reproducible in both replicates A and B. Protein '5' seemed to sustain a significant variation in quantity, with higher intensity of protein spots noted in replicate gel A. Area 2 (MW 50-70 kDa; pI 6.5-6.8) contained 7 proteins in replicate gel A and 6 proteins in gel replicate gel B. Protein '7' was present in replicate gel A but absent in replicate gel B . Area 3 (MW 35-40 kDa; pI 5.7-6.6) show similar migration pattern in replicates A and B, but spot 5 and 6 were more intense in replicate A.

4.3.1.2 Analytical reproducibility: statistical estimate

The analytical reproducibility of the 2-DE system was also assessed by estimating the coefficient of variation between the analytical replicates. This was carried out by comparing the average number of protein spots from two different batches of protein preparation but from the same cell passage for each of the breast cell lines. The representative 2-DE separated protein gels are shown in figure 4.6. The spot detection software revealed a similar number of spots in the 6 replicates gels from the two different batches of proteins prepared from the breast cell lines (table 4.1). HMT3522 and T47D (both with CV=0.01) were the most reproducible whilst BT474 (CV=0.02) and MCF-7 (CV=0.03) were least reproducible. The highest number of protein species separated was noted in T47D and the lowest in HMT3522 (figure 4.5). This observation was consistent with the metastatic phenotype of the cell lines and is suggestive that an increase in the number of proteins accompanies cells with a metastatic phenotype. In summary, these data validate the efficacy of the solubilisation buffers and showed that the capacity of the 2DE system to enable reproducible separation of high as well as low abundance protein isoforms with high resolution in all the four cell lines.

Table 4.1: Total number of protein spots from same batch of cells. The total number of protein spots observed in 6 replicate gels from two different protein preparations (1 and 2), from the same passage of breast cell lines HMT3522, BT474, MCF-7 or T47D. The mean average proteins species detected from each cell line, standard deviation and coefficient of variation (CV) is shown. Highest variability in spot numbers were observed for MCF-7 (CV=0.03) and lowest in HMT3522 (CV=0.01)

	Total number of protein spots								
	preparation 1			preparation 2					
Cell line	**Gel 1**	**Gel 2**	**Gel 3**	**Gel 4**	**Gel 5**	**Gel 6**	**Mean (x̄)**	**±SD**	**CV**
HMT3522	425	432	428	437	424	430	429	4.4	0.01
BT474	501	525	534	507	511	525	517	11.6	0.02
MCF-7	575	580	586	587	550	555	572	14.5	0.03
T47D	580	598	587	588	589	590	589	5.3	0.01

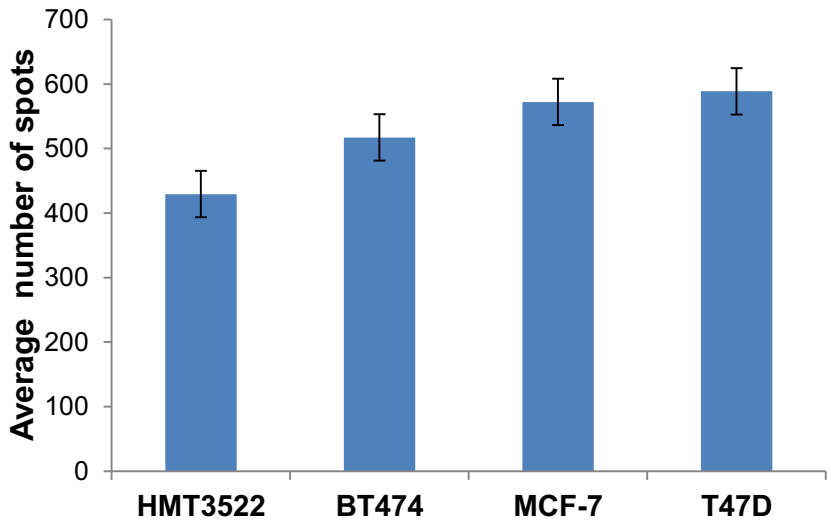


Figure 4.5: The average number of proteins in the breast cell lines. The average number of protein species detected following 2-DE separation of cell membrane proteins in U1T1 buffer. 100 µg of protein was loaded onto mini-SDS-PAGE gel and stained with CBB. Mean +/- SD shown.

4.3.2 Biological reproducibility of the 2-DE system in the breast cell lines

In this experiment we aimed to assess the biological reproducibility of the 2-DE system using protein preparations from two different cell passage numbers (batch A and batch B) from MCF-7 and T47D (table 4.2). The variation in spots resolved was higher when proteins were extracted from different cellular passage (CV≥0.05) than when proteins were extracted from same passage (CV≤0.03). Parker *et al*(2006) have reported a biological difference where the coefficient of variation was within a range of 0.7 again using a 2-DE approach. In this study it was considered that the difference in average number of protein spots from two different passages of cells was fairly low, therefore the system was considered to be sufficiently robust to allow the comparison of protein levels across the cell lines.

Table 4.2: Total number of protein spots from different batches of cells. The average number of protein spots The average number of protein spots observed from two different cellular passages (batch A and batch B) of the breast cell lines MCF-7 and T47D. The average numbers of proteins in each cell line are shown with standard deviation (SD) and coefficient of variation (CV). A higher variability of spot number was observed in T47D (CV=0.06).
.

	Average number of protein spots			
Cell line	**Batch A**	**Batch B**	**(±)SD**	**CV**
T47D	589	651	35.5	0.06
MCF-7	572	578	26.2	0.05

4.4 Membrane protein profile of the breast cell lines

To compare the protein levels across the cell lines, it is necessary to identify 'landmark' or 'reference' proteins which are present across all the cells and use the location of these to 'anchor' the profiles and enable them to be overlaid. HMT3522, BT474, MCF-7 and T47D, all originate from breast tissue and would, therefore, be expected to contain similar structural proteins, these might be used as reference proteins. A total of ten reference/landmark protein species which were present across all the four breast cell lines were used for the overlay analysis of the protein separations. The protein profiles across the four breast cell lines were compared and this enabled the characterisation of the proteins found in elevated levels in the breast cancer cell lines with metastatic phenotype (T47D and MCF-7).

An overview of protein separations from the four breast cell lines is shown in figure 4.6. The general pattern of migration of proteins from the four cell lines was similar, althrough variations in intensity of protein spots were noted. The region framed in red represent high abundance proteins (50-125 kDa) and these were noted in the all the cell lines, however estimation by eye showed that more high abundance proteins was present in T47D, consistent with a increased number of proteins separated from T47D. The overlay analysis revealed proteins which were elevated in T47D and MCF-7 compared to HMT3522 and BT474 ($P \leq 0.005$; fold difference ≥ 1.5). These proteins were identified by MALDI-MS and MS-MS. The proteins in elevated levels are shown in figure 4.7 and are numbered 1-18. The average normalised volume, fold difference and P values of each protein are shown in Appendix 2. The identity of the proteins are shown in table 4.3, the protein sequence and sequence coverage of proteins with the peptides are shown in Appendix 3.

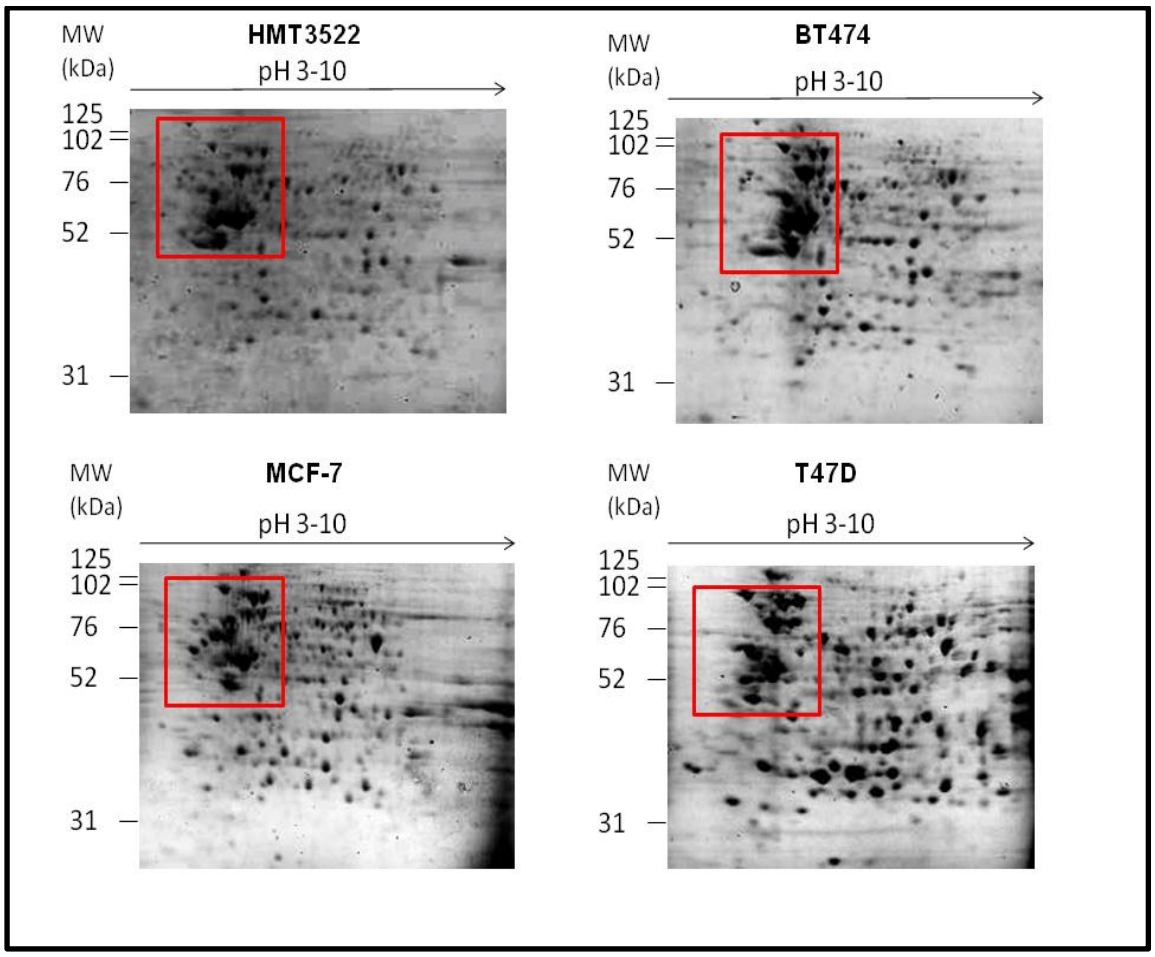


Figure 4.6: Representative 2-DE gels of membrane proteins of HMT3522, BT474, MCF-7 and T47D. 100 µg of proteins were separated on a pH 3-10 IPG strip and then by SDS-PAGE.The cells were stained with CBB. Similar protein migration profile was observed across all the four cell lines. The region framed in red represent high abundance proteins (50-125 kDa).

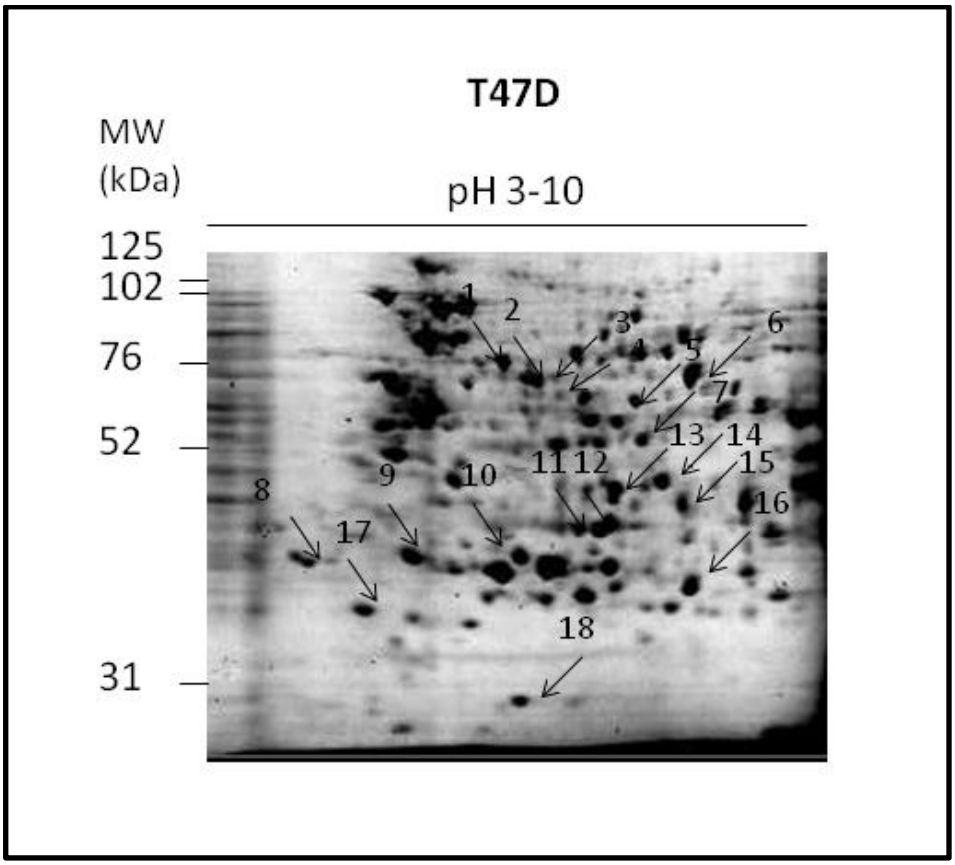


Figure 4.7: Identification of proteins in elevated levels in T47D cell lines. Comparison of gels of membrane proteins separated 2-DE and stained with CBB showed that proteins numbered 1-18 were elevated in metastatic T47D cells. The proteins identities are shown in table 4.3.

4.5 Characterisation of HPA binding glycoproteins: 2-DE and lectin blotting analysis

In chapter 3, a strategy for the separation and analysis of HPA binding membrane proteins from the breast cell lines was described. HPA bound most intensely to the membrane proteins of T47D. To further characterise the HPA binding partners of the cell lines, membrane proteins were separated by 2-DE, replicate gels were prepared, transferred to nitrocellulose and probed with HPA, in this system the resolving power of 2-DE was employed to investigate in detail the individual polypeptides recognised by HPA. All the four cell lines were subjected to this analysis (figure 4.8).

HPA binding proteins were most abundant in the T47D cells. High molecular weight protein species were observed at approximately 80 kDa and 130 kDa (figure 4.8, marked with red arrows), and several proteins species of ≤ 80kDa were also observed to interact with HPA. These observations were consistent with those of the 1-DE analysis where GP80 and GP130 comprised the most intense HPA binding components of T47D (section 3.3.3). The protein species that showed the highest abundance in terms of HPA binding were identified on the equivalent CBB stained gels and were subjected to MALDI-MS and MS-MS (table 4.3). GP80 and GP130 as well as a number of other low abundance proteins were not detected by MS. Other HPA binding proteins of ≤80 kDa were identified as: heterogeneous nuclear ribonuclear protein H1 (hnRNP H1), heterogeneous nuclear ribonuclearprotein D-like (HnRNP-D like), heterogeneous nuclear ribonuclear protein A2/B1: isoform 1 (hnRNP A2/B1), enolase 1 (ENO1), heat shock protein 27 (hsp 27) and glial fibrillary acidic protein isoform 1 (GFAP). These cytoplasmic proteins were assumed to complex *in vitro* with the membrane proteins. Interestingly, the protein analysis of the previous section (section 4.4) showed that these glycoproteins were also elevated in the metastatic MCF-7 and T47D cells (P ≤0.005 ; fold difference ≥1.5). The most notable changes in protein levels was observed in Hsp27 (p=3.10E-07; fold difference = 6.1). Similar HPA binding glycoproteins were also observed in MCF-7. The proteins at approximately 75 kDa and 29 kDa (circled in red) represented HnRNP H1 and Hsp 27 respectively. The blue arrows show HPA binding glycoproteins at approximately 80 kDa and 130

kDa in MCF-7, however, the number of the protein species observed in this region was less than those observed for T47D. In the 1-DE experiment an HPA binding component at approximately 80 kDa was also observed in MCF-7 and it appears that the GP80 was the principal HPA binding component in MCF-7 whilst GP130 was predominant in the T47D cell line. Unlike the observations made in 1-DE where there were low molecular weight proteins observed in the region of approximately 29-75 kDa in HMT3522 and BT474, the 2-DE analysis showed only two protein species (circled in blue) at approximately 80 kDa which weakly interacted with HPA.

The experiments using 2-DE and HPA with Western blotting confirmed that the higher molecular proteins (GP80 and GP130) that bind HPA in T47D and MCF-7 are not recognised in the non-metastatic cell lines HMT3522 and BT474. Furthermore, the 2-DE system enabled the separation of individual polypeptides hidden within, for example, the GP130 protein band observed in the 1-DE separation, which may represent a number of differentially glycosylated forms of the same polypeptide sequence.

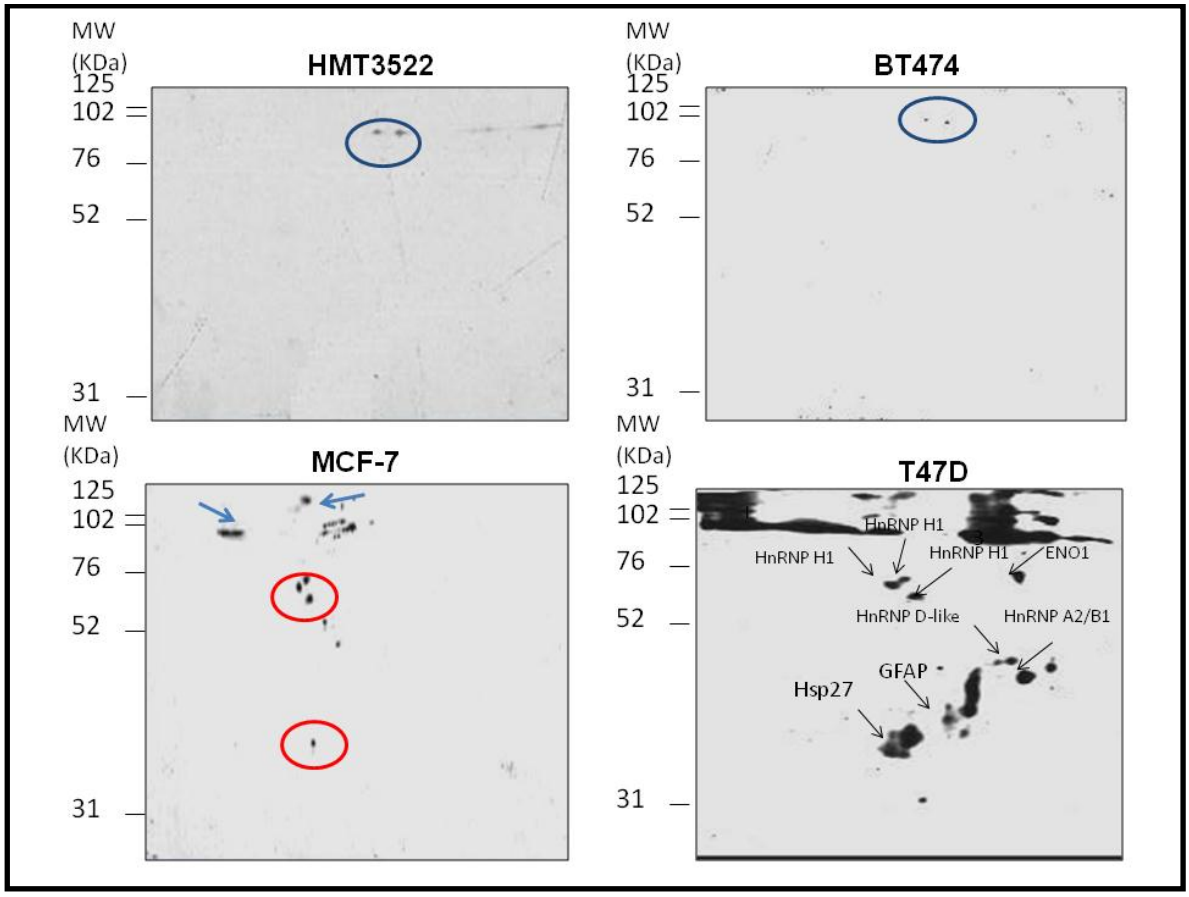


Figure 4.8: HPA binding proteins of T47D, MCF-7, BT474, and HMT3522. 100 µg of proteins were separated by 2-DE, transferred to nitrocellulose by Western blotting and probed with HPA. HPA interacted with 2 protein species in HMT3522 and BT474 at MW of ~80 kDa. A wide range of proteins were recognised below 80 kDa in MCF-7 and T47D. In MCF-7, HPA bound strongly to several protein species at MW ~80 (marked in blue arrows), whereas HPA bound strongly to GP130 and GP80 protein species in T47D (marked in red arrows). HPA did not bind to these proteins in the non-metastatic cell lines. The proteins circled in red in MCF-7 closely corresponded to the MW and position of the characterised protein in T47D. In the negative control, where the lectin step was omitted and blots were probed with streptavidin alone no protein species were detected (data not shown).

Table 4.3: Identification of spots with protein accession number, Mascot score, MW and predicted pI.

Spot No	Protein ID	HPA bound	↑ levels	Accession No	Mascot score	Mr (kDa)	Predicted pI	References
1	glutamate synthetase	-	+	gi|31831	82	42.8	6.61	Liver cancer, Christa *et al* (1994)
2	hnRNP H1	+	-	gi|5031753	73	49.5	5.89	Zhang *et al*.(2008)
3	hnRNP H1	+	-	gi|5031753	124	49.5	5.89	N/A
4	hnRNP H1	+	-	gi|5031753	134	49.5	5.89	N/A
5	Elongation Factor Tu	-	+	gi|704416	139	49.9	7.7	N/A
6	Eno1	+	+	gi|4503571	50	47.5	7.01	Head and neck cancer, Tsai *et al*.(2010)
7	Elongation Factor Tu	-	+	gi|704416	151	49.9	7.7	N/A
8	Macropain subunit delta	-	+	gi|296734	44	19.6	5.01	N/A
9	Proteasome subunit alpha type 2	-	+	gi|4506181	44	25.9	6.9	Colorectal cancer Arlt *et al*. (2009)
10	Heat shock protein 27	+	+	gi|662841	75	22.4	7.83	Breast cancer , Storm *et al*. (1996)
11	Enoyl Coenzyme A hydratase 1 peroxisomal	-	+	gi|16924265	165	36.1	8.5	Stomach and Heptacellular cancer (Yue *et al*., 2008)
12	Glial fibrillary acidic protein isoform 1	+	+	gi| 4503979	66	49.9	5.42	maglignant myoepithelial of the breast , Shirashi *et al* (1999)
13	Prosomal Protein P30-33K	-	+	gi|190447	97	30.4	6.5	Breast cancer, Bhui-Kaur *et al* (1998)
14	hnRNP-D like	+	+	gi|14110407	126	46.5	9.6	N/A
15	hnRNP- A2/B1 (isoform A2)	+	+	gi| 4504497	62	36.0	8.67	Breast and lung cancer, (Zhou *et al*., 2001)
16	Chaperonin containing TCP1 subunit alpha type 2	-	+	gi| 4504497	62	36.0	8.67	Colorectal cancer , (Coghlin *et al*., 2006)

4.6 Anti-CD49f antibody (integrinα6) binding to T47D membrane proteins

The lectin blotting experiments have suggested that GP130 was amongst the most important HPA binding protein of T47D. Saint-Guirons *et al* (2007) previously reported that integrin α6 was the major HPA binding protein component of GP130 in the CRC cell line HT29. To identify whether the protein species at approximately 130 kDa was integrin α6, T47D membrane proteins were separated by SDS-PAGE, blotted to nitrocellulose and probed with a murine anti-human integrin α6 (anti-CD49f) antibody, this was subsequently detected with a peroxidase labelled goat anti-mouse IgG and incubated with the ECL reagent, (as described in section 2.13.3). The predominant 130 kDa protein species observed to bind HPA in T47D was also recognised by the anti-CD49f antibody (figure 4.9). A blot in which the primary antibody step was omitted and IgG-HRP was probed alone confirmed that binding was not due to non-specific interaction of the secondary IgG-HRP with the T47D membrane proteins.

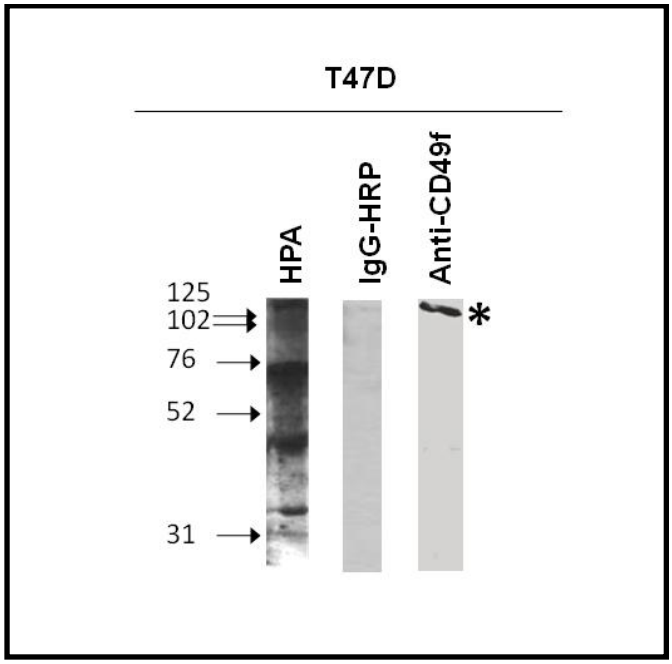


Figure 4.9: Detection of anti-integrin α6 binding to membrane proteins in T47D. T47D membrane proteins were separated and analysed by 1-DE and Western blotting. Detection of anti-CD49f antibody was performed using a secondary IgG-HRP antibody. The equivalent Western blot probed with HPA/streptavidin-HRP is shown as a reference blot. The anti-CD49f antibody detected one major protein species at ~130 kDa in the membrane fraction of T47D (marked by asterisks). No bands were observed when the blots were incubated with IgG-HRP.

4.7 Discussion

4.7.1 Solubilisation of membrane proteins

In this part of the work the aim was to evaluate four solubilisation buffers to enable the separation of cell membrane proteins by 2-DE. Buffers containing urea and thiourea were important for the solubilisation of a fraction enriched in membrane proteins, and these observations are consistent with previous findings (Rabilloud, 1998; Weiss & Gorg, 2008). To date, only a few studies have investigated HPA binding membrane glycoproteins in cancer cells. Redondo *et al* (2004) reported membrane proteins recognised by lectins HPA and WGA in HCT116 and the CaCo2 CRC cell lines, using high speed centrifugation to purify the proteins and detergent based approaches. Recently Saint-Guirons *et al* (2007) used a similar approach to investigate the HPA binding proteins of the HT29 CRC cell line. The methods used to prepare membrane proteins in this study were slightly different with respect to the composition of the buffers but the key step was the use of high speed centrifugation and this was a common feature in these studies. In this work, the solubilisation buffer enabled preparation and separation of membrane proteins from the breast cell lines, which showed similar migration patterns and therefore constitute an ideal model for studying HPA-binding glycoproteins in breast cancer.

4.7.2 Reproducibility of the 2-DE system

The robustness of the 2-DE separation was assessed using two analytical replicates of T47D (different protein preparations but the same cellular passage of the cell line). The protein migration pattern was similar in both replicates, although small variations in terms of intensity of protein spots were noted across the gels due to gel to gel variability. This has been observed in other studies (Valcu & Valcu, 2007). The protein map of the membrane proteins from T47D separated using 2-DE with pH 3-10 IPG strips was comparable to the 2-DE map generated for the HT29 cell line by Saint-Guirons *et al* (2007) and Tan *et al* (2002), although different samples were used in these experiments such as total lysate and membrane preparation. Several high abundance proteins ranging from 40-80 kDa

were noted in these 2-DE protein maps. For example, the high abundance proteins that were analysed in area 2 (50-70 kDa) were also observed in the 2-DE map described by Saint-Guirons *et al* (2007). This may be because these protein preparations contain membrane proteins which were enriched using similar approaches (high speed centrifugation and solubilisation buffers) and the same 7 cm IEF strips were used in both studies.

The analytical reproducibility of the 2-DE system was evaluated with respect to the protein species detected from each cell line preparation. Replicates gels were compared with protein extracted on two different occasions from cell lines with the same cellular passage. The coefficient of variation was low (CV≤ 0.03) in these cell lines; this data validates 2-DE as a system to enable separation of high as well as low abundance proteins with high resolution and with good reproducibility in all the four cell lines. This is consistent with findings that the separation method using immobilised pH gradient (IPGs) minimises gel to gel variation (Weiss & Gorg, 1998). Comparing the average number of protein spots detected in the four breast cell lines, showed that T47D had the highest number of proteins and HMT3522 the lowest, perhaps indicating that the expression of numerous proteins is accompanied with the development of a metastatic phenotype (Welsh *et al.,* 2003).

The biological reproducibility of the 2-DE system was also evaluated using proteins extracted from different cells with different cellular passage. The coefficient of variation was higher than those obtained for the analytical variation but it is worth noting that Parker *et al* (2006) have also shown that biological differences result in coefficient of variation of around 0.7. In this study the coefficient of variation was ≤0.06 for all analysis.

4.7.3 Protein 'profiles' across breast the cell lines

The protein profiles obtained from the 2-DE experiments allowed the comparison of between 400-500 protein 'spots' across all the cell lines. In this evaluation, proteins from T47D showing an average of 1.5 fold difference in levels or P≤0.05, were subjected to MALDI-MS and MS-MS analysis. The MASCOT search program compared the resulting peptide mass fingerprint with the theoretical peptide fingerprint generated *in silico* and produced a MOWSE score to give an evaluation of the significance of the matches, mass scores with P≤0.05 which were considered significant (Pappin *et al*., 1993). In addition proteins that were identified with MOWSE scores (≥40) were considered significant. Most of the proteins with elevated levels for example glutamate synthetase, HnRNPs, ENO 1, proteasome subunits, Hsp 27, GFAP, chaperonin subunits and Nm23 have previously been reported in malignancy (Arlt *et al*., 2009; Bhui-Kaur *et al*., 1998; Christa *et al*., 1994; Coghlin *et al*., 2006; Shiraishi *et al*., 1992; Steeg *et al*., 1993; Storm *et al*., 1996; Tsai *et al*., 2010; Zhang *et al*., 2008; Zhou *et al*., 2001). In this study for the first time the following proteins were reported as elevated in T47D; elongation factor Tu, Enoyl Coenzyme A hydratase 1 peroxisomal and macropain subunits.

4.7.4 Characterisation of HPA binding proteins

A 2-DE proteomic based approach was used to identify the HPA binding glycoproteins of T47D cells. This analysis identified molecules involved in apoptosis (Hsp 27) (Rane *et al.,* 2003), pre-mRNA splicing (HnRNP H1, HnRNP D-like, HnRNP A2/B1) (Gallinaro *et al.,* 1986), cellular remodelling (GFAP) (Kohama *et al.,*1995) and cell migration (ENO1) (Wygrecka *et al.,* 2009). It is assumed that the cytoplasmic HPA binding glycoproteins were present because they were complexed with HPA membrane proteins and thus were pelleted during membrane protein separation but the intracellular location of these proteins needs confirmation using other techniques. The proteins identified in this 2-DE work have been observed in other cancers (see table 4.3): for example heat shock protein (Hsp 27) has been reported to be increased in brain, breast and prostate cancer

(Andrieu *et al*., 2010 ; Kato *et al*., 1992); heterogenous nuclear ribonuclear protein H1 (HnRNP H1) in breast (Zhang *et al*., 2008); HnRNP A2/B1 in breast and lung (Zhou *et al*., 2001); glial fibrillary acidic protein (GFAP) in breast (Shiraishi *et al*., 1999) and enolase 1 (ENO1) in head and neck cancer (Tsai *et al*. 2010). Whilst many of these proteins have previously been linked to cancer, there has not been any reagent other than HPA which will bind all of these proteins simultaneously. It remains however unclear as to whether HPA recognises these cytoplasmic proteins via binding to GalNAc or GlcNAc containing epitopes. Prediction of potential glycosylation sites showed no N-linked, very few (≤8) O-GalNAc, and (≥4) O-GlcNAc sites within these proteins (Appendix 4). It will be interesting to investigate whether HPA recognises these proteins via O-GlcNAcylation.

In this study, the protein components of GP80 and GP130 were not identified with MADLDI-TOF-MS and MS-MS, because these were low abundance proteins and strategies such as affinity chromatography were not used to purify these proteins due to a number of cells required as starting material for this type of analysis. Previously Saint-Guirons *et al* (2007) identified the major components of that GP80 and GP130 to be integrin α6/αV and Hsp90 respectively. In an attempt to further investigate whether GP130 in T47D corresponded to integrin α6, T47D membrane proteins were blotted and probed with a monoclonal antibody directed against integrin α6. This experiment revealed that HPA recognises proteins of the same molecular weight as the integrin α6 subunit, consistent with the report of Saint-Guirons *et al* (2007) that the major component of GP130 was integrin α6. GP80 was common to both MCF-7 and T47D and it will be interesting to further investigate whether this protein corresponds to Hsp 90. These observations support the hypothesis that HPA recognises the same glycoproteins across both breast and CRC.

4.8 Conclusion

In conclusion, the work described the development of a sensitive, reproducible and robust methodology (based on 2-DE) enabling the separation and analysis of proteins. HPA was observed to recognise a range of proteins with diverse cellular functions ranging from apoptosis, pre-mRNA splicing, remodelling, to cell migration. To date, no reagent other than HPA has been described that can bind simultaneously to these proteins.

Chapter 5

Intracellular localisation and identification of HPA binding antigens in breast cancer cell lines

5.0 Intracellular localisation and identification of HPA binding partners in breast cancer cell lines

5.1 Introduction

In the previous chapter, the HPA binding glycoproteins of T47D were characterised. In this model, HPA recognised proteins from different cellular compartments: the cytoplasm and membrane. Despite the significance of HPA as a tool for predicting poor prognosis breast (Alam *et al*., 1990; Brooks & Leathem, 1991; Fenlon *et al*., 1987; Fukutomi *et al*., 1989; Leathem & Brooks, 1987) and other cancers (Kakeji *et al*., 1991; Schumacher *et al*., 1994; Shiraishi *et al*., 1992), the precise nature of the HPA-binding ligands and their defined role in the metastatic process has remained obscure.

It is known that native HPA preferentially binds terminal alpha α-N-acetylgalactosamine (αGalNAc) residues (Hammarstrom & Kabat, 1971; Lescar *et al*., 2007) and recent studies using a recombinant form of HPA have supported this observation (Markiv *et al*., 2011). Initially, therefore it was throught that HPA bound αGalNAc containing tumour associated antigens such as the Tn antigen (αGalNAc1 -O- Ser/Thr) (Springer, 1989) the Forssman antigen (αGalNAc1-3GalNAc) (Baker *et al*., 1983) or blood group A substance (GalNAcα1-3Galβ1-4GlcNAcβ1-Fucα1-2) (Anderson & Haas, 1984; Grundbacher, 1987; Mourali *et al*., 1980) in cancer cells. However, other studies have shown that its utility may lie in its ability to bind a much wider and heterogeneous array of N-acetyl galactosaminylated glycoproteins (Brooks & Leathem, 1995). On the other hand, further reports have highlighted that HPA binds GlcNAc (Wu & Sugii, 1991) and sialic acid containing epitopes (Dwek *et al*., 2001).

Aberrant expression of the Tn antigen has been observed in cancer with approximately 90% of human carcinomas having detectable levels of these epitopes (Springer *et al*., 1990). Whilst the presence of the Tn antigen has been shown in cancer, its functional role of in tumour metastasis has remained unclear. Springer (1989) proposed that the Tn antigen is the predominant structure

recognised by HPA in breast and other cancers, however, Grundbacher (1987) suggested that HPA recognises tumour cells by virtue of binding to blood group A antigen, the increase expression of which has been also reported in several cancers. Brooks and Leathem (1995), reported that in addition to the Tn epitope and blood group A antigen, HPA recognises related but distinct carbohydrate moeities sharing similar terminal αGalNAc groups. This absence of defined HPA epitopes has limited the application of HPA in the clinical oncology setting and there is, therefore, an opportunity to further investigate the complex epitopes recognised by HPA.

The synthesis of tumour associated carbohydrate structures such as Tn, is mediated through alterations in expression or activity of one or more of the enzymes of glycosylation. The events involved in the glycosylation processes are orchestrated by glycosidases and glycosyltransferases, these enzymes catalyse, respectively, the degradation and biosynthesis of glycans. Several studies have demonstrated altered expression of glycosyl epitopes and changes in levels of enzymes of glycosylation as related to tumour invasion and the metastatic process (Brockhausen, 2006; Dennis *et al*., 1999; Hakomori, 1984).

One group of enzymes that are of key importance in the glycobiology of breast cancer are the UDP-N-acetyl-α-D-galactosamine polypeptide: GalNAc-N-acetylgalactosaminyl transferases (ppGalNAc-Ts), the over-expression of which, has been reported in the malignancies of the breast (Berois *et al*., 2006; Freire *et al*., 2006). ppGalNAc-T*s* regulate the attachment of N-acetylgalactosamine (GalNAc) to Ser/Thr on the polypeptide backbone initiating mucin type *O*-linked glycosylation in the Golgi apparatus (Brooks *et al*., 2007; Cardone *et al*., 2005; Hassan *et al*., 2000). To date, fifteen enzymes of the ppGalNAc-T have been functionally identified in mammalian cells, and *in silico* studies suggest that as many 20 may exist (Ten Hagen *et al*., 2003; Ten Hagen *et al*., 1998). An increasing body of evidence suggests that while some members of the ppGalNAc-T family are broadly expressed in normal tissues, the distribution of others is restricted to certain tissues or organs (Bennett *et al*., 1998; Mandel *et al*., 1999; Young *et al*., 2003) . O-linked glycosylation in the Golgi apparatus is regulated by

differential expression of the enzymes that initiate O-linked glycosylation, thus ppGalNAc-Ts may be the initiating factor that ultimately gives rise the aberrant synthesis of immature carbohydrate structures such as Tn or sialyl Tn, frequently observed in cancer (Brooks *et al*., 2007; Cardone *et al*., 2005; Hassan *et al*., 2000). If this is indeed the case then glycosylated epitopes in cancer would be expected to be labelled with HPA in the Golgi apparatus.

Recently, Freire *et al* (2006), investigated the mRNA expression of 9 enzymes (ppGalNAc T1-T9) in breast tumours specimen and breast cancer cell lines (including MCF-7 and T47D using RT-PCR) and reported the expression of only ppGalNAc-T1, T2, T3 and T6. The increase expression of ppGalNAc-T6 correlated with breast cancer, thus the authors postulated that increase expression ppGalNAc-T6 may mark an early event associated with development of a metastatic phenotype. Immunolocalisation studies confirmed these observations with ppGalNAc T6 readily detectable in MCF-7 and T47D and non-detectable in HMT3522 and BT474 cells (Brooks *et al*., 2007). These findings suggest that increase expression of ppGalNAc T6 enzyme may be consistent with the presence of a wide range of immature GalNAc glycoforms in the Golgi apparatus.

Another family of enzymes of key interest in breast cancer biology are the sialyltransferases. The sialyltransferase α2,6-sialyltransferase I (ST6GalNAc I) transfers a sialic acid residue in α2,6-linkage onto GalNAcα1-*O*-Ser/Thr resulting in the formation of the sialyl-Tn antigen (STn) in the Golgi apparatus (Ikehara *et al*., 1999). The synthesis of STn has been reported to be increased in several different types of epithelial cancer (Kuwabara *et al*., 1997; Pinho *et al*., 2007; Schuessler *et al*., 1991; Wang *et al*., 2001) including breast cancer (Leivonen *et al*., 2001).

In an attempt to unravel the HPA binding partners observed in the breast cancer cell lines, we undertook several different approaches. The aim was to determine whether the ligands recognised by HPA include the blood group A determinant. Secondly, the binding of HPA in the perinuclear region of T47D and MCF-7 (reported in chapter 3) was evaluated to determine if it represented GalNAc

glycoforms in the Golgi apparatus. Thirdly, the finding (reported chapter 4) that HPA binds cytosol resident proteins, including transcription factors, led to an investigation as to whether HPA recognises O-GlcNacylated proteins in this model of cancer. Fourthly, a preliminary analysis of the expression of mRNA for key enzymes regulating O-glycosylation, such as ppGalNAc transferase (T1, T2 and T3, T6) and sialyltransferases (ST6GalNAc I and ST6GalNAc II) was performed in the cell lines. Lastly, the presence of STn (SAα2,6-GalNAcα-Ser/Th) and TF (Galβ1-3-GalNAcα-Ser/Th) were considered using *Sambucus nigra* agglutinin (SNA) and *Arachis hypogaea* agglutinin (peanut lectin, PNA) binding.

5.2 Binding of blood group A antibody to Western blots of T47D membrane proteins

The first step was to determine whether the glycoproteins recognised by HPA in T47D (established from individual with unknown blood group) contain the blood group A antigen or if they also carry other GalNAc containing glycans. Saint-Guirons *et al* (2007) previously reported a number of glycoproteins bearing the blood group A substance in the range of approximately 20-80 kDa in HT29 cells (established from blood group A individual). The work with HT29 cells revealed that some of the glycoproteins bearing the blood group A substance also interact with HPA.

To identify the glycoproteins bearing the blood group A substance in T47D membrane proteins, a murine monoclonal antibody directed against the blood group A antigen was used to probe Western blot of membrane proteins separated by SDS-PAGE from T47D cell line (as described in section 2.13.4). Detection was made with a secondary horseradish peroxidase (HRP) labelled goat anti-mouse antibody. The anti-A Western blot was compared with an HPA blot. A negative control (where the lectin and monoclonal antibody step was omitted) was also considered in this investigation. The antibody reacted weakly with a single glycoprotein band at approximately 85 kDa (GP 85), this broadly corresponded to the molecular weight of the HPA binding glycoprotein that was observed at approximately 85 kDa in T47D membrane protein fraction (both marked with

asterisks, figure 5.1). Interestingly, the predominant HPA binding glycoprotein did not interact with the antibody. Saint-Guirons *et al* (2007) previously showed an HPA binding glycoprotein at approximately 82 kDa which reacted with the blood group A antibody in HT29 cells. Whilst it had been considered possible that the same glycoprotein bearing the blood group A determinant would be recognised in T47D, only one protein carrying the blood group A epitope was observed in this cell line. The findings indicate that the predominant proteins recognised by HPA in T47D, bear glycosylation motifs distinct from blood group A antigen.

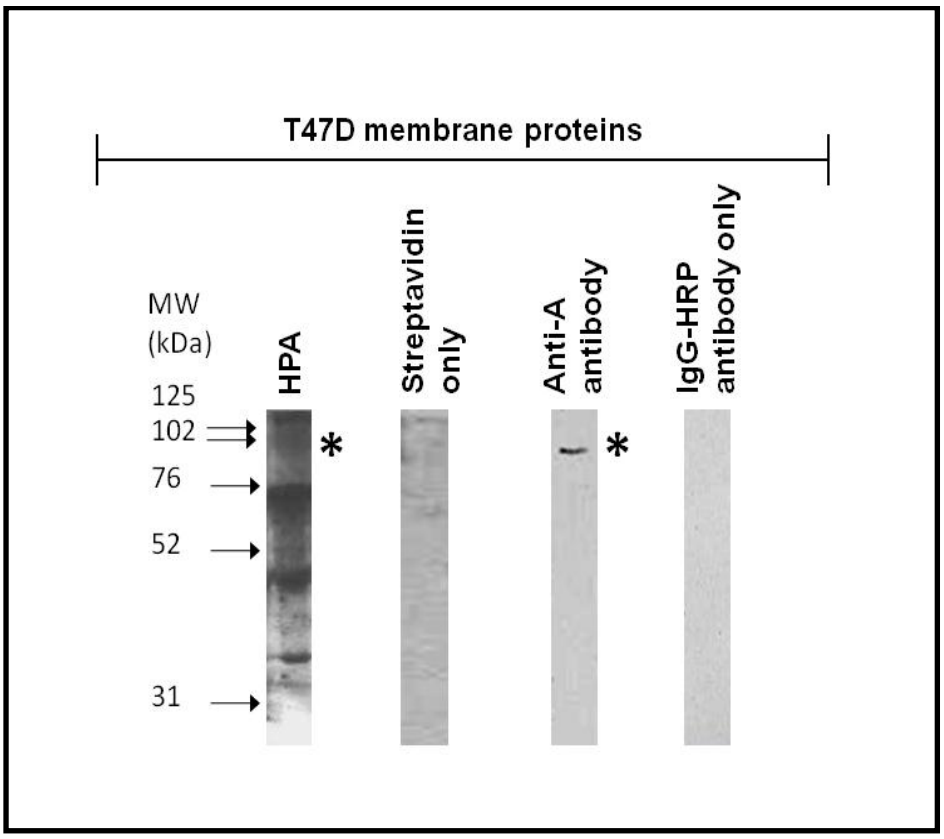


Figure 5.1: Anti-A antibody binding to T47D membrane proteins. Proteins (20 µg) were separated by SDS-PAGE, transferred by nitrocellulose and the Western blot was probed with biotinylated HPA followed by streptavidin-HRP or a goat anti-mouse HRP labelled antibody. The anti-blood group A antibody detected a single protein band (showed by asterisks) in the T47D membrane protein fraction.

5.3 O-GlcNAc epitopes in the breast cancer cell lines

Aberrant O-GlcNAcylation of cytoplasmic and nuclear pore proteins is a common feature accompanying metastatic transformation (Chou *et al*., 2001; Slawson *et al*., 2010 Slawson *et al*., 2008). The occurence of O-GlcNAcylated epitopes and their HPA binding properties was studied in the breast cells lines. In the first section of this investigation, an anti-human O-GlcNAc antibody was used to investigate the O-GlcNAc containing staining pattern across the breast cancer cell lines, HMT3522, BT474, MCF-7 and T47D, using confocal microscopy. Secondly, a co-localisation study using HPA and anti-O-GlcNAc antibody was undertaken using T47D cells. Finally, the O-GlcNAcylated protein of the membrane and cytoplasmic fraction of T47D were assessed following separation by SDS-PAGE and probing of Western blots. The 2-DE approach was also used to further investigate the membrane glycoproteins recognised by both HPA and the O-GlcNAc antibody of T47D.

5.3.1 O-GlcNAc labelling in breast cells

The cells were incubated with a murine antibody against human O-GlcNAc followed by a FITC labelled IgG antibody (as described in section 2.3.2), the O-GlcNAc labelling was observed using confocal microscopy. The O-GlcNAc cytochemistry revealed that all the four breast cancer cell lines synthesised O-GlcNAc containing glycoconjugates (figure 5.2). However, a marked quantitative difference in O-GlcNAc labelling pattern was observed across the cell lines, varying from intense (T47D), to moderate (MCF-7), to weak or negligible (HMT3522 and BT474). Interestingly, the levels of synthesis of O-GlcNAcylated glycoconjugates correlated with the metastatic potential of the cells, with the metastatic cell line T47D exhibiting higher levels of synthesis of O-GlcNAcylated glycoconjugates than the non-metastatic cell line HMT3522. The binding was mainly cytoplasmic but some occasional binding to the cell surface of T47D was also noted. Omission of the antibody resulted in no labelling in those cells (Appendix 6).

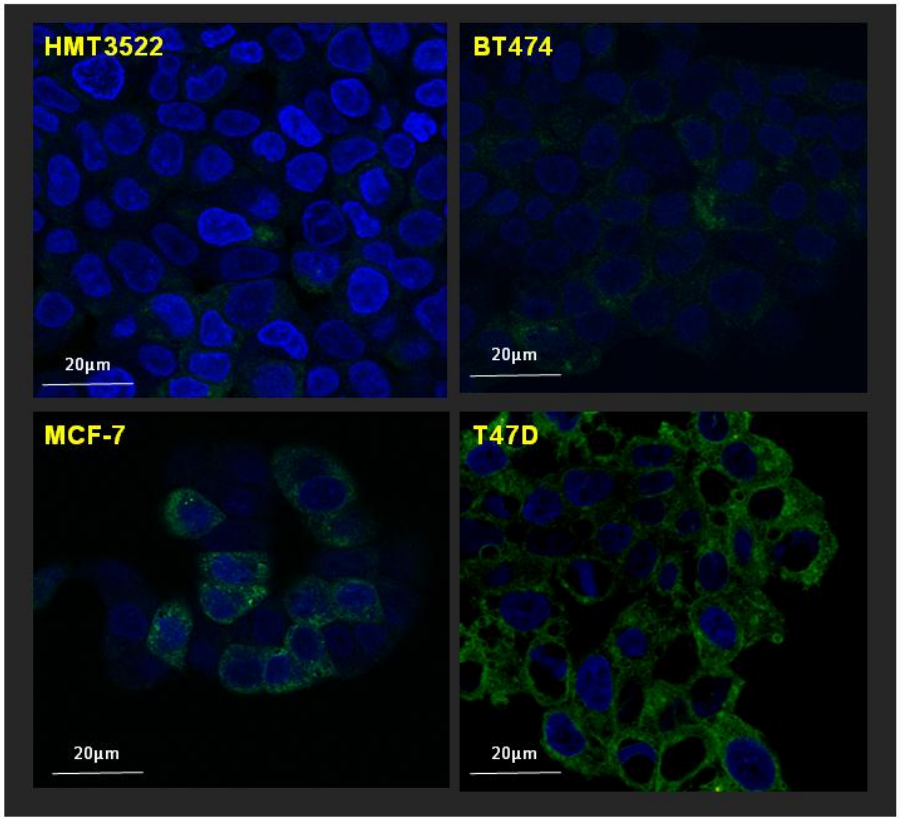


Figure 5.2: Confocal images showing O-GlcNAc/IgG-FITC labelling in the four breast cell lines. The nucleii was counter-stained with the nuclear label To-Pro-3 (blue). A murine antibody against human O-GlcNAc followed by a FITC labelled IgG antibody (green) was used. HMT3522 and BT474 showed occasional weak granular staining, whereas MCF-7 showed moderate binding and T47D showed intense cytoplasmic as well occasional binding to the cell surface. Scale bars = 20µm

5.3.2 HPA co-localisation with O-GlcNAcylated epitope

In this section, we aimed to further assess HPA binding and correlate this with O-GlcNAcylated epitopes through co-localisation studies in T47D cells. Cells were grown and fixed (as described in section 2.3). The cells were incubated with TRITC labelled HPA followed by an O-GlcNAc/IgG-FITC antibody step. The binding was observed using confocal microscopy (figure 5.3).

Intense HPA labelling (A) and anti-O-GlcNAc antibody (B) was observed in T47D cells. Strong co-localisation of HPA and O-GlcNAc binding epitopes were noted in the cytoplasmic region of T47D (C, mark with green arrows) with binding also observed on the cell surface (mark with yellow arrows). This result suggested that HPA has the ability to recognise cytoplasmic O-GlcNAcylated glycoproteins associated with the metastatic phenotype of T47D.

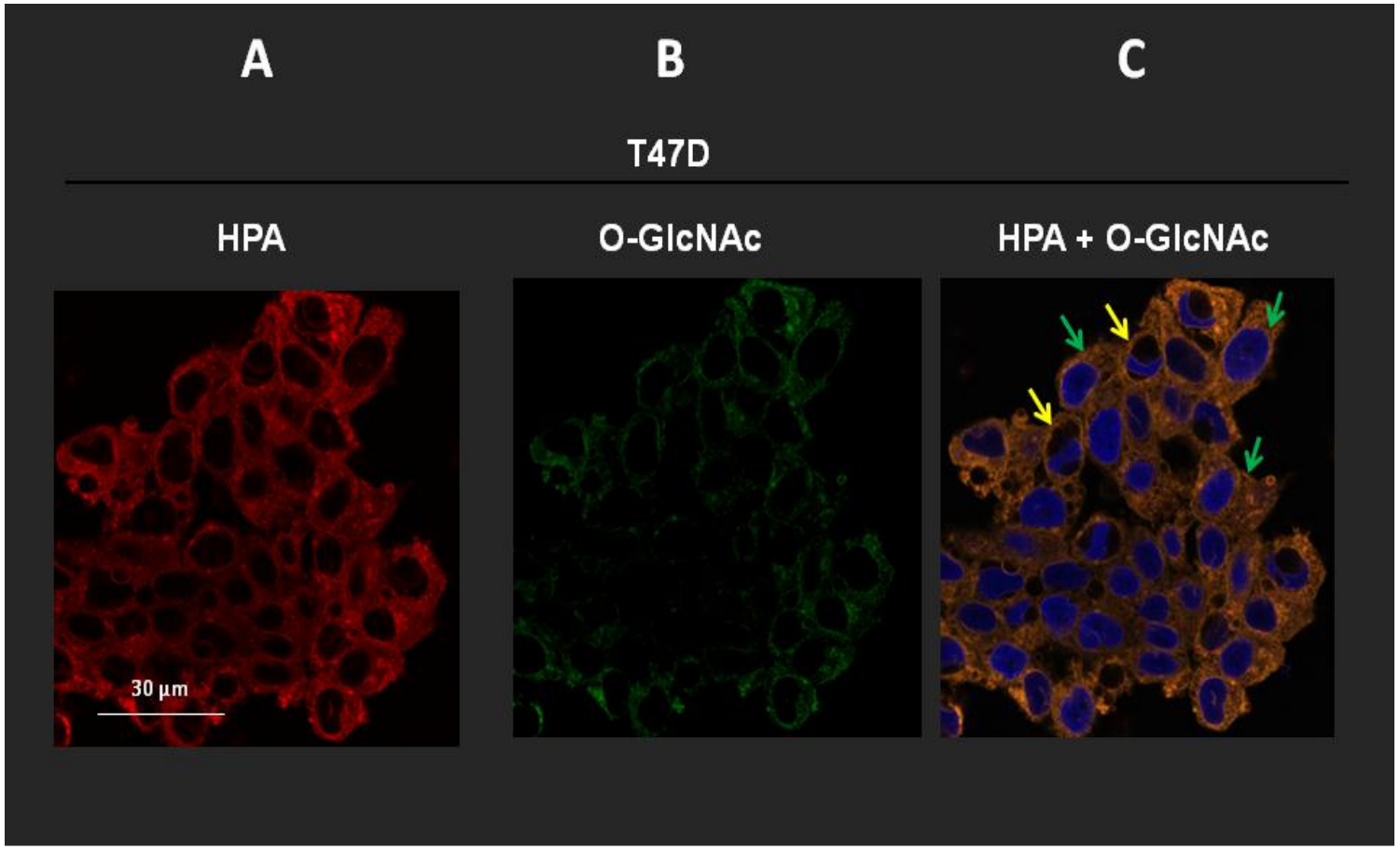


Figure 5.3: Co-localisation of HPA and O-GlcNAc epitopes in T47D cells. The images showed strong HPA binding in T47D (red, panel A) and anti-O-GlcNAc antibody binding (green, panel B) and overlayed image of HPA and anti-O-GlcNAc antibody binding (orange, panel C), intense cytoplasmic (green arrows) and occasional cell surface labelling (yellow arrows) in T47D. Scale bars = 30µm.

5.3.3 Anti-O-GlcNAc antibody probing of Western blots of T47D membrane proteins

Cytoplasmic and membrane proteins of T47D were separated by SDS-PAGE, blotted to nitrocellulose and probed with a murine anti-O-GlcNAc antibody and detected using an HRP labelled secondary antibody raised in goat (as described in section 2.13.5). A negative control where the antibody step was omitted was also considered.

The Western blot probed with the anti-O-GlcNAc antibody was compared with the reference HPA blot (figure 5.4). The anti-O-GlcNAc antibody reacted with a number of proteins species in the range of approximately 30-150 kDa, with three major species (marked in asterisks) noted at approximately 35 kDa, 50 kDa and 70 kDa in the cytoplasmic fraction of T47D. In the membrane fraction only three weakly binding protein bands were observed to interact with the antibody, the molecular weights of these corresponded with three major HPA binding proteins (marked in asterisks) at approximately 35 kDa, 50kDa and 70 kDa. These observations suggest that most of the HPA binding proteins at lower molecular weights in the membrane fraction are O-GlcNAcylated and may represent contamination of membrane proteins with cytoplasmic proteins.

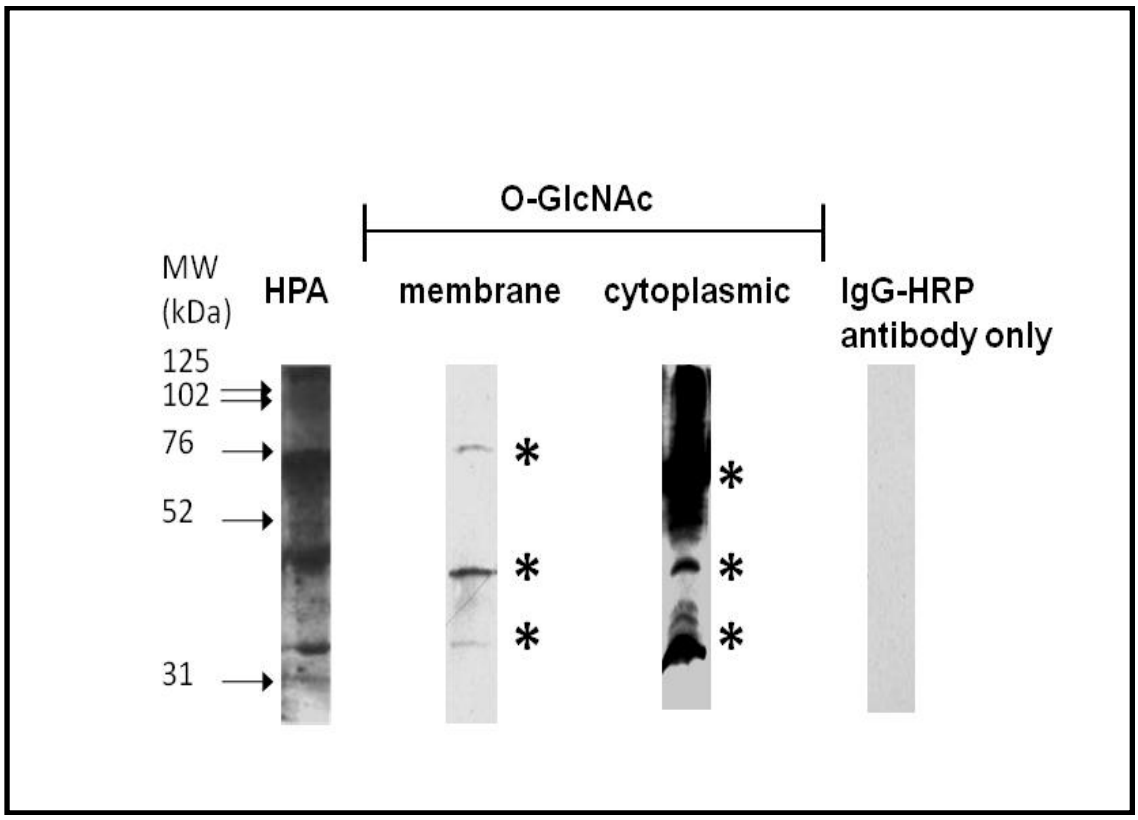


Figure 5.4: Anti-O-GlcNAc antibody binding to T47D cytoplasmic and membrane proteins. T47D cytoplasmic and membrane proteins (20 µg) were separated by SDS-PAGE and transferred to nitrocellulose and the Western blots were probed with an anti-O-GlcNAc antibody followed by a secondary HRP labelled antibody. The anti-O-GlcNAc antibody detected proteins in the range of approximately 30 to 150 kDa with three intense bands at MW of 35 kDa, 50 kDa and 75 kDa, these O-GlcNAcylated proteins were also observed in the membrane fraction of T47D (marked by asterisks). No bands were observed when the blots were incubated with secondary antibody alone.

Figure 5.4 showed three O-GlcNAc containing glycoproteins in the membrane protein of T47D. To characterise these further, Western blots of 2-DE separated proteins were probed using the strategy described above. This approach revealed eight proteins that were recognised by the O-GlcNAc antibody (Figure 5.5). The HPA binding proteins circled in blue (panel A) closely corresponded to the molecular weight and position of O-GlcNAcylated protein species (also circled in blue). For instance the protein species at 75 kDa correlated with position of HnRNP H1 and the ENO 1 and those at the lower molecular weight correlated with HnRNP D-like, HnRNP A2/B1 and Hsp27 proteins identified in chapter 4. However, GFAP (circled in yellow) were not detected on the O-GlcNAc blot and other proteins, not recognised by HPA (circled in red) were detected by the O-GlcNAc antibody. In summary, in this experiment, HPA binding glycoproteins, HnRNPs, Hsp27 and ENO1 appeared to be recognised by virtue of their O-GlcNAcylation.

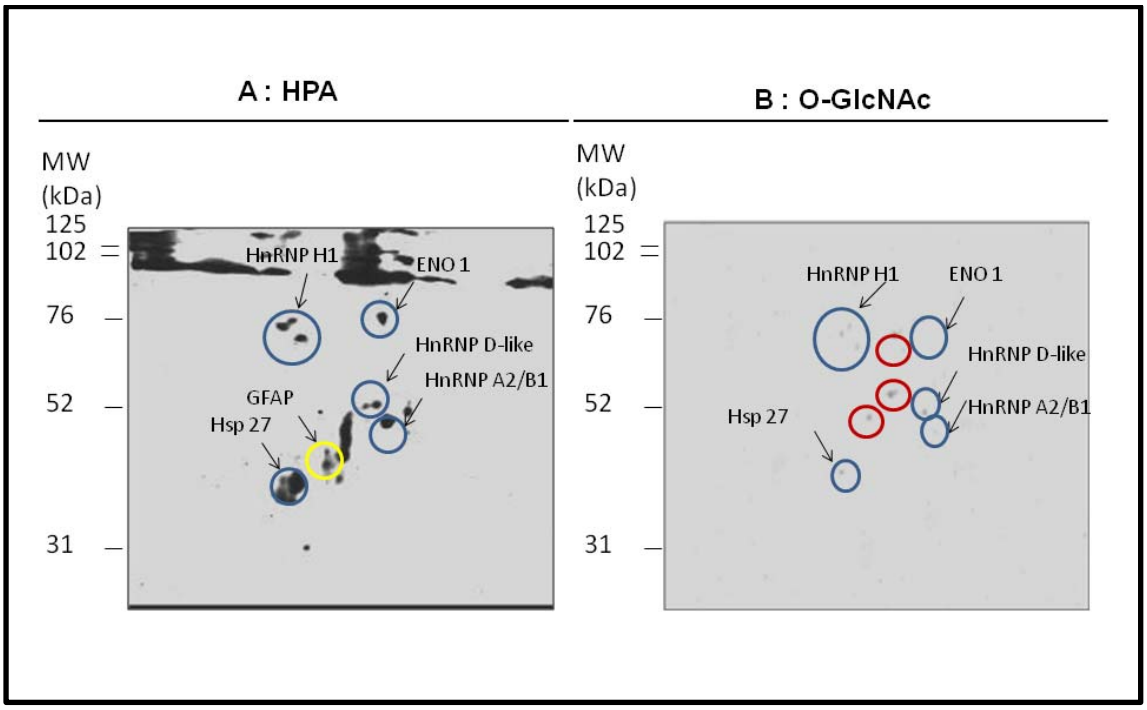


Figure 5.5 Anti-O-GlcNAc antibody binding to T47D membrane proteins. T47D membrane proteins were separated by 2-DE and analysed by Western blotting. Representative HPA and O-GlcNAc blot is shown in panel A and B respectively. The blots were detected with streptavidin-HRP and IgG-HRP respectively. The anti-O-GlcNAc antibody bound to eight proteins. The proteins circled in blue were noted both in the HPA and O-GlcNAc blot, whereas proteins circled in yellow and red were detected only in HPA and O-GlcNAc blots respectively.

5.3.4 Specificity of GlcNAc binding

The specificity of HPA binding was assessed by preincubation of the lectin with 100 mM freshly prepared GlcNAc. In this approach pre-incubation of the lectin with GlcNAc was observed to abrogate the binding of HPA to almost all the binding partners in T47D (figure 5.6). Many of the HPA binding proteins, including the O-GlcNAcylated proteins were not detectable when HPA was pre-incubated with GlcNAc (panel B). This is consistent with the observations made in the 1-DE Western blot analysis, where HPA binding was completely abolished following pre-incubation of the lectin with 100 mM GlcNAc (chapter 3). Unlike the observations made previously, GP80 was still observed. These observations suggest that the majority of the HPA binding proteins, except GP80, encompass glycans which appear to be recognised by the lectin through GlcNAc: HPA interactions.

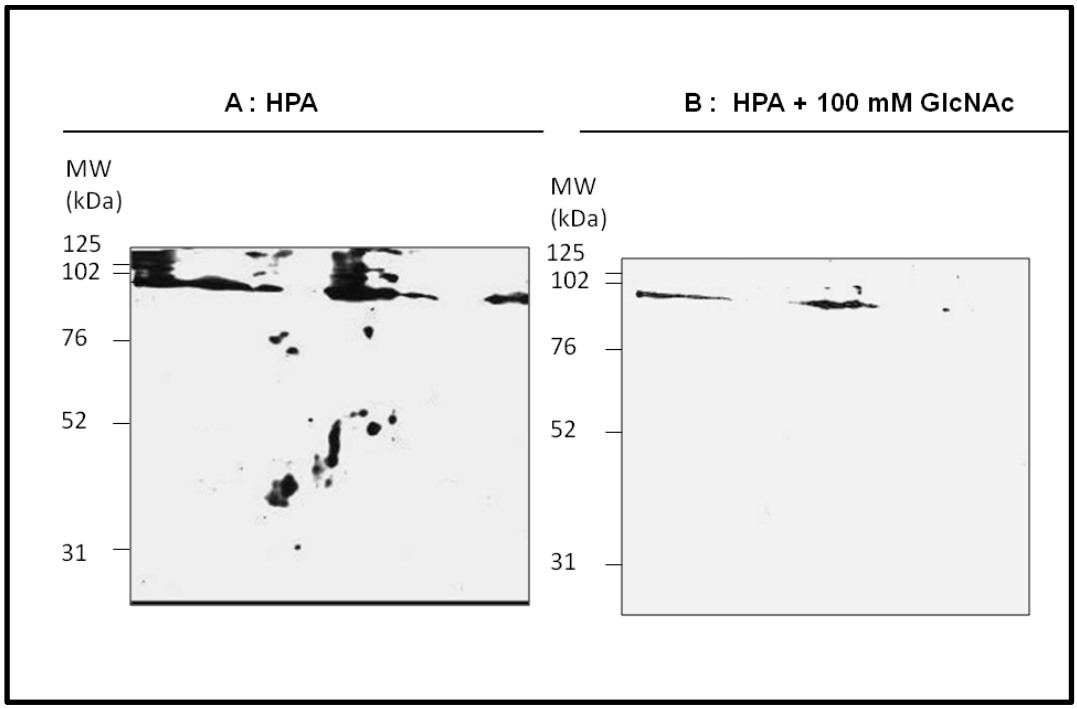


Figure 5.6: Inhibition of HPA binding to O-GlcNAcylated glycoproteins. Representative HPA (panel A) and HPA pre-incubated with 100 mM GlcNAc (panel B) blots of T47D membrane proteins. T47D membrane proteins were separated by 2-DE transferred to nitrocellulose by Western blotting and probed with biotinylated HPA which has pre-incubated with 100 mM of GlcNAc and detected using streptavidin-HRP. Most of the HPA binding proteins, except GP80 were no longer detectable when the HPA was pre-incubated with GlcNAc.

5.4 Co-localisation of HPA binding in the Golgi apparatus using a Golgi tracker dye

Initiation of mucin type *O*-linked glycosylation occurs in the Golgi apparatus and aberrant activity of the glycosyltransferase or glycosidase enzymes may be reflected by an increase in the levels of a wide range of GalNAc containing glycoforms in this cellular compartment (Brooks & Leathem, 1995). In chapter 3, punctuate HPA labelling in the perinuclear region was observed in MCF-7 and T47D cells. Here, we set out to identify whether HPA binding localised to the Golgi apparatus in MCF-7 and T47D by co-localisation of HPA binding with a Golgi tracker dye.

A Golgi tracker dye ($NBDC_6$-Ceramide) was used to selectively stain the sphingolipids of the trans–Golgi compartments (Ktistakis *et al.,* 1995). Cells were grown, fixed and stained (as described in section 2.3.3), the co-localisation of HPA-TRITC and $NBDC_6$-ceramide was assessed using fluorescence microscopy. A 3D model of HPA binding was prepared to assist visualisation of the interaction.

HPA and $NBDC_6$-ceramide dye binding was located in the same region of the cells. T47D exhibited slightly more intense binding than MCF-7 (indicated by arrows, figure 5.7 and figure 5.8). This finding indicated that HPA binding in the perinuclear region of these cell lines was associated, at least in part, with the localisation of HPA binding epitopes in the Golgi apparatus.

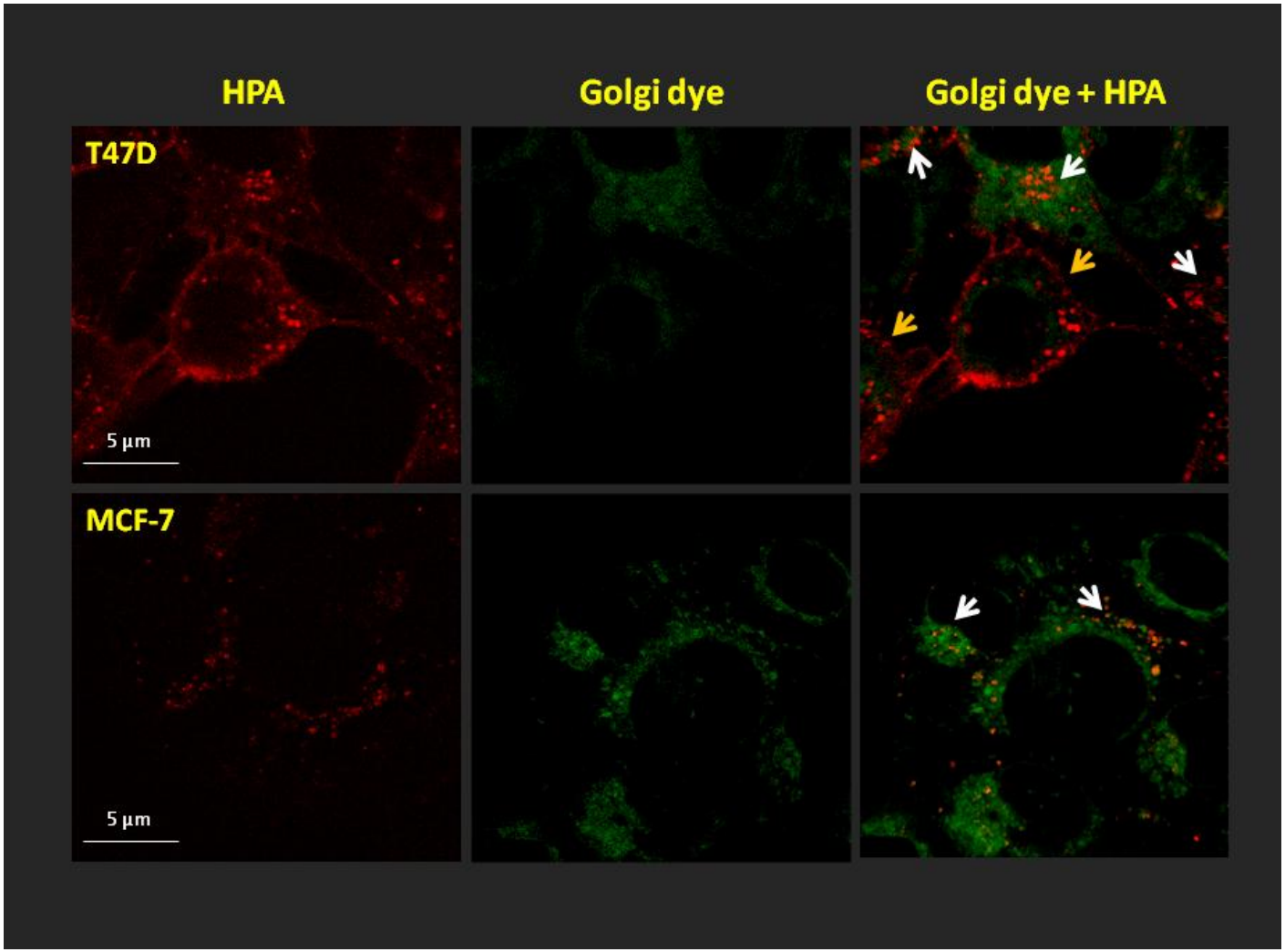


Figure 5.7: Localisation of HPA and Golgi tracker dye binding in T47D and MCF-7. Confocal images showing the binding of TRITC-labelled HPA (red) to the cell surface (yellow arrows), perinuclear region (white arrows) and the Golgi tracker dye (green) in MCF-7 and T47D cell lines. Images show an overlay of HPA and Golgi tracker dye binding. Areas of co-localisation are shown in (orange). Scale bars = 5µm.

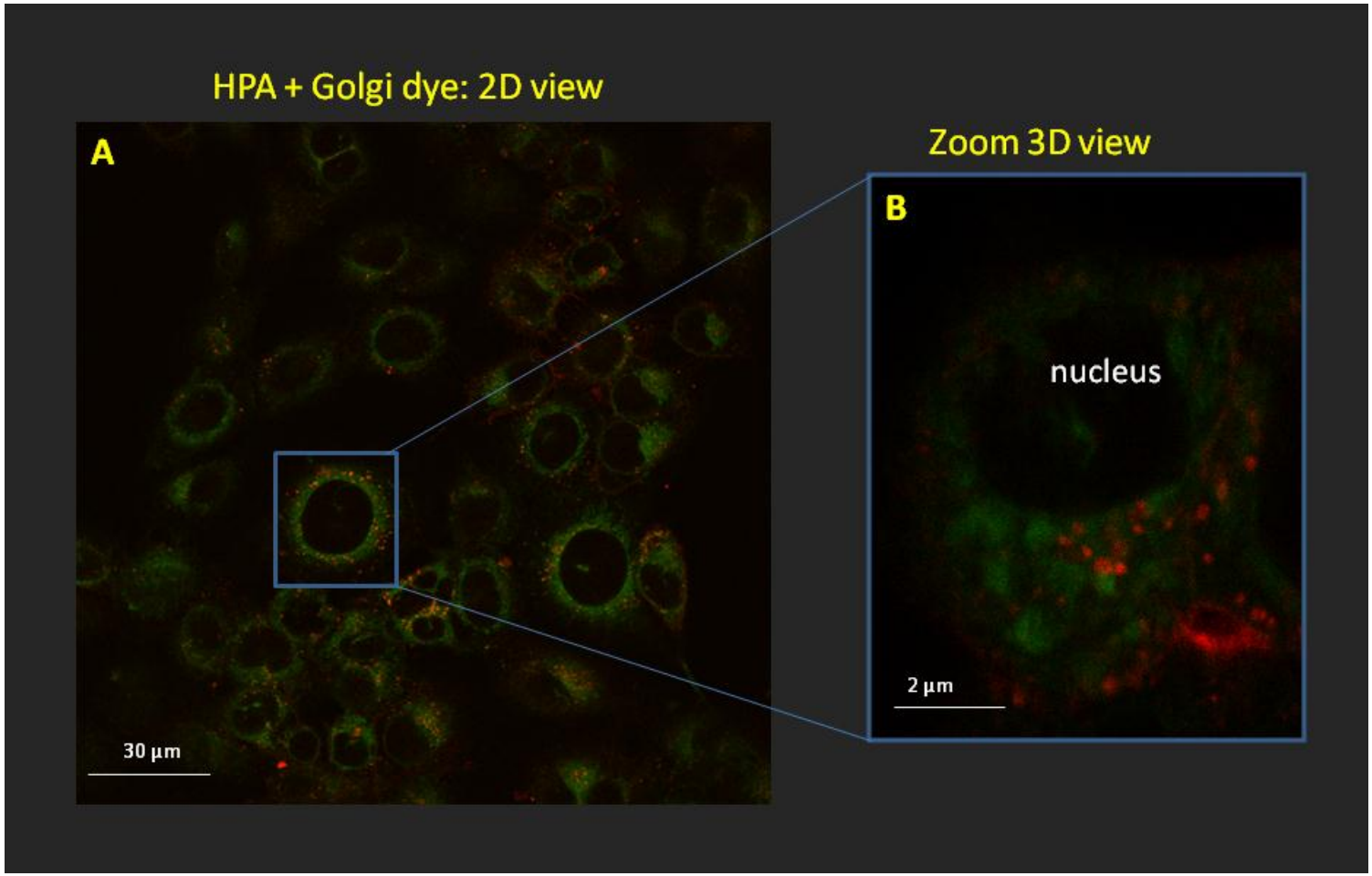


Figure 5.8: 3D view of HPA and Golgi tracker dye binding in MCF-7. Zoom image showing the binding of TRITC-labelled HPA (red) and the Golgi tracker dye (green) in MCF-7 cell lines. Panel A: 2D Surpass view images showing HPA and Golgi tracker dye binding in MCF-7 cells Panel B: 3D view of a single MCF-7 cell showing co-localisation of HPA with a Golgi tracker dye (orange). Scale bars A= 30µm and B=2 µm.

5.5 Glycosyltransferases expression in the breast cell lines

As described in section 5.1 ppGalNAcTs catalyse the addition of a GalNAc residue to an O-linkage to give rise to the Tn epitope (Brooks *et al*., 2007; Cardone *et al*., 2005; Hassan *et al*., 2000). ST6GalNAc I and β3GalT (T-synthase) transfer a sialic acid or a galactose residue respectively to the Tn structure to give rise to STn and the TF antigen (core 1) (Ikehara *et al*., 1999; Ju *et al*., 2002). Whilst the formation of the core 1 structure is most common in normal human cells (Ju *et al*., 2002), formation of STn is less frequently observed and its occurence in cancer has been shown to correlate with the development of a metastatic phenoytype (Sewell *et al*., 2006). Abrogation of glycosyltransferase expression, for example, up-regulation of expression of mRNA for *ppGalNAcTs* or down-regulation of mRNA for *ST6GalNAc I* genes or *β3GalT* may lead to increased exposure of the Tn epitope or incomplete synthesis of the sialyl Tn or TF antigens, respectively. The schematic diagram below (figure 5.9) illustrates the enzymatic reactions that catalyse the formation of Tn and STn or TF antigen.

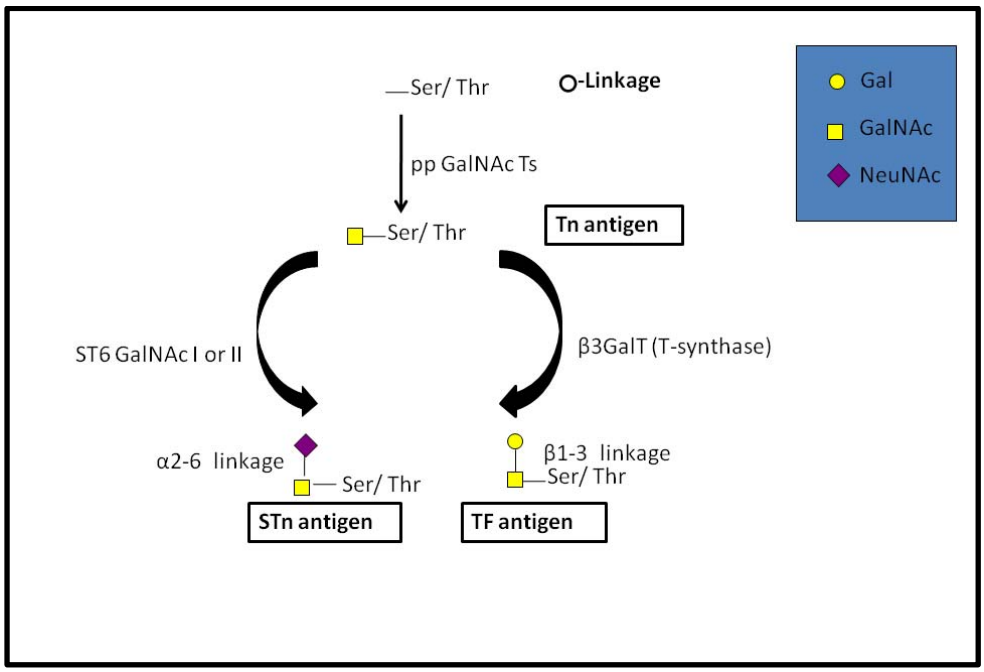


Figure 5.9: Schematic illustration of the formation of Tn , Sialyl Tn, TF antigen.

To correlate the expression of *ppGalNAcTs* and *ST6GalNAcs* with HPA binding, a study of gene expression study was undertaken using the HMT3522, BT474, MCF-7 and T47D cells. The breast cells were grown to near confluency and mRNA was extracted and converted into cDNA (as described in section 2.18). A quantitative system using q-RT-PCR was used to assess the expression of the genes. The cycle theshold (C_T) was determined at a theshold fluorescence value of 0.2. The expression levels of the *glyceraldehyde-3-phosphate dehydrogenase* (*GAPDH*) gene were compared with *β-actin* gene in the breast cancer cell lines. Next expression levels of *ppGalNAc T1, T2,T3, T6* and *ST6GalNAc I/II* genes were also examined in all these breast cell lines and compared with the reference gene. The melting curve of the amplicons for all the six genes is shown in Appendix 8.

5.5.1 Validation of reference genes as internal control for quantitative RT-PCR: *β-actin* compared with *GAPDH*

Validation of q-RT- PCR results requires accurate normalisation of the PCR data against a reference gene (internal control). Ideally, the reference gene should display uniform expression, regardless of the phenotypic characteristics of the cell lines or tissues (Berois *et al.,* 2006).

The expression of *GAPDH* varied across the breast cell lines, with relatively high levels of expression observed in HMT3522 and BT474 and comparatively lower expression observed in MCF-7 and T47D (figure 5.10). However the *β-actin* gene appeared to be uniformly expressed across the four breast cells studied; this is consistent with previous observations made by Freire *et al*(2006) in breast cell lines, including T47D and MCF-7 used in this study. As the *β-actin* gene was constitutively expressed in the breast cell lines, it was considered to be an appropriate reference gene for the q-RT-PCR data analysis undertaken in the current study.

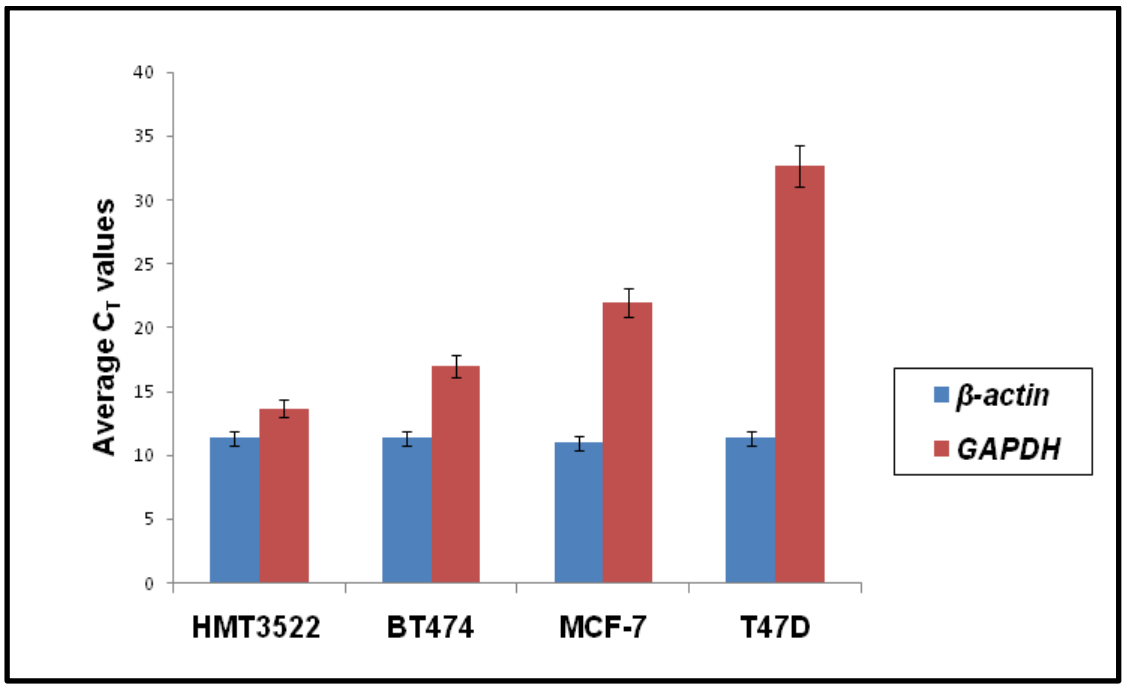


Figure 5.10: mRNA expression of *β-actin* and *GAPDH* reference genes in the breast cell lines. RT-PCR cycle theshold (C_T) values for *β-actin* and *GAPDH* reference genes for the breast cell lines., mean average and +/- SD of 6 replicates. C_T values are inversely proportional to the transcript level (gene expression). Expression of *GAPDH* was variable whilst stable expression of *β-actin* was observed in the breast cell lines.

5.5.2 Expression of *ppGalNAc T1*, *T2, T3* and *T6* genes in breast cells

Freire *et al* (2007) previously reported that ppGalNAc T1, T2, T3 and T6 enzymes were expressed in breast cancer cells, with an elevated mRNA expression of *ppGalNAc T6* also noted in breast cancer tissues. Brooks *et al* showed that ppGalNAc T3 and T6 enzymes were detected only in metastatic cell lines with metastatic phenotype, including, MCF-7 and T47D used in this study. In this experiment, we investigated the gene expression of four *pp-GalNAc-Ts*: *T1*, *T2*, *T3* and *T6* in the breast cell lines by RT-PCR. The expression of the target genes was normalised to the *β-actin* reference gene to take into account any differences in the amount of cDNA in the starting mixture.

In this study the *ppGalNAc-T1*, *T2*, *T3* and *T6* genes were expressed across all the four cell lines (figure 5.11), consistent with previous reports where those genes

are heterogeneously expressed in breast tissues (Berois *et al*., 2006, Freire *et al*., 2006). Unlike previous findings, this study did not observe an increase in the expression of *ppGalNAc-T6* in the metastatic breast cancer cells (T47D and MCF-7), instead, HMT3522 expressed highest level of all the four genes, BT474 slightly lower and all the cell lines derived from metastatic breast expressing comparatively lower level of expression. All the four genes showed similar expression levels across all the breast cell lines with *ppGalNAc-T2* expressed in higher levels followed by *ppGalNAc-T1*, then *ppGalNAc-T3* and *ppGalNAc-T6*. The product of the RT-PCR was run on an agarose gel to confirm the length of the amplicons (figure 5.12). The intensity of the bands for *ppGalNAc-T3* and *ppGalNAc-T6* was however higher in the all the breast cell lines and this does not correlate with the q-RT-PCR work.

In summary, the experimental data did not correlate with the HPA binding status of the cell lines and hence the phenotype of the cells. Furthermore, these findings contradicted previous reports and there is clearly a need for further investigation to validate these findings more fully.

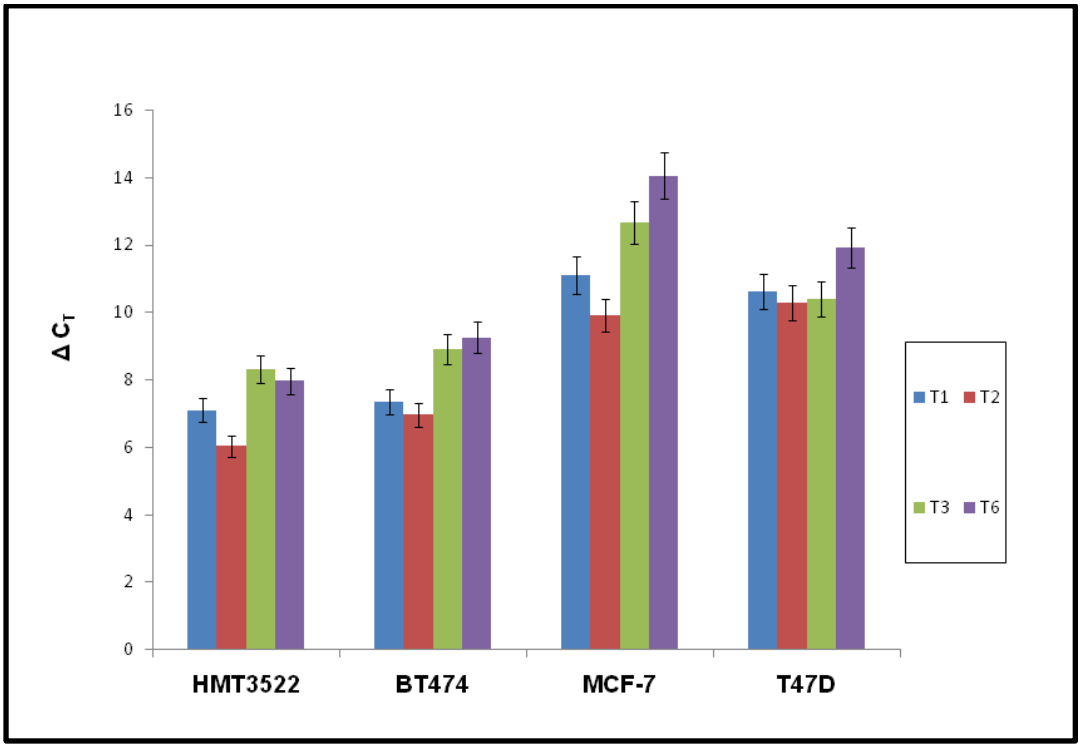


Figure 5.11: mRNA expression of *ppGalNAc T1, T2, T3 and T6* genes in the breast cell lines. ΔC_T values, mean average and +/- SD of 6 replicates. C_T values are inversely proportional to the transcript level (gene expression). RT-PCR data was normalised to the β-actin gene expression. In all cases, *T2* was more highly expressed, followed by *T1*, then *T3* and *T6* in the breast cell lines. Expression of all the four genes was higher in HMT3522, followed by BT474 and then T47D and MCF-7.

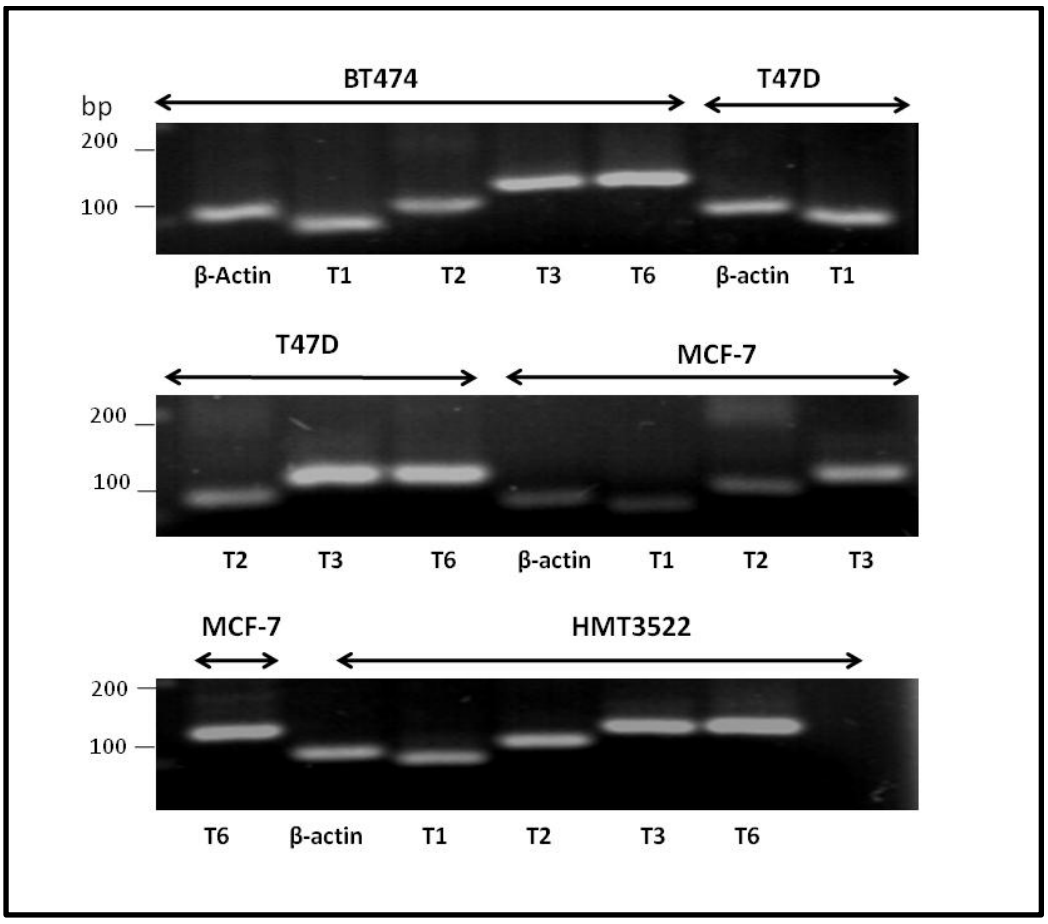


Figure 5.12: Representative agarose gel for pp-*GalNAcTs* RT-PCR product separation. 5 µg of DNA amplicon was loaded on agarose gel. Detection was performed with 1% w/v ethidium bromide. The RT-PCR products for *pp-GalNAc T3 and T6* appeared as the sharpest bands in the breast cell lines HMT3522, BT474, MCF-7 and T47D. The length of amplicon for *ppGalNAc T1, T2, T3 T6* and *β-actin* was calculated as 96 bp, 114 bp, 129 bp, 129 bp and 98 bp respectively.

5.5.3 Expression of *ST6GalNAc I* and *II* genes in breast cells

The expression levels of core 1 enzymes, *ST6GalNAc I* and *II* in the breast cancer cell lines was assessed. ST6GalNAc I transfers a sialic acid residue in alpha 2-6 linkage onto the Tn structure to yield sialyl Tn and failure of this enzyme may results in an immature Tn structure. Althrough, ST6GalNAc I is the primary synthase of the STn epitope (Marcos *et al*., 2004), it has also been reported that ST6GalNAc II may also act as a secondary candidate synthase. In this experiment, we sought to investigate whether these enzymes are down-regulated in metastatic breast cancer, thus leaving exposed Tn antigen.

Gene expression analysis revealed an absence of expression of *ST6GalNAc I* enzymes in all the four breast cell lines consistent with previous findings (Julien *et al*., 2001). However, the *ST6GalNAc II* gene was expressed across all the four breast cell lines with the highest level observed in MCF-7 and lowest in T47D (figure 5.13). These observations showed no correlation with the HPA binding status of the cell lines. The product of RT-PCR was run on an agarose gel to confirm the length of the amplicons (figure 5.14).

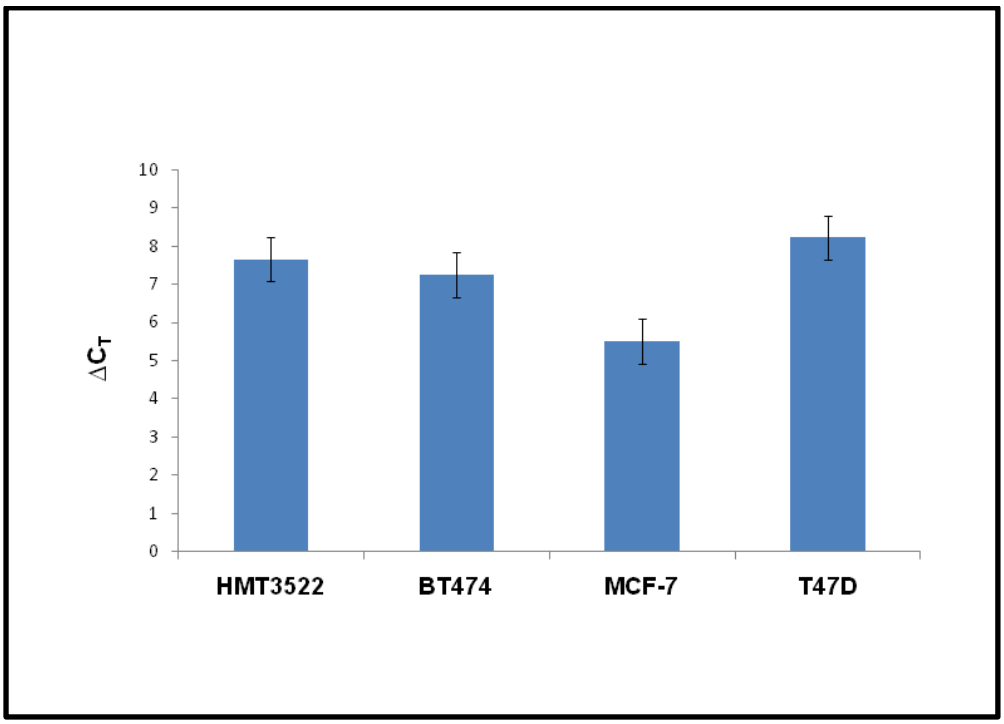


Figure 5.13: mRNA expression of *ST6GalNAc II* gene in the breast cell lines. ΔC_T values, mean average and +/- SD of 6 replicates. C_T values are inversely proportional to the transcript level (gene expression). RT-PCR data was normalised to the β-actin gene expression. *ST6GalNAc II* gene was expressed in all the breast cell lines, however higher expression was noted in MCF-7 and lowest expression was observed in T47D.

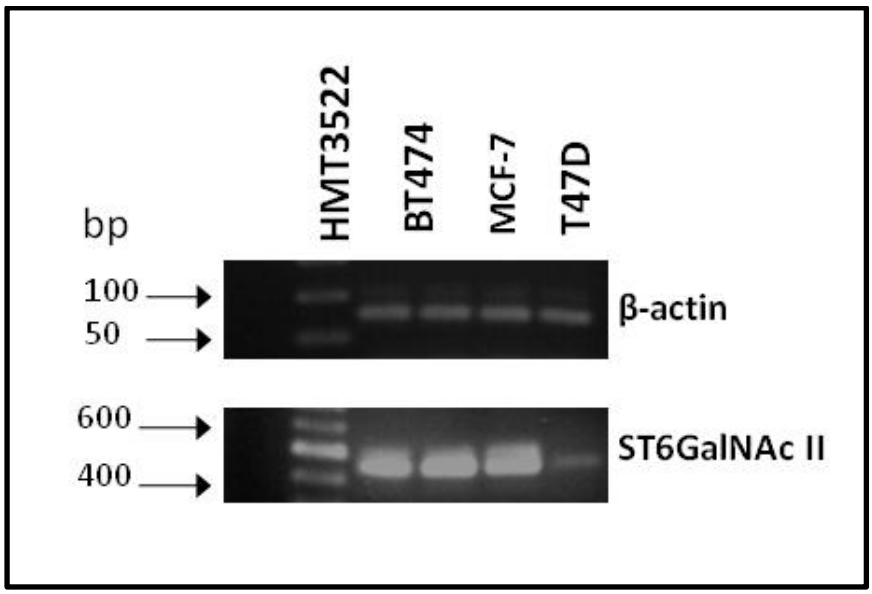


Figure 5.14 Representative agarose gel for ST6GalNAc *II* product separation. 5 µg of DNA amplicon was loaded on agarose gel. Detection was performed with 1% w/v ethidium bromide. The RT-PCR products for *ST6GalNAc II* and *β-actin* appeared as the sharp bands. The length of amplicon for *ST6GalNAc II* and *β-actin* was calculated as 550 bp and 98 bp respectively.

5.5 SNA and PNA labelling in breast cells

In this part of the study, the binding of *Sambucus nigra* agglutinin (SNA) and *Arachis hypogaea* agglutinin (PNA) to the cell lines was assessed. SNA and PNA have affinity for 2,6 linked sialic acid (such as those found in STn) and Galβ1-3-GalNAcα-Ser/Th (Alam *et al*.,1990) containing epitopes, respectively, the synthesis of which have also been correlated with the development of a metastatic phenotype (Campbell *et al*., 1995; Murayama *et al*., 1997; Ryder *et al*., 1992; Sata *et al*., 1991; Slovin *et al*., 2005).

Marked differences in cytoplasmic SNA and PNA labelling was observed across the breast cell lines HMT3522, BT474, MCF-7 and T47D (figure 5.15). SNA labelling ranged from intense (T47D), moderate (MCF-7) to negligible or weak in HMT3522 and BT474. Similarly, PNA labelling ranged from intense in T47D and moderate in MCF-7 and weak in BT474 and HMT3522. Both of these observations were consistent with the HPA binding status as well as the phenotypic characteristics of the cell lines with the metastatic cell lines binding to SNA and PNA and the non-metastatic cell lines HMT3522 and BT474 showing reduced levels of binding to the lectins. This study suggests that the metastatic cell lines T47D and MCF-7 may contain T antigen and sialyl Tn epitopes.

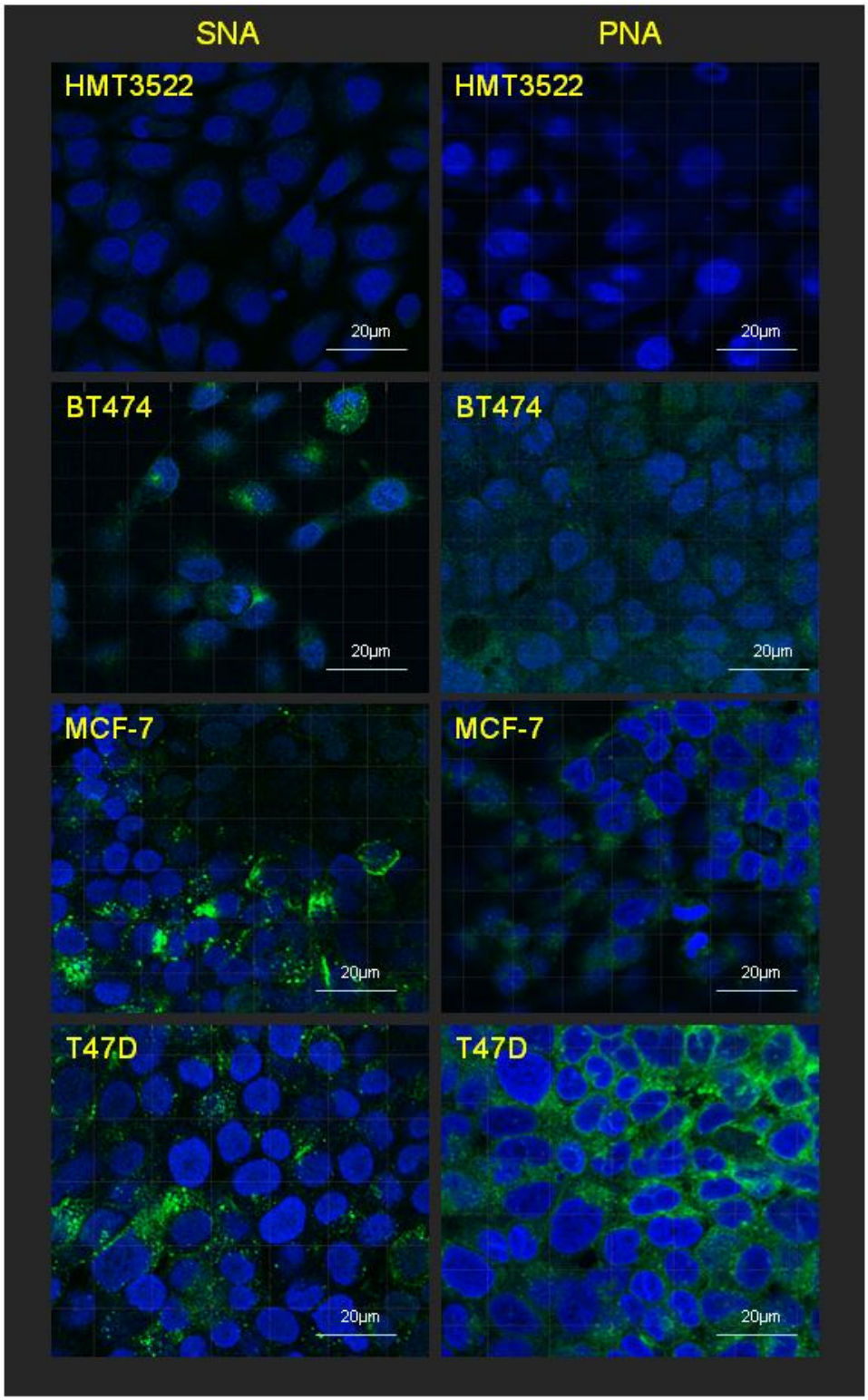


Figure 5.15: SNA and PNA labelling in HMT3522, BT474, MCF-7 and T47D. The cells were counter-stained with To-Pro-3. HMT3522 and BT474 showed occasional or 'negligible' granular staining. MCF-7 showed moderate labelling whereas T47D showed intense granular staining. MCF-7 showed intense granular staining in the perinuclear region. Scale bars = 20µm

5.6 Discussion

5.6.1 HPA interaction with blood group A antigen

It has long been known that HPA has the ability to recognise the blood group A determinant (Anderson & Haas, 1984; Grundbacher, 1987; Mourali *et al*., 1980). Whilst the aberrant synthesis of blood group antigens has been well established in cancer, it has been suggested that the prognostic utility of HPA lies in its ability to predominantly recognise blood group A determinant in tumours (Grunbacher *et al*., 1984). This study has shown that the majority of HPA binding glycoproteins in T47D (with the exception of one glycoprotein of approximately 85kDa which interacted with the anti-A antibody), encompassed structures other than the blood group A epitope. Hirota *et al* (1992) previously reported a number of membrane proteins bearing the blood group A antigen in pancreatic cancer and these ranged in molecular weight from 20 to 200 kDa, of which, one glycoprotein at 85 kDa was amongst the most prominent glycoprotein (Hirota *et al*., 1992). Similarly, Saint-Guirons *et al* (2007) showed that membrane glycoproteins bearing the blood group A substance ranged in molecular weight from 20-82 kDa in the HT29 colorectal cell line, and of these, intriguingly, only one membrane glycoprotein at 82 kDa interacted with the lectin HPA. Furthermore, a recent study has shown an exhaustive number of glycoproteins in the sera of breast cancer patients (including a glycoprotein at 85 kDa), bound HPA, irrespective of the blood group of the patient from whom the sera was derived (Welinder *et al*., 2009). Taken together, these observations suggest that the presence of the blood group A determinant on GP85 of T47D cells is likely to be a tumour -associated phenomenon.

5.6.2 HPA binding to O-GlcNAcylated proteins

O-GlcNAcylation is a common form of glycosylation amongst cytoplasmic proteins and has been shown to be an alternative pathway to phosphorylation functioning in intracellular signalling processes (Slawson *et al*., 2008). Aberrant O-GlcNAcylation is a notable feature associated with chonic diseases such as diabetes (Copeland *et al*., 2008; Akimoto *et al*., 2005), cardiovascular disease (Laczy *et al*., 2009;

Jones *et al*, 2008), neurodegenerative disorders (Lazarus *et al*; 2009; Lefebvre *et al*., 2005) and cancer (Chou *et al*., 2001). In this study intense binding of an anti-O-GlcNAc antibody was noted in MCF-7 and T47D consistent with the metastatic phenotype of these cells. Immunofluorescence studies with an O-GlcNAc antibody and HPA showed that most of the binding co-localised in the cytoplasm, confirming the hypothesis that aberrant O-GlcNAcylation occurs mostly on cytoplasmic proteins. Beside these findings made at the cellular level, the 1-DE immunoblots showed three major protein species which interacted with the antibody at approximately 35 kDa, 50kDa and 70 kDa in the both the cytoplasmic and membrane protein fractions of T47D, all of which closely correspond to the molecular weight of the HPA binding glycoproteins. It is therefore reasonable to propose that the proteins at 35 kDa, 50 kDa and 70 kDa are likely to be Hsp27, HnRNP D-like or HnRNP A2/B1 and HnRNP H1, respectively. Furthermore inhibition studies showed that the interaction of HPA with the O-GlcNAcylated protein species could be abrogated with GlcNAc. Recent studies using click chemistry based tagging has confirmed the O-GlcNAcylation of ENO1, Hsp27 in MCF-7 (Gurcel *et al*., 2008), however, this is the first report to show the O-GlcNAc modification of the transcription factors HnRNP H1, HnRNP D-like and HnRNP A2/B1. Interestingly, the HPA binding glycoproteins identified in this study also exist as O-phosphorylated proteins (Mayrand et la., 1993; Huot *et al*, 1996; Karl *et al*2000; Trojanowicz *et al*., 2008). Bioinformatic analysis was undertaken to evaluate the potential O-GlcNAcylation and phosphorylation of the HPA binding glycoproteins identified in this study. This analysis confirmed that all the proteins have potential O-GlcNAc (≥4) and O-phosphate (≥12) modification sites (Appendix 5). For instance, ten potential O-GlcNAcylation and twelve potential O-phosphorylation sites respectively were predicted in Hsp27, ten of which occurred on, or near, the same Ser or Thr residue on the polypeptide chain (Appendix 5), indicating that indeed these proteins may exist both in an O-GlcNAcylated and O-phosphorylated form.

The 2DE immunoblot also showed other O-GlcNAcylated proteins that were not observed in the HPA blot. It is interesting to speculate as to whether these proteins are simply proteins that are in transition to an O-phosphorylation state. Studies have shown that the transition of proteins between O-GlcNAcylation and

phosphorylation states enables exquisite control to be exerted on cell regulatory mechanisms and that a disruption in this cross-talk may be implicated in many disease states (Bork *et al*., 2008; Copeland *et al*., 2008). It will be interesting to examine the whether the cross-talk between O-GlcNAcylation and O-phosphorylation bears any relation to the function of these proteins in cancer.

5.6.3 HPA binding in the Golgi apparatus

The initial observation that HPA labelled the perinuclear region of the metastatic cell lines MCF-7 and T47D (Chapter 3) led to co-localisation work using HPA-TRITC and a Golgi tracker dye. The observations highlighted that HPA recognises proteins resident in or in transit through, the Golgi apparatus, consistent with previous observations where HPA labelling has been localised in this cellular compartment in non-malignant cells from sources other than the breast (Laitinen *et al*., 1990; Roth, 1984), but this finding contradicts ultrastructural studies with breast cells lines which failed to report HPA localisation in the Golgi apparatus in human mammary tumours (Calafat & Janssen, 1984) and breast cancer cell lines (Mitchell *et al*., 1995).

Wang *et al*. (2001) suggested that disruptions in the glycosylation pathway in colorectal adenocarcinoma is manifested by an increase in the levels of STn in the trans-Golgi compartment, possibly corresponding to aberrant expression of enzymes such as sialyltransferases, acting later in the biosynthetic pathways.

It appears that the HPA labelling observed in MCF-7 and T47D in this study represents a mixture of immature GalNAc bearing glycan epitopes which may be consistent with increasing disruption in glycosylation pathways of enzymes such as ppGalNAc-Ts or ST6GalNAc (Brooks *et al.,* 2001; Cardone *et al.,* 2005; Hassan *et al.,*2000) leading to an increased concentration of immature N-acetylgalactosaminylated glycoforms in these cellular compartments.

5.6.4 Glycosyltransferases

To understand the basis of HPA binding in MCF-7 and T47D, an analysis of the gene expression of *of ppGalNAc-T1, T2, T3* and *T6* was undertaken. In the current investigation, however no correlation between mRNA levels of ppGalNAc-T and the HPA binding status of the cell lines was noted using the q-RT-PCR. It is possible that mRNA expression level of *ppGalNAc T1, T2, T3* and *T6* do not reflect the level of these enzyme in the cells, as has been proposed elsewhere where mRNA levels did not predict protein level (Chen *et al*., 2002; Gygi *et al*., 1999; Lidgren *et al*., 2008; Nie *et al*., 2006; Al-Mulla *et al*., 2005). From what is known about the relationship between protein levels and mRNA expression levels it appears plausible the mRNA expression level of *ppGalNAcTs* and *ST6GalNAcs* may not reflect the level of these enzymes in the cells.

Another factor to consider is the confluency of the starter culture and other aspects of cell culture which have shown to be important for the expression of cell surface glycoconjugates (Breen & Ronayne, 1994). Thus, the results observed in this study in terms of the mRNA analysis may reflect differences across laboratories with respect to experimental conditions.

To date, a family of fifteen of ppGalNAc Ts have been identified in mammalian cells and according to Ten Hagen *et al* (2003) as many as 20 ppGalNAc-Ts may exist. Out of the 11 ppGalNAc-Ts not studied in the current investigation, some of these enzymes might be responsible for the addition of the GalNAc residue to the proteins, as opposed to the four ppGalNAc-T studied in this investigation.

Another possibility considered was that increased levels of other enzymes acting later (such as core 1 transferase) might lead a failure in completion of the glycosylation process and leaving immature carbohydrate moieties exposed. ST6GalNac I transfers a sialic acid residue in alpha 2-6 linkage onto GalNAc-O-Ser/Thr (Pinho *et al*. 2007) to yield sialyl Tn. Failure of this enzyme to function may result in exposed GalNAc residues (Brooks *et al*., 2001). Althrough ST6GalNAc I

is the primary synthase of STn epitope, ST6GalNAc II has also been reported as a secondary candidate synthase for this epitope.

In an attempt to correlate a relationship between HPA labelling and ST6GalNAc I and ST6GalNAc II expression levels in breast cells, mRNA expression was assessed using q-RT-PCR. Julien *et al* (2001) previously reported that breast cancer cell lines (including the MCF-7 and T47D) cells used in this study express neither ST6GalNAc I nor sialyl-Tn antigen and indeed in this investigation no expression of ST6GalNAc I was observed, in any of the breast cell lines suggesting that STn is unlikely to be HPA binding candidate antigen in metastatic T47D or MCF-7 cells. Gene expression levels of ST6GalNAc II also failed to show correlation with the HPA binding status and the phenotypic characteristic of the cell lines, thus indicating that ST6GalNAc II is unlikely to be a candidate synthase for STn antigen in the breast cell lines. This may be consistent with the previous observations that ST6GalNAc I is the primary synthase of STn in cancer (Marcos *et al*., 2004).

Another possibility is that the substrate for the enzyme might be limited or as discussed previously, the enzyme may be mislocalised intracellularly. Studies have shown that changes in the localisation of transferases residing in the Golgi apparatus occurs in malignancy and this results in increased synthesis of cancer associated carbohydrate epitopes by disrupting the location or the optimal activity of several glycosylation enzymes (Kellokumpu *et al*., 2002; Rivinoja *et al*., 2006). Specifically it has been shown that inappropriate Golgi pH resulted in fragmented Golgi apparatus and lowered glycosylation potential of the Golgi apparatus in MCF-7 cells (Kellokumpu *et al*., 2002). It may be possible that the differences in glycosylation observed between normal and cancer cells is associated with an alteration in enzymatic activity due to changes in the Golgi apparatus pH.

The glycosyltransferase work is at relatively early stage and there is clearly a need for further investigations to validate and further explore these observations. Initial work with SNA and PNA suggested that sialic acid and TF containing epitopes are present in the MCF-7 and T47D cells as well as to a lesser degree in HMT3522

and BT474 cells. This is consistent with reports where increased levels of STn antigen detected by SNA were observed in several malignancies (Kuwabara *et al*., 1997; Leivonen *et al*., 2001; Pinho *et al*., 2007; Schuessler *et al*., 1991; Wang *et al*., 2001). These lectins as well as others, alongside mRNA analysis of a wider range of ppGalNAc-Ts and ST6GalNAc-Ts will facilitate further understanding of the changes in these cells as associated with HPA binding.

5.7 Conclusion

In summary, the HPA binding partners in the breast cancer cell lines were shown to encompass antigens other than blood group A. HPA staining in the perinuclear region of the cytoplasm of MCF-7 and T47D was localised in the Golgi apparatus and this was assumed to be by virtue of interaction with proteins in transit through the secretory pathways of the cell or with glycosyltransferases resident in this organelle. This study revealed that the cells with a metastatic phenotype contain O-GlcNAcylated proteins on a range of cytoplasmic proteins. For the first time HPA has been shown to recognise the HnRNPs, Hsp 27 and ENO1 and the evidence with the confocal microscopy work suggests that this was via O-GlcNAcylated residues. In an attempt to unravel the changes in the glycosylation machinery, as related to HPA binding, expression levels of *ppGalNAc-T1, T2, T3* and *T4* in T47D and MCF-7 were assessed, but in the current study, no correlation between the HPA binding status of the cells and the expression of the enzymes was found. An absence of ST6GalNAc I and varied expression of ST6GalNAc II in the breast cells sparks new interest in other ppGalNAc Ts not considered in this study and also other enzymes that are involved at later stages in the glycosylation process.

Chapter 6

General Discussion

6.0 General Discussion

6.1 Novelty of the research outcomes

The significance of HPA as a tool for identifying poor prognosis cancer has been shown in several retrospective studies of resected primary tumour tissues (Ikeda *et al*., 1994; Leathem & Brooks, 1987) and the role of HPA binding epitopes in the metastatic process has been established using cancer cell lines derived from human tumour tissues (Schumacher *et al*., 1994). However, the precise nature of the HPA-binding partners that are involved in cancer cell metastasis not yet been identified (Brooks, 2000; Lescar *et al*., 2007). The major aim of this study was to use proteomic technologies to explore the hypothesis that HPA recognises the same proteins in breast cancer as have previously been reported in CRC (Saint-Guirons *et al*., 2007). The proteomic approach that was adopted has enabled hundreds of proteins to be compared across a range of breast cancer cell lines and to be correlated with the HPA binding properties of the cells.

The cell lines used in this study have previously been assessed for their HPA binding status and the observations made previously were confirmed here (Brooks *et al*., 2001; Schumacher & Adam, 1997). A range of novel HPA binding proteins were identified as described in Chapter 4. These included molecules involved in apoptosis (Hsp 27) (Rane *et al.,* 2003), pre-mRNA splicing (HnRNP H1, HnRNP D-like, HnRNP A2/B1) (Gallinaro *et al.,* 1986), cellular remodelling (GFAP) (Kohama *et al.,* 1995) and cell migration (ENO1) (Wygrecka *et al.,* 2009).

One of the concerns surrounding the use of HPA in cancer-related studies is that the lectin has been shown to recognise the blood group A epitope (Prokop *et al*., 1965; Uhlenbruck, 1966). To address whether this is the case in the breast cancer cells, the HPA binding glycoproteins of the cells were compared with those recognised by an anti-blood group A antibody. From this work it was clear that HPA binds glycoproteins that contain structures distinct from the blood group A determinant. Further work has shown that HPA appears to bind to cancer cells

via recognition of O-GlcNAcylated epitopes (Chapter 5). HPA also bound proteins of similar molecular weight to the integrin α6 subunit, concurring with previous findings in CRC. The results from the present study show that, as in the CRC model, the integrin α6 subunit was detectable in the breast cell lines and appeared to be the most abundant HPA binding glycoprotein in the cells with a metastatic phenotype.

Many of the proteins that were identified in this work have previously been described to be altered in cancer (Arlt *et al*., 2009; Bhui-Kaur *et al*., 1998; Chista *et al*., 1994; Coghlin *et al*., 2006; Shiraishi *et al*., 1992; Steeg *et al*., 1993; Storm *et al*., 1996; Tsai *et al*., 2010; Zhang *et al*., 2008; Zhou *et al*., 2001). The proteomic study also identified elevated levels of the elongation factor Tu; enoyl coenzyme A hydratase 1 peroxisomal; and macropain subunits (Chapter 4), none of these proteins have previously been reported as altered in cancer and therefore offer potential for future studies.

HPA labelling of the cancer cells was shown to partly localise to the Golgi apparatus. This is consistent with previous observations (Brooks *et al*., 2001; Laitinen *et al*., 1990; Roth, 1984) and supports the hypothesis that changes in the glycosylation pathways and location of enzymes, such as ppGalNAc-Ts ST6GalNAcs occur in cancer.

6.2 Glycoproteins as biomarkers and targets for cancer treatments

Most tumour markers that are used in the clinical setting are glycoproteins. For example: CA125 (Bast *et al*., 2005); CA15.3 (Schmidt-Rhode *et al*., 1987); carcinoembryonic antigen, CEA, (Carmignani *et al*., 2004); HER2 (Konecny *et al*., 2004) and PSA (Stenman *et al*., 2005). Some of these proteins have been used to monitor the clinical progression of cancer (PSA) whilst others are used to help define treatment regimens (HER2). Whilst many of these markers play an important role, they are of limited use as diagnostic tools, principally because of their low specificity and sensitivity, two factors important in a tumour marker

(Pannall & Kotasek., 1997). Therefore, there remains an opportunity to identify new biomarkers for the diagnosis of breast (and other) cancers and these may also offer utility for predicting response to treatment or for prognostication.

Lectins have frequently been used for identify glycosylation changes in cancer (Baldus & Hanisch, 2000; Blonski *et al*., 2007; Handerson *et al*., 2005; Takano *et al*., 2000), and there has been considerable interest in identifying the epitopes recognised by the lectin HPA. As described above, one of the key findings of this investigation was that HPA bound to a protein of the cancer cells that correlated with the migration position of the integrin α6 subunit on SDS-PAGE. This finding concurs with observations previously made in CRC cells (Saint-Guiros *et al*., 2007) and the work of Prokopishyn *et al*(1999) who have proposed that the integrins are the major carrier of oncodevelopmental carbohydrates in CRC. Taken together, these observations support the hypothesis that members of the integrin family, which exhibit abnormal glycosylation, may play a significant role in the metastatic process, for example by facilitating cancer cell invasion through the ECM (Chao *et al*., 1996).

To extend the findings of this study it would be of value to assess whether the integrin α6 subunit is detectable in the serum of breast cancer patients, as the glycosylation of this protein appears to be altered in a cancer-specific manner and this may, therefore, be a useful tumour marker. An alternative approach, given the cell surface localisation of the integrin α6 subunit, may be to evaluate this protein as a potential candidate for biological based approaches for cancer targeting. A similar strategy was adopted for targeting cErb-B2/HER2 in breast cancer. Clearly, many more samples from different cell lines and human cancer samples will require detailed peptide mapping and functional analysis in an initial step toward testing this hypothesis.

6.3 HPA recognition of O-GlcNAcylated proteins

One of the major novel findings of this work, supported by evidence from confocal microscopy and Western blot analysis, is that HPA recognises ENO1, HnRNPs

and Hsp 27 and that this recognition appears to be via O-GlcNAcylation of the proteins. Bioinformatic analysis has shown the likely addition of phosphate and O-GlcNAc groups to these proteins. ENO1 and Hsp27 have been shown to be modified with O-GlcNAc in previous studies in breast cancer (Gurcel *et al*., 2008). This is the first observation of the modification of the transcription factors HnRNP H1, HnRNP D-like and HnRNP A2/B1 with O-GlcNAc, although elevated levels of these transcription factors has previously been shown in cancer (Zhang *et al*., 2008, Zhou *et al*., 2001). The HPA binding glycoproteins identified as being modified with O-GlcNAc residues have also been shown to exist as O-phosphorylated proteins (Mayrand *et al.,* 1993; Huot *et al*, 1996; Trojanowicz *et al*., 2008).

The O-GlcNAc modification has been shown to play a role in regulating protein stability (Bode *et al*., 2004), protein-protein interactions (Golks *et al*., 2007) and in modifying the DNA binding ability of a number of transcription factors including c-myc, p53, Pdx-1, FoxO-1, NF-κB STAT5 (Gu *et al*., 1994) by competing with protein kinases and modulating protein phosphorylation of the same Ser/Thr residues (Slawson *et al*., 2008).

Modification of proteins with O-GlcNAc also regulates their sub-cellular localisation, particularly the shuttling of transcription factors between the nucleus and cytoplasm (Lefterova *et al*., 1994; Petersen *et al*., 2007; Kawamura *et al*., 2002). It is unclear as to the precise cellular localisation of the HPA binding glycoproteins found in this study, whether these exist as cytosolic, nuclear or cell membrane resident molecules and if the intracellular localisation is influenced by their O-GlcNAcylation and O-phosphorylation status. There is an opportunity to examine the potential cross-talk between O-GlcNAcylation and O-phosphorylation for these proteins using this well-characterised *in vitro* model and to determine if the O-GlcNAc modification of the HnRNPs influences intracellular signalling pathways and DNA transcription.

6.3.1 Inhibitors of O-GlcNAcylation as cancer therapies

The O-GlcNAc modification is catalysed by the enzyme O-linked N-acetylglucosaminyl transferase (OGT) which functions to transfer O-GlcNAc to Ser/Thr residues on the protein backbone; β-N-acetylglucosaminidase (O-GlcNAcase or OGA) removes O-GlcNAc residues. There has been considerable interest over the past few years in understanding how these two enzymes control the O-GlcNAcylation of the hundreds of individual protein substrates found intracellularly and how this system regulates cellular processes. A monoclonal antibody raised against the O-GlcNAc-Ser epitope (CTD110.6) has been shown to block the addition of free GlcNAc to Ser/Thr and appears to be of value for the study of the role of O-GlcNAc in cell physiology (Comer *et al*., 2001).

The crystal structures of OGT (Lazarus *et al*., 2011) and OGA (Dennis *et al*., 2006) have been solved and this has provided an insight into the substrate recognition and the mechanism by which the enzymes function. Using the data from the crystal structure alongside molecular modelling (for example with the Autodock system) will enable large numbers of inhibitor molecules to be screened from databases of ligand molecules (Heindl *et al*., 2011). The relative affinity of ligands to the enzyme can be determined and promising candidate molecules are then subject to detailed laboratory analyses (Dorfmueller *et al*., 2011, Haltiwanger *et al*., 1998). To date, two OGA inhibitors have been explored in the laboratory (PUGNAc and streptozotocin) by raising cellular O-GlcNAc levels and hence favouring O-GlcNAc modification (Haltiwanger *et al*., 1998). UDP-GlcNAc analogues such as alloxan (a glucose analogue) have also been explored as OGT inhibitors and are in the pre-clinical testing phase (Dorfmueller *et al*., 2011). In summary, the development of inhibitors will enable studies of the cellular functions mediated by the OGT and OGA enzymes. Inhibitors of the O-GlcNAcylation process may later find clinical application as anti-cancer drugs, potentially offering scope alongside the protein kinase inhibitors such as tyrosine kinases inhibitors.

6.4 Conclusion

The results presented in this thesis provide clear evidence of the utility of HPA as a tool for identification of proteins showing changes in glycosylation in a model of metastatic breast cancer. The proteomic approach allowed the characterisation of HPA binding glycoproteins and non-HPA binding proteins which were differentially glycosylated in the cells with a metastatic phenotype. The work has shown that integrin α6 is an HPA binding protein that is recognised in both CRC and breast cancer cells. HPA also recognises O-GlcNAcylated cytoplasmic proteins (ENO1, HnRNP H1, HnRNP A2/B1, HnRNP D-like). The O-GlcNAcylation of the HnRNPs transcription factors opens up the possibility of glycosylation affecting the regulation and, therefore, function of these proteins suggesting a new mechanism by which HPA may detect poor prognosis cancers. New approaches aimed at targeting these changes might have broad application for the treatment of breast, colorectal and possibly other epithelial cancers.

6.5 Future work

There is an opportunity to investigate the following:

- The HPA binding component described as GP80 remained poorly characterised. A strategy for this may include HPA affinity chromatography followed by LC-MS/MS to identify the HPA binding partners.

- It will be helpful to investigate the protein levels of ppGAlNAc-Ts and ST6GalNAcs enzymes rather than the mRNA levels in the breast cells to validate the observations made in this investigation.

- The specificity of HPA binding to proteins of T47D was assessed by competitive inhibition using GalNAc and GlcNAc but this may be further investigated using cancer cell derived glyconjugates and HPA in conjunction with Surface Plasmon Resonance measurements (BiaCore System).

- It would be of interest to compare the protein profile for O-GlcNAcylation of the metastatic and non-metastatic cell lines using‘click chemistry’ based tagging approach. It will also be interesting to study the inter-relationship between O-GlcNAcylation and O-phosphorylation in the function of the HnRNPs in pre-mRNA splicing.

References

References

Abiatari, I., Esposito, I., Oliveira, T.D., Felix, K., Xin, H., Penzel, R., Giese, T., Friess, H. & Kleeff, J. (2010). Moesin-dependent cytoskeleton remodelling is associated with an anaplastic phenotype of pancreatic cancer. *J Cell Mol Med*, **14**, 1166-79.

Aebi, M. & Hennet, T. (2001). Congenital disorders of glycosylation: genetic model systems lead the way. *Trends Cell Biol*, **11,** 136-41.

Afrough, B. (2008). Characterisation of O-glycans from HOC544 and the challenges in targeting pre-defined complex carbohydrate conformation of mucin type glycoprotein. PhD thesis, University of Westminster, London.

Akimoto, Y., Comer, F.I., Cole, R.N., Kudo, A., Kawakami, H., Hirano, H. & Hart, G.W. (2003). Localization of the O-GlcNAc transferase and O-GlcNAc-modified proteins in rat cerebellar cortex. *Brain Res*, **966,** 194-205.

Akimoto, Y., Hart, G.W., Hirano, H. & Kawakami, H. (2005). O-GlcNAc modification of nucleocytoplasmic proteins and diabetes. *Med Mol Morphol*, **38,** 84-91.

Al-Mulla, F., Abdulrahman, M., Varadharaj, G., Akhter, N. & Anim, J.T. (2005). BRCA1 gene expression in breast cancer: a correlative study between real-time RT-PCR and immunohistochemistry. *J Histochem Cytochem*, **53**, 621-9.

Alam, S.M., Whitford, P., Cushley, W., George, W.D. & Campbell, A.M. (1990). Flow cytometric analysis of cell surface carbohydrates in metastatic human breast cancer. *Br J Cancer*, **62**, 238-42.

Altevogt, P., Fogel, M., Cheingsong-Popov, R., Dennis, J., Robinson, P. & Schirrmacher, V. (1983). Different patterns of lectin binding and cell surface sialylation detected on related high- and low-metastatic tumor lines. *Cancer Res*, **43**, 5138-44.

Anderson, D.E. & Haas, C. (1984). Blood type A and familial breast cancer. *Cance*r, **54**, 1845-9.

Andrieu, C., Taieb, D., Baylot, V., Ettinger, S., Soubeyran, P., De-Thonel, A., Nelson, C., Garrido, C., So, A., Fazli, L., Bladou, F., Gleave, M., Iovanna, J.L. & Rocchi, P. (2010). Heat shock protein 27 confers resistance to androgen ablation and chemotherapy in prostate cancer cells through eIF4E. *Oncogene*, **29**, 1883-96.

Apweiler, R., Hermjakob, H. & Sharon, N. (1999). On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta*, **1473,** 4-8.

Arlt, A., Bauer, I., Schafmayer, C., Tepel, J., Muerkoster, S.S., Brosch, M., Roder, C., Kalthoff, H., Hampe, J., Moyer, M.P., Folsch, U.R. & Schafer, H. (2009). Increased proteasome subunit protein expression and proteasome activity in colon cancer relate to an enhanced activation of nuclear factor E2-related factor 2 (Nrf2). *Oncogene*, **28,** 3983-96.

Asada, M., Furukawa, K., Segawa, K., Endo, T. & Kobata, A. (1997). Increased expression of highly branched N-glycans at cell surface is correlated with the malignant phenotypes of mouse tumor cells. *Cancer Res*, **57**, 1073-80.

Aub, J.C., Sanford, B.H. & Cote, M.N. (1965). Studies on reactivity of tumor and normal cells to a wheat germ agglutinin. *Proc Natl Acad Sci U S A*, **54**, 396-9.

Avigdor, A., Goichberg, P., Shivtiel, S., Dar, A., Peled, A., Samira, S., Kollet, O., Hershkoviz, R., Alon, R., Hardan, I., Ben-Hur, H., Naor, D., Nagler, A. & Lapidot, T. (2004). CD44 and hyaluronic acid cooperate with SDF-1 in the trafficking of human CD34+ stem/progenitor cells to bone marrow. *Blood*, **103**, 2981-9.

Baker, D.A., Sugii, S., Kabat, E.A., Ratcliffe, R.M., Hermentin, P. & Lemieux, R.U. (1983). Immunochemical studies on the combining sites of Forssman hapten reactive hemagglutinins from Dolichos biflorus, Helix pomatia, and Wistaria floribunda. *Biochemistry*, **22,** 2741-50.

Baldus, S., Park, Y., Kotlarek, G., Hell, K. & Fischer, R. (1996). Expression of Gal beta 1-4GlcNAc sequences by human gastrointestinal neoplasms and their precursors as detected by Erythrina cristagalli and Erythrina corallodendron lectins. *Int J Oncol*, **9**, 43-8.

Baldus, S.E. & Hanisch, F.G. (2000). Biochemistry and pathological importance of mucin-associated antigens in gastrointestinal neoplasia. *Adv Cancer Res*, **79,** 201-48.

Bannister, L. H., Berry, M. M., Collins, P., Dyson, M. & Dussek, J. E. (1995). *Gray's Anatomy*, 38. Churchill Livingstone. New York, USA, 417-424.

Barnes, Y.C., Skelton, T.P., Stamenkovic, I. & Sgroi, D.C. (1999). Sialylation of the sialic acid binding lectin sialoadhesin regulates its ability to mediate cell adhesion. *Blood*, **93,** 1245-52.

Barondes, S.H., Cooper, D.N., Gitt, M.A. & Leffler, H. (1994). Galectins. Structure and function of a large family of animal lectins. *J Biol Chem*, **269,** 20807-10.

Bartolazzi, A., Nocks, A., Aruffo, A., Spring, F. & Stamenkovic, I. (1996). Glycosylation of CD44 is implicated in CD44-mediated cell adhesion to hyaluronan. *J Cell Biol*, **132,** 1199-208.

Bast, R.C., Jr., Badgwell, D., Lu, Z., Marquez, R., Rosen, D., Liu, J., Baggerly, K.A., Atkinson, E.N., Skates, S., Zhang, Z., Lokshin, A., Menon, U., Jacobs, I. & Lu,

K. (2005). New tumor markers: CA125 and beyond. *Int J Gynecol Cancer*, **15 Suppl 3,** 274-81.

Baum, M. & Schipper, H. (2005). *Breast Cancer*. 3rd Edition. Health Press Ltd. Oxford.

Beckmann, M.W., Niederacher, D., Schnurch, H.G., Gusterson, B.A. & Bender, H.G. (1997). Multistep carcinogenesis of breast cancer and tumour heterogeneity. *J Mol Med*, **75,** 429-39.

Bedard, K., Szabo, E., Michalak, M. & Opas, M. (2005). Cellular functions of endoplasmic reticulum chaperones calreticulin, calnexin, and ERp57. *Int Rev Cytol*, **245,** 91-121.

Bennett, E.P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M.A., Merkx, G., van Kessel, A.G., Eiberg, H., Steffensen, R. & Clausen, H. (1998). Cloning of a human UDP-N-acetyl-alpha-D-Galactosamine:polypeptide N-acetylgalactosaminyltransferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. *J Biol Chem*, **273**, 30472-81.

Berinstein, N.L. (2002). Carcinoembryonic antigen as a target for therapeutic anticancer vaccines: a review. *J Clin Oncol*, **20,** 2197-207.

Berois, N., Mazal, D., Ubillos, L., Trajtenberg, F., Nicolas, A., Sastre-Garau, X., Magdelenat, H. & Osinaga, E. (2006). UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-6 as a new immunohistochemical breast cancer marker. *J Histochem Cytochem*, **54,** 317-28.

Berry, N., Jones, D.B., Smallwood, J., Taylor, I., Kirkham, N. & Taylor-Papadimitriou, J. (1985). The prognostic value of the monoclonal antibodies HMFG1 and HMFG2 in breast cancer. *Br J Cancer*, **51,** 179-86.

Bhui-Kaur, A., Therwath, A., Henry, L., Chiesa, J., Kurkure, A., Scherrer, K. & Bureau, J.P. (1998). Increased prosomal proteins in breast cancer cells and in neighboring normal cells in Parsi and non-Parsi populations. *J Cancer Res Clin Oncol*, 124, 117-26.

Bickel, T., Lehle, L., Schwarz, M., Aebi, M. & Jakob, C.A. (2005). Biosynthesis of lipid-linked oligosaccharides in Saccharomyces cerevisiae: Alg13p and Alg14p form a complex required for the formation of GlcNAc(2)-PP-dolichol. *J Biol Chem*, **280,** 34500-6.

Bies, C., Lehr, C.M. & Woodley, J.F. (2004). Lectin-mediated drug targeting: history and applications. *Adv Drug Deliv Rev*, **56,** 425-35.

Blass, S.L., Pure, E. & Hunter, C.A. (2001). A role for CD44 in the production of IFN-gamma and immunopathology during infection with Toxoplasma gondii. *J Immunol*, **166,** 5726-32.

Bloom, H.J.G & Richardson, W.W (1957). Histological Grading and Prognosis in Breast cancer: A study of 1409 cases of which 359 have been followed for 15 years. *British Journal of cancer.* **11**, 359-377.

Blonski, K., Milde-Langosch, K., Bamberger, A.M., Osterholz, T., Utler, C., Berger, J., Loning, T. & Schumacher, U. (2007). Ulex europeus agglutinin-I binding as a potential prognostic marker in ovarian cancer. *Anticancer Res*, **27,** 2785-90.

Bloom, H., and Richardson, W. (1957). Histological grading and prognosis in breast cancer. *Brit. J. Cancer*, **11**, 359-377.

Bode, A.M. & Dong, Z. (2004). Post-translational modification of p53 in tumorigenesis. *Nat Rev Cancer*, **4,** 793-805.

Bodian, C.A. (1993). Benign breast diseases, carcinoma in situ, and breast cancer risk. *Epidemiol Rev*, **15,** 177-87.

Bogenrieder, T. & Herlyn, M. (2003). Axis of evil: molecular mechanisms of cancer metastasis. *Oncogene*, **22,** 6524-36.

Bork, K., Kannicht, C., Nohing, S., Reutter, W., Weidemann, W., Hart, G.W. & Horstkorte, R. (2008). N-propanoylmannosamine interferes with O-GlcNAc modification of the tyrosine 3-monooxygenase and stimulates dopamine secretion. *J Neurosci Res*, **86**, 647-52.

Boyd, W.C. & Shapleigh, E. (1954). Separation of individuals of any blood group into secretors and non-secretors by use of a plant agglutinin (lectin). *Blood*, **9,** 1194-8

Breen, K.C. & Ronayne, E. (1994). The effect of cell confluency state on the expression of neural cell surface glycoconjugates. *Neuroreport*, 5, 970-2.

Briand, P., Petersen, O.W. & Van Deurs, B. (1987). A new diploid nontumorigenic human breast epithelial cell line isolated and propagated in chemically defined medium. *In Vitro Cell Dev Biol*, **23,** 181-8.

Brieher, W.M., Yap, A.S. & Gumbiner, B.M. (1996). Lateral dimerization is required for the homophilic binding activity of C-cadherin. *J Cell Biol*, **135,** 487-96.

Brinkley, D. and Haybittle, J. L. (1975). The curability of cancer. *Lancet* ii, 95-97.

Brockhausen, I. (1999). Pathways of O-glycan biosynthesis in cancer cells. *Biochim Biophys Acta*, **1473,** 67-95.

Brockhausen, I. (2006 a). Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions. *EMBO Rep*, **7,** 599-604.

Brockhausen, I. (2006 b). The role of galactosyltransferases in cell surface functions and in the immune system. *Drug News Perspect*, **19**, 401-9.

Brooks, S.A. (2000). The involvement of Helix pomatia lectin (HPA) binding N-acetylgalactosamine glycans in cancer progression. *Histol Histopathol*, **15,** 143-58.

Brooks, S.A., Carter, T.M., Royle, L., Harvey, D.J., Fry, S.A., Kinch, C., Dwek, R.A. & Rudd, P.M. (2008). Altered glycosylation of proteins in cancer: what is the potential for new anti-tumour strategies. *Anticancer Agents Med Chem*, **8,** 2-21.

Brooks, S.A., Carter, T.M., Bennett, E.P., Clausen, H. & Mandel, U. (2007). Immunolocalisation of members of the polypeptide N-acetylgalactosaminyl transferase (ppGalNAc-T) family is consistent with biologically relevant altered cell surface glycosylation in breast cancer. *Acta Histochem*, **109**, 273-84.

Brooks, S. A., Dwek, M. V. and Schumacher, U. (2002). *Functional and Molecular Glycobiology*. Bios Scientific Publishers Ltd. Oxford.

Brooks, S.A. & Hall, D.M. (2002). Investigations into the potential role of aberrant N-acetylgalactosamine glycans in tumour cell interactions with basement membrane components. *Clin Exp Metastasis*, **19,** 487-93.

Brooks, S.A., Hall, D.M.S. and Buley, I. (2001). GalNAc glycoprotein expression by breast cell lines, primary breast cancer and normal breast epithelial membrane. *British Journal of Cancer*, vol 85: 1014-1022.

Brooks, S.A. & Leathem, A.J. (1991). Prediction of lymph node involvement in breast cancer by detection of altered glycosylation in the primary tumour. *Lancet*, **338,** 71-4.

Brooks, S.A. & Leathem, A.J. (1995). Expression of alpha-GalNAc glycoproteins by breast cancers. *Br J Cancer*, **71,** 1033-8.

Brooks, S.A. & Leathem, A.J. (1998). Expression of N-acetyl galactosaminylated and sialylated glycans by metastases arising from primary breast cancer. *Invasion Metastasis*, **18,** 115-21.

Brooks, S.A., Lymboura, M., Schumacher, U. & Leathem, A.J. (1996). Histochemistry to detect Helix pomatia lectin binding in breast cancer: methodology makes a difference. *J Histochem Cytochem*, **44,** 519-24.

Brower, S. T., Cunningham J. D., Tartter P. I. Staging and assessing prognosis. *Breast cancer*. Ed/s: Roses, D. F. Elsevier Churchill Livingstone. Philadelphia.

Buko, A.M., Kentzer, E.J., Petros, A., Menon, G., Zuiderweg, E.R. & Sarin, V.K. (1991). Characterization of a posttranslational fucosylation in the growth factor domain of urinary plasminogen activator. *Proc Natl Acad Sci U S A*, **88,** 3992-6.

Burchell, J., Durbin, H. & Taylor-Papadimitriou, J. (1983). Complexity of expression of antigenic determinants, recognized by monoclonal antibodies HMFG-1 and

HMFG-2, in normal and malignant human mammary epithelial cells. *J. Immunol*, **131,** 508-13.

Burchell, J.M., Mungul, A. & Taylor-Papadimitriou, J. (2001). O-linked glycosylation in the mammary gland: changes that occur during malignancy. *J Mammary Gland Biol Neoplasia*, **6,** 355-64.

Burda, P. & Aebi, M. (1999). The dolichol pathway of N-linked glycosylation. *Biochim Biophys Acta*, **1426,** 239-57.

Burger, M.M. & Goldberg, A.R. (1967). Identification of a tumor-specific determinant on neoplastic cell surfaces. *Proc Natl Acad Sci U S A*, **57**, 359-66.

Calafat, J. & Janssen, H. (1984). Binding of lectins to human mammary tumors: ultrastructural study. *Breast Cancer Res Treat*, **4**, 169-79.

Campbell, B.J., Finnie, I.A., Hounsell, E.F. & Rhodes, J.M. (1995). Direct demonstration of increased expression of Thomsen-Friedenreich (TF) antigen in colonic adenocarcinoma and ulcerative colitis mucin and its concealment in normal mucin. *J Clin Invest*, **95**, 571-6.

Cancer Research UK (2008). Cancer Stats. Key Facts on Breast Cancer. http://info.cancerresearchuk.org/cancerstats/keyfacts/breast-cancer/

Cardone, R.A., Casavola, V. & Reshkin, S.J. (2005). The role of disturbed pH dynamics and the Na+/H+ exchanger in metastasis. *Nat Rev Cancer*, **5**, 786-95.

Carmignani, C.P., Hampton, R., Sugarbaker, C.E., Chang, D. & Sugarbaker, P.H. (2004). Utility of CEA and CA 19-9 tumor markers in diagnosis and prognostic assessment of mucinous epithelial cancers of the appendix. *J Surg Oncol*, **87,** 162-6.

Carter, W.G. & Wayner, E.A. (1988). Characterization of the class III collagen receptor, a phosphorylated, transmembrane glycoprotein expressed in nucleated human cells. *J Biol Chem*, **263,** 4193-201.

Cavallaro, U. & Christofori, G. (2004). Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nat Rev Cancer*, **4,** 118-32.

Chambers, A.F., Groom, A.C. & MacDonald, I.C. (2002). Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer*, **2,** 563-72.

Chantret, I., Dancourt, J., Barbat, A. & Moore, S.E. (2005). Two proteins homologous to the N- and C-terminal domains of the bacterial glycosyltransferase Murg are required for the second step of dolichyl-linked oligosaccharide synthesis in Saccharomyces cerevisiae. *J Biol Chem*, **280,** 9236-42.

Chao, C., Lotz, M.M., Clarke, A.C. & Mercurio, A.M. (1996). A function for the integrin alpha6beta4 in the invasive properties of colorectal carcinoma cells. *Cancer Res*, **56**, 4811-9.

Chen, E., Hermanson, S. & Ekker, S.C. (2004). Syndecan-2 is essential for angiogenic sprouting during zebrafish development. *Blood*, **103,** 1710-9.

Chen, G., Gharib, T.G., Huang, C.C., Taylor, J.M., Misek, D.E., Kardia, S.L., Giordano, T.J., Iannettoni, M.D., Orringer, M.B., Hanash, S.M. & Beer, D.G. (2002). Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics*, **1**, 304-13.

Chen, X., Jiang, J., Yang, J., Chen, C., Sun, M., Wei, Y., Guang, X. & Gu, J. (2006). Down-regulation of the expression of beta1,4-galactosyltransferase V promotes integrin beta1 maturation. *Biochem Biophys Res Commun*, **343,** 910-6.

Cheresh, D.A., Reisfeld, R.A. & Varki, A.P. (1984). O-acetylation of disialoganglioside GD3 by human melanoma cells creates a unique antigenic determinant. *Science*, **225,** 844-6.

Cheung, K.L. (2007). Endocrine therapy for breast cancer: An overview. The breast, **16,** 327-243.

Chou, T.Y. & Hart, G.W. (2001). O-linked N-acetylglucosamine and cancer: messages from the glycosylation of c-Myc. *Adv Exp Med Biol*, **491,** 413-8.

Christa, L., Simon, M.T., Flinois, J.P., Gebhardt, R., Brechot, C. & Lasserre, C. (1994). Over-expression of glutamine synthetase in human primary liver cancer. *Gastroenterology*, **106**, 1312-20.

Cichy, J., Bals, R., Potempa, J., Mani, A. & Pure, E. (2002). Proteinase-mediated release of epithelial cell-associated CD44. Extracellular CD44 complexes with components of cellular matrices. *J Biol Chem*, **277,** 44440-7.

Clark, G.M. & McGuire, W.L. (1988). Steroid receptors and other prognostic factors in primary breast cancer. *Semin Oncol*, **15,** 20-5.

Clausen, H. & Bennett, E.P. (1996). A family of UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferases control the initiation of mucin-type O-linked glycosylation. *Glycobiology*, **6,** 635-46.

Coghlin, C., Carpenter, B., Dundas, S.R., Lawrie, L.C., Telfer, C. & Murray, G.I. (2006). Characterization and over-expression of chaperonin t-complex proteins in colorectal cancer. *J Pathol*, **210**, 351-7.

Colognato, H. & Yurchenco, P.D. (2000). Form and function: the laminin family of heterotrimers. *Dev Dyn*, **218,** 213-34.

Comer, F.I., Vosseller, K., Wells, L., Accavitti, M.A. & Hart, G.W. (2001). Characterization of a mouse monoclonal antibody specific for O-linked N-acetylglucosamine. *Anal Biochem*, **293,** 169-77.

Condeelis, J. & Pollard, J.W. (2006). Macrophages: obligate partners for tumor cell migration, invasion, and metastasis. *Cell*, **124,** 263-6.

Cooper, A. P. (1840). *The Anatomy of the Breast.* Longman, Orme, Green, Brown and Longmans. London.

Copeland, R.J., Bullen, J.W. & Hart, G.W. (2008). Cross-talk between GlcNAcylation and phosphorylation: roles in insulin resistance and glucose toxicity. *Am J Physiol Endocrinol Metab*, **295**, E17-28.

Couldrey, C. & Green, J.E. (2000). Metastases: the glycan connection. *Breast Cancer Res*, **2,** 321-3.

Crocker, P.R., Paulson, J.C. & Varki, A. (2007). Siglecs and their roles in the immune system. *Nat Rev Immunol*, **7,** 255-66.

Cunha, G.R. (1994). Role of mesenchymal-epithelial interactions in normal and abnormal development of the mammary gland and prostate. *Cancer*, **74,** 1030-44.

Cunningham, L. (1977). The anatomy of the arteries and veins of the breast. *J Surg Oncol*, **9,** 71-85.

Dasgupta, A., Takahashi, K., Cutler, M. & Tanabe, K.K. (1996). O-linked glycosylation modifies CD44 adhesion to hyaluronate in colon carcinoma cells. *Biochem Biophys Res Commun*, **227,** 110-7.

Dawson EK. (1934). A histological study of the normal mamma in relation to tumour growth: I. Early development to maturity. *Edinburgh Med J* **41**, 653-682.

Dedhar, S., Saulnier, R., Nagle, R. & Overall, C.M. (1993). Specific alterations in the expression of alpha 3 beta 1 and alpha 6 beta 4 integrins in highly invasive and metastatic variants of human prostate carcinoma cells selected by in vitro invasion through reconstituted basement membrane. *Clin Exp Metastasis*, **11,** 391-400.

De Graffenried, C.L. & Bertozzi, C.R. (2004). The roles of enzyme localisation and complex formation in glycan assembly within the Golgi apparatus. *Curr Opin Cell Biol*, **16,** 356-63.

DeGrendele, H.C., Estess, P. & Siegelman, M.H. (1997). Requirement for CD44 in activated T cell extravasation into an inflammatory site. *Science*, **278,** 672-5.

Denda-Nagai, K. & Irimura, T. (2000). MUC1 in carcinoma-host interactions. *Glycoconj J*, **17,** 649-58.

Dennis, J.W. (1992). Changes in glycosylation associated with malignant transformation and tumor progression. In Cell Surface and Cell Development. Ed. Fukuda, M. CRC Press. 161–194.

Dennis, J.W., Granovsky, M. & Warren, C.E. (1999). Glycoprotein glycosylation and cancer progression. *Biochim Biophys Acta*, **1473,** 21-34.

Dennis, J.W., Laferte, S., Waghorne, C., Breitman, M.L. & Kerbel, R.S. (1987). Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis. *Science*, **236,** 582-5.

Dennis, R.J., Taylor, E.J., Macauley, M.S., Stubbs, K.A., Turkenburg, J.P., Hart, S.J., Black, G.N., Vocadlo, D.J. & Davies, G.J. (2006). Structure and mechanism of a bacterial beta-glucosaminidase having O-GlcNAcase activity. *Nat Struct Mol Biol*, **13,** 365-71.

Donohue, T.M., Osna, N.A. & Clemens, D.L. (2006). Recombinant Hep G2 cells that express alcohol dehydrogenase and cytochrome P450 2E1 as a model of ethanol-elicited cytotoxicity. *Int J Biochem Cell Biol*, **38,** 92-101.

Dorfmueller, H.C., Borodkin, V.S., Blair, D.E., Pathak, S., Navratilova, I. & van Aalten, D.M. Substrate and product analogues as human O-GlcNAc transferase inhibitors. *Amino Acids*, **40,** 781-92.

Douek, M. & Hall-Craggs, M.A. (1999). Can the use of preoperative MR imaging reduce local recurrence rates in patients with retroareolar breast cancer who undergo breast-conservation surgery? *Radiology*, **210,** 880-1.

Dove, A. (2001). The bittersweet promise of glycobiology. *Nat Biotechnol*, **19,** 913-7.

Dube, D.H. & Bertozzi, C.R. (2005). Glycans in cancer and inflammation--potential for therapeutics and diagnostics. *Nat Rev Drug Discov*, **4,** 477-88.

Duenas-Decamp, M.J., Peters, P., Burton, D. & Clapham, P.R. (2008). Natural resistance of human immunodeficiency virus type 1 to the CD4bs antibody b12 conferred by a glycan and an arginine residue close to the CD4 binding loop. *J Virol*, **82,** 5807-14.

Duffy, M. J. (2006). Serum Tumor Markers in Breast Cancer: Are They of Clinical Value? *Clinical Chemistry*, **52**, 345-351.

Dupont, W.D., Parl, F.F., Hartmann, W.H., Brinton, L.A., Winfield, A.C., Worrell, J.A., Schuyler, P.A. & Plummer, W.D. (1993). Breast cancer risk associated with proliferative breast disease and atypical hyperplasia. *Cancer*, **71,** 1258-65.

Dwek, M.V. (1998). Mapping of oligosaccharides of aggressive breast cancer. PhD thesis, University of London, London.

Dwek, M.V. & Brooks, S.A. (2004). Harnessing changes in cellular glycosylation in new cancer treatment strategies. *Curr Cancer Drug Targets*, **4,** 425-42.

Dwek, M.V., Ross, H.A., Streets, A.J., Brooks, S.A., Adam, E., Titcomb, A., Woodside, J.V., Schumacher, U. & Leathem, A.J. (2001). Helix pomatia agglutinin lectin-binding oligosaccharides of aggressive breast cancer. *Int J Cancer*, **95,** 79-85.

Dwek, R.A. (1995). Glycobiology: "towards understanding the function of sugars". *Biochem Soc Trans*, 23, 1-25.

Ehnis, T., Dieterich, W., Bauer, M., Lampe, B. & Schuppan, D. (1996). A chondroitin/dermatan sulfate form of CD44 is a receptor for collagen XIV (undulin). *Exp Cell Res*, **229,** 388-97.

Elledge,R.M and Osborne, C.K. (1997). Oestrogen receptors and breast cancer: It is time for indivisualised treatment based on oestrogen receptor status. *British Medical Journal,* **314,** 1843-1844.

Ellis, I.O., Bell, J., Todd, J.M., Williams, M., Dowle, C., Robins, A.R., Elston, C.W., Blamey, R.W. & Baldwin, R.W. (1987). Evaluation of immunoreactivity with monoclonal antibody NCRC 11 in breast carcinoma. *Br J Cancer*, **56,** 295-9.

Ellis, I.O., Hinton, C.P., MacNay, J., Elston, C.W., Robins, A., Owainati, A.A., Blamey, R.W., Baldwin, R.W. & Ferry, B. (1985). Immunocytochemical staining of breast carcinoma with the monoclonal antibody NCRC 11: a new prognostic indicator. *Br Med J (Clin Res Ed)*, **290,** 881-3.

Elola, M.T., Capurro, M.I., Barrio, M.M., Coombs, P.J., Taylor, M.E., Drickamer, K. & Mordoh, J. (2007). Lewis x antigen mediates adhesion of human breast carcinoma cells to activated endothelium. Possible involvement of the endothelial scavenger receptor C-type lectin. *Breast Cancer Res Treat*, **101,** 161-74.

Elston,C.W & Ellis, I.O. (1991). Pathological prognostic factors in breast cancer. The value of histological grade in breast cancer:experience from a large study with long-term follow-up. *Histopathology.* **19** (5): 403-410.

Ernster, V.L., Ballard-Barbash, R., Barlow, W.E., Zheng, Y., Weaver, D.L., Cutter, G., Yankaskas, B.C., Rosenberg, R., Carney, P.A., Kerlikowske, K., Taplin, S.H., Urban, N. & Geller, B.M. (2002). Detection of ductal carcinoma in situ in women undergoing screening mammography. *J Natl Cancer Inst*, **94,** 1546-54.

Faiz, O. & Fentiman, I.S. (1998). Hormone replacement therapy and breast cancer. *Int J Clin Pract*, **52,** 98-101.

Fenlon, S., Ellis, I.O., Bell, J., Todd, J.H., Elston, C.W. & Blamey, R.W. (1987). Helix pomatia and Ulex europeus lectin binding in human breast carcinoma. *J Pathol*, **152,** 169-76.

Fernandes, B., Sagman, U., Auger, M., Demetrio, M. & Dennis, J.W. (1991). Beta 1-6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. *Cancer Res*, **51,** 718-23.

Ferrara, N. (2001). Role of vascular endothelial growth factor in regulation of physiological angiogenesis. *Am J Physiol Cell Physiol*, **280,** C1358-66.

Fidler, I.J. (1970). Metastasis: guantitative analysis of distribution and fate of tumor embolilabeled with 125 I-5-iodo-2'-deoxyuridine. *J Natl Cancer Inst*, **45,** 773-82.

Fidler, I.J. & Bucana, C. (1977). Mechanism of tumor cell resistance to lysis by syngeneic lymphocytes. *Cancer Res*, **37,** 3945-56.

Fischer, P.B., Karlsson, G.B., Dwek, R.A. & Platt, F.M. (1996). N-butyldeoxynojirimycin-mediated inhibition of human immunodeficiency virus entry correlates with impaired gp120 shedding and gp41 exposure. *J Virol*, **70,** 7153-60.

Folkman, J. (2006). Angiogenesis. *Annu Rev Med*, **57,** 1-18.

Fraser, H.M. & Lunn, S.F. (2000). Angiogenesis and its control in the female reproductive system. *Br Med Bull*, **56,** 787-97.

Freire, T., Berois, N., Sonora, C., Varangot, M., Barrios, E. & Osinaga, E. (2006). UDP-N-acetyl-D-galactosamine:polypeptide N-cetylgalactosaminyltransferase 6 (ppGalNAc-T6) mRNA as a potential new marker for detection of bone marrow-disseminated breast cancer cells. *Int J Cancer*, **119,** 1383-8.

Frisch, S.M. & Ruoslahti, E. (1997). Integrins and anoikis. *Curr Opin Cell Biol*, **9,** 701-6.

Fukushima, K. & Yamashita, K. (2001). Interleukin-2 carbohydrate recognition modulates CTLL-2 cell proliferation. *J Biol Chem*, **276,** 7351-6.

Fukutomi, T., Itabashi, M., Tsugane, S., Yamamoto, H., Nanasawa, T. & Hirota, T. (1989). Prognostic contributions of Helix pomatia and carcinoembryonic antigen staining using histochemical techniques in breast carcinomas. *Jpn J Clin Oncol*, **19,** 127-34.

Galea, M. H., Ellis, I. O., Bell, J., Elston, C. W & Blamley, R. W. (1991). Prediction of lymph node involvement in breast cancer. *Lancet,* **338**, 392-293.

Gallagher, H. S. (1984). Pathological types of breast cancer; their pronoses. *Cancer*, **53**, 623-629.

Gallinaro, H., Sittler, A. & Jacob, M. (1986). In vivo splicing of the premRNAs from early region 3 of adenovirus-2: association of precursors, intermediates and products with hnRNP. *Nucleic Acids Res*, **14,** 4171-85.

Gao, X.D., Nishikawa, A. & Dean, N. (2004). Physical interactions between the Alg1, Alg2, and Alg11 mannosyltransferases of the endoplasmic reticulum. *Glycobiology*, **14,** 559-70.

Gerdes, J., Li, L., Schlueter, C. Duchrow, M., Wohlenberg, C., Stahmer, I., Kloth, S., Brandt, E and Flad, H.D. (1991). Immunobiochemical and molecular biological characterisation of the cell proliferation-associated nuclear antien that is defined by monoclonal antibody Ki-67. *The American Journal of Pathology*, **138**, 867-873.

Glinsky, V.V., Huflejt, M.E., Glinsky, G.V., Deutscher, S.L. & Quinn, T.P. (2000). Effects of Thomsen-Friedenreich antigen-specific peptide P-30 on beta-galactoside-mediated homotypic aggregation and adhesion to the endothelium of MDA-MB-435 human breast carcinoma cells. *Cancer Res*, **60,** 2584-8.

Gold, P., and Freeman, S. S. (1956). Specific carcinoembryonic antigen of the digestive system. *J Exp Med*, **122**, 467-475.

Golks, A., Tran, T.T., Goetschy, J.F. & Guerini, D. (2007). Requirement for O-linked N-acetylglucosaminyltransferase in lymphocytes activation. *Embo J*, **26,** 4368-79.

Gonzalez-Amaro, R. & Sanchez-Madrid, F. (1999). Cell adhesion molecules: selectins and integrins. *Crit Rev Immunol*, **19,** 389-429.

Greenough,R.B.(1925). Varying degrees of malignancy in cancer of the breast. *Journal of Cancer Research*, **9**, 453-463.

Grundbacher, F.J. (1987). Helix pomatia lectin-binding and predictive value in breast-cancer. *Lancet*, **2,** 1145.

Gu, J., Sumida, Y., Sanzen, N. & Sekiguchi, K. (2001). Laminin-10/11 and fibronectin differentially regulate integrin-dependent Rho and Rac activation via p130(Cas)-CrkII-DOCK180 pathway. *J Biol Chem*, **276,** 27090-7.

Gu, J. & Taniguchi, N. (2004). Regulation of integrin functions by N-glycans. *Glycoconj J*, **21,** 9-15.

Gu, W., Bhatia, K., Magrath, I.T., Dang, C.V. & Dalla-Favera, R. (1994). Binding and suppression of the Myc transcriptional activation domain by p107. *Science*, **264,** 251-4.

Gumbiner, B.M. (2000). Regulation of cadherin adhesive activity. *J Cell Biol*, **148,** 399-404.

Guo, Y., Feinberg, H., Conroy, E., Mitchell, D.A., Alvarez, R., Blixt, O., Taylor, M.E., Weis, W.I. & Drickamer, K. (2004). Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR. *Nat Struct Mol Biol*, **11,** 591-8.

Gurcel, C., Vercoutter-Edouart, A.S., Fonbonne, C., Mortuaire, M., Salvador, A., Michalski, J.C. & Lemoine, J. (2008). Identification of new O-GlcNAc modified proteins using a click-chemistry-based tagging. *Anal Bioanal Chem*, **390**, 2089-97.

Gusterson, C.W., Taylor, K.N & Gelber, J. (1993). Prognostic value of Helix pomatia in breast cancer. *Br J Cancer,* **68**,146-150.

Gygi, S.P., Rochon, Y., Franza, B.R. & Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. *Mol Cell Biol*, **19**, 1720-30.

Haagensen,C.D. (1986). Diseases of the breast. 3rd Edition, WB Saunders, Philadelphia PA.

Hakomori, S. (1984). Tumor-associated carbohydrate antigens. *Annu Rev Immunol*, **2,** 103-26.

Hakomori, S. (1996). Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer Res*, **56,** 5309-18.

Halsted, W. S. (1898). A clinical study of certain adenocarcinomata of the breast . *Annals of surgery,* **28**, 557.

Haltiwanger, R.S., Blomberg, M.A. & Hart, G.W. (1992). Glycosylation of nuclear and cytoplasmic proteins. Purification and characterization of a uridine diphospho-N-acetylglucosamine:polypeptide beta-N-acetylglucosaminyltransferase. *J Biol Chem*, **267,** 9005-13.

Haltiwanger, R.S., Grove, K. & Philipsberg, G.A. (1998). Modulation of O-linked N-acetylglucosamine levels on nuclear and cytoplasmic proteins in vivo using the peptide O-GlcNAc-beta-N-acetylglucosaminidase inhibitor O-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-N-phenylcarbamate. *J Biol Chem*, **273,** 3611-7.

Haltiwanger, R.S. & Stanley, P. (2002). Modulation of receptor signaling by glycosylation: fringe is an O-fucose-beta1,3-N-acetylglucosaminyltransferase. *Biochim Biophys Acta*, **1573,** 328-35.

Hammarstrom, S. & Kabat, E.A. (1969). Purification and characterization of a blood-group A reactive hemagglutinin from the snail Helix pomatia and a study of its combining site. *Biochemistry*, **8,** 2696-705.

Hammarstrom, S. & Kabat, E.A. (1971). Studies on specificity and binding properties of the blood group A reactive hemagglutinin from Helix pomatia. *Biochemistry*, **10,** 1684-92.

Hammarstrom, S., Lindberg, A.A. & Robertsson, E.S. (1972). Precipitation of lipopolysaccharides from rough mutants of Salmonella typhimurium by an A-hemagglutinin from Helix pomatia. *Eur J Biochem*, **25,** 274-82.

Hammer,C., Fanning, A & Crowe, J. (2008). Overview of breast cancer staging and surgical treatment options. *Cleveland Clinic Journal of Medicine* **75** (S1):10-16

Hanahan, D. & Weinberg, R.A. (2000). The hallmarks of cancer. *Cell*, **100,** 57-70.

Handerson, T., Camp, R., Harigopal, M., Rimm, D. & Pawelek, J. (2005). Beta1,6-branched oligosaccharides are increased in lymph node metastases and predict poor outcome in breast carcinoma. *Clin Cancer Res*, **11,** 2969-73.

Hardan, I., Weiss, L., Hershkoviz, R., Greenspoon, N., Alon, R., Cahalon, L., Reich, S., Slavin, S. & Lider, O. (1993). Inhibition of metastatic cell colonization in murine lungs and tumor-induced morbidity by non-peptidic Arg-Gly-Asp mimetics. *Int J Cancer*, **55,** 1023-8.

Harris, J., & Hellman, S. (1996). Natural history of breast cancer. In diseases of the breast. Lippincott-Raven, Philadelphia PA, 376.

Harris, R.J., Ling, V.T. & Spellman, M.W. (1992). O-linked fucose is present in the first epidermal growth factor domain of factor XII but not protein C. *J Biol Chem*, **267,** 5102-7.

Harris, R.J. & Spellman, M.W. (1993). O-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology*, **3,** 219-24.

Hart, G.W. (1997). Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. *Annu Rev Biochem*, **66,** 315-35.

Hassan, H., Reis, C.A., Bennett, E.P., Mirgorodskaya, E., Roepstorff, P., Hollingsworth, M.A., Burchell, J., Taylor-Papadimitriou, J. & Clausen, H. (2000). The lectin domain of UDP-N-acetyl-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-T4 directs its glycopeptide specificities. *J Biol Chem*, **275,** 38197-205.

Heindl, H., Greenwell, P., Weingarten, N., Kiss, T., Terstyanszky, G. & Weinzierl, R.O. Cation-pi interactions induce kinking of a molecular hinge in the RNA polymerase bridge-helix domain. *Biochem Soc Trans*, **39,** 31-5.

Helenius, A. & Aebi, M. (2004). Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem*, **73,** 1019-49.

Helenius, J., Ng, D.T., Marolda, C.L., Walter, P., Valvano, M.A. & Aebi, M. (2002). Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature*, **415,** 447-50.

Henke, C., Bitterman, P., Roongta, U., Ingbar, D. & Polunovsky, V. (1996). Induction of fibroblast apoptosis by anti-CD44 antibody: implications for the treatment of fibroproliferative lung disease. *Am J Pathol*, **149,** 1639-50.

Hennighausen, L. & Robinson, G.W. (2005). Information networks in the mammary gland. *Nat Rev Mol Cell Biol*, **6,** 715-25.

Hens, J.R. & Wysolmerski, J.J. (2005). Key stages of mammary gland development: molecular mechanisms involved in the formation of the embryonic mammary gland. *Breast Cancer Res*, **7,** 220-4.

Hernandez, J.D. & Baum, L.G. (2002). Ah, sweet mystery of death! Galectins and control of cell fate. *Glycobiology*, **12,** 127R-36R.

Hirota, M., Pour, P.M., Tempero, M.A. & Chaney, W.G. (1992). Purification and analysis of glycoproteins bearing blood group-A determinants from hamster pancreatic ductal adenocarcinomas. *Carcinogenesis*, **13,** 1829-33.

Hogdall, E. (2008). Cancer antigen 125 and prognosis. *Curr Opin Obstet Gynecol*, **20,** 4-8.

Holt, J.T., Jensen, R.A. & Page, D.L. (1993). Histopathology: old principles and new methods. *Cancer Surv*, **18,** 115-33.

Homans, S.W., Dwek, R.A. & Rademacher, T.W. (1987). Solution conformations of N-linked oligosaccharides. *Biochemistry*, **26,** 6571-8.

Horton, M.A., Lewis, D., McNulty, K., Pringle, J.A. & Chambers, T.J. (1985). Monoclonal antibodies to osteoclastomas (giant cell bone tumors): definition of osteoclast-specific cellular antigens. *Cancer Res*, **45,** 5663-9.

Howard, B.A. & Gusterson, B.A. (2000). Human breast development. *J Mammary Gland Biol Neoplasia*, **5,** 119-37.

Hsu, D.K., Dowling, C.A., Jeng, K.C., Chen, J.T., Yang, R.Y. & Liu, F.T. (1999). Galectin-3 expression is induced in cirrhotic liver and hepatocellular carcinoma. *Int J Cancer*, **81,** 519-26.

Hughes, R.C. (2001). Galectins as modulators of cell adhesion. *Biochimie*, **83,** 667-76.

Huot, J., Houle, F., Spitz, D.R. & Landry, J. (1996). HSP27 phosphorylation-mediated resistance against actin fragmentation and cell death induced by oxidative stress. *Cancer Res*, **56,** 273-9.

Hutson, T.L & Osborne, M.P.(2005). Evaluating and Staging the Patient with Breast Cancer. In: Breast cancer. Ed/s: Roses, D.F. 2nd Edition. Elsevier Churchill Livinstone. Philadelphia.

Hynes, R.O. (1992). Integrins: versatility, modulation, and signaling in cell adhesion. *Cell*, **69,** 11-25.

Ikeda, Y., Mori, M., Adachi, Y., Matsushima, T. & Sugimachi, K. (1994). Prognostic value of the histochemical expression of helix pomatia agglutinin in advanced

colorectal cancer. A univariate and multivariate analysis. *Dis Colon Rectum*, 37, 181-4.

Ikehara, Y., Kojima, N., Kurosawa, N., Kudo, T., Kono, M., Nishihara, S., Issiki, S., Morozumi, K., Itzkowitz, S., Tsuda, T., Nishimura, S.I., Tsuji, S. & Narimatsu, H. (1999). Cloning and expression of a human gene encoding an N-acetylgalactosamine-alpha2,6-sialyltransferase (ST6GalNAc I): a candidate for synthesis of cancer-associated sialyl-Tn antigens. *Glycobiology*, **9,** 1213-24.

Imperiali, B. & Hendrickson, T.L. (1995). Asparagine-linked glycosylation: specificity and function of oligosaccharyl transferase. *Bioorg Med Chem*, **3,** 1565-78.

Iozzo, R.V. & San Antonio, J.D. (2001). Heparan sulfate proteoglycans: heavy hitters in the angiogenesis arena. *J Clin Invest*, **108,** 349-55.

Irimura, T., Matsushita, Y., Sutton, R.C., Carralero, D., Ohannesian, D.W., Cleary, K.R., Ota, D.M., Nicolson, G.L. & Lotan, R. (1991). Increased content of an endogenous lactose-binding lectin in human colorectal carcinoma progressed to metastatic stages. *Cancer Res*, **51,** 387-93.

Iwakawa, K., Ueda, N., Murao, S. & Kobayashi, N. (1996). Altered carbohydrate composition in colorectal adenomas and carcinomas: histochemical characterization of N-acetylgalactosamine, L-fucose, and o-acetylated sialic acid. *J Gastroenterol*, **31**, 24-32.

Jacobson, B.S., Stolz, D.B. & Schnitzer, J.E. (1996). Identification of endothelial cell-surface proteins as targets for diagnosis and treatment of disease. *Nat Med*, **2,** 482-4.

Jaeken, J., Carchon, H. & Stibler, H. (1993). The carbohydrate-deficient glycoprotein syndromes: pre-Golgi and Golgi disorders? *Glycobiology*, **3,** 423-8.

Jalkanen, S. & Jalkanen, M. (1992). Lymphocyte CD44 binds the COOH-terminal heparin-binding domain of fibronectin. *J Cell Biol*, **116,** 817-25.

Jamal B, Nita Lazar M, Zhennan G, Bokr A, J, w. & M, k. (2009). N-glycosylation status of E-cadherin controls cytoskeletal dynamics through the organization of distinct β-catenin- and γ-catenin-containing AJs. *Cell Health and Cytoskeleton***,** 67-80.

Johansen, P.G., Marshall, R.D. & Neuberger, A. (1961). Carbohydrates in protein. 3 The preparation and some of the properties of a glycopeptide from hen's-egg albumin. *Biochem J*, **78,** 518-27.

Jones, J., Krag, S.S. & Betenbaugh, M.J. (2005). Controlling N-linked glycan site occupancy. *Biochim Biophys Acta*, **1726,** 121-37.

Jones, M., Tussey, L., Athanasou, N. & Jackson, D.G. (2000). Heparan sulfate proteoglycan isoforms of the CD44 hyaluronan receptor induced in human

inflammatory macrophages can function as paracrine regulators of fibroblast growth factor action. *J Biol Chem*, **275,** 7964-74.

Jones, S.P., Zachara, N.E., Ngoh, G.A., Hill, B.G., Teshima, Y., Bhatnagar, A., Hart, G.W. & Marban, E. (2008). Cardioprotection by N-acetylglucosamine linkage to cellular proteins. *Circulation*, **117,** 1172-82.

Ju, T., Brewer, K., D'Souza, A., Cummings, R.D. & Canfield, W.M. (2002). Cloning and expression of human core 1 beta1,3-galactosyltransferase. *J Biol Chem*, **277,** 178-86.

Julien, S., Adriaenssens, E., Ottenberg, K., Furlan, A., Courtand, G., Vercoutter-Edouart, A.S., Hanisch, F.G., Delannoy, P. & Le Bourhis, X. (2006). ST6GalNAc I expression in MDA-MB-231 breast cancer cells greatly modifies their O-glycosylation pattern and enhances their tumourigenicity. *Glycobiology*, **16,** 54-64.

Julien, S., Krzewinski-Recchi, M.A., Harduin-Lepers, A., Gouyer, V., Huet, G., Le Bourhis, X. & Delannoy, P. (2001). Expression of sialyl-Tn antigen in breast cancer cells transfected with the human CMP-Neu5Ac: GalNAc alpha2,6-sialyltransferase (ST6GalNac I) cDNA. *Glycoconj J*, **18,** 883-93.

Jung, T.Y., Jung, S., Ryu, H.H., Jeong, Y.I., Jin, Y.H., Jin, S.G., Kim, I.Y., Kang, S.S. & Kim, H.S. (2008). Role of galectin-1 in migration and invasion of human glioblastoma multiforme cell lines. *J Neurosurg*, **109,** 273-84.

Juul, A., Teilmann, G., Scheike, T., Hertel, N.T., Holm, K., Laursen, E.M., Main, K.M. & Skakkebaek, N.E. (2006). Pubertal development in Danish children: comparison of recent European and US data. *Int J Androl*, **29,** 247-55; discussion 286-90.

Kakeji, Y., Maehara, Y., Tsujitani, S., Baba, H., Ohno, S., Watanabe, A. & Sugimachi, K. (1994). Helix pomatia agglutinin binding activity and lymph node metastasis in patients with gastric cancer. *Semin Surg Oncol*, **10,** 130-4.

Kakeji, Y., Tsujitani, S., Mori, M., Maehara, Y. & Sugimachi, K. (1991). Helix pomatia agglutinin binding activity is a predictor of survival time for patients with gastric carcinoma. *Cancer*, **68,** 2438-42.

Kaplan, R.N., Psaila, B. & Lyden, D. (2006). Bone marrow cells in the 'pre-metastatic niche': within bone and beyond. *Cancer Metastasis Rev*, **25,** 521-9.

Kariya, Y., Kato, R., Itoh, S., Fukuda, T., Shibukawa, Y., Sanzen, N., Sekiguchi, K., Wada, Y., Kawasaki, N. & Gu, J. (2008). N-Glycosylation of laminin-332 regulates its biological functions. A novel function of the bisecting GlcNAc. *J Biol Chem*, **283,** 33036-45.

Kariya, Y., Kawamura, C., Tabei, T. & Gu, J. Bisecting GlcNAc residues on laminin-332 down-regulate galectin-3-dependent keratinocyte motility. *J Biol Chem*, **285,** 3330-40.

Kato, M., Herz, F., Kato, S. & Hirano, A. (1992). Expression of stress-response (heat-shock) protein 27 in human brain tumors: an immunohistochemical study. *Acta Neuropathol*, 83, 420-2.

Kawai, T., Suzuki, M., Torikata, C. & Suzuki, Y. (1991). Expression of blood group-related antigens and Helix pomatia agglutinin in malignant pleural mesothelioma and pulmonary adenocarcinoma. *Hum Pathol*, **22,** 118-24.

Kawamura, H., Tomozoe, Y., Akagi, T., Kamei, D., Ochiai, M. & Yamada, M. (2002). Identification of the nucleocytoplasmic shuttling sequence of heterogeneous nuclear ribonucleoprotein D-like protein JKTBP and its interaction with mRNA. *J Biol Chem*, **277,** 2732-9.

Kelleher, D.J. & Gilmore, R. (2006). An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology*, **16,** 47R-62R.

Kellokumpu, S., Sormunen, R. & Kellokumpu, I. (2002). Abnormal glycosylation and altered Golgi structure in colorectal cancer: dependence on intra-Golgi pH. *FEBS Lett*, **516,** 217-24.

Key, T.J. & Pike, M.C. (1988). The role of oestrogens and progestagens in the epidemiology and prevention of breast cancer. *Eur J Cancer Clin Oncol*, **24,** 29-43.

Keydar,I., Chen, L., Karby,S., Weiss, F.R., Delarea, J., Radu, M., Chaitcik, S & Brenner, H.J. (1979). Establishmentand characterisation of a cell line of human breast carcinoma origin. *European Journal of cancer*,**15**, 659-670

Kilpatrick, D.C. (2002). Animal lectins: a historical introduction and overview. *Biochim Biophys Acta*, **1572,** 187-97.

Kim, J. D., Manderino, G. L., Gooch, G. T., Gould, V. E., Koukoulis, G. K. And Tomita, J. T. (1991). Isolation and characterization of the tumor-associated mucin-like glycoprotein defined by monoclonal antibody A-80*. Cancer Ther,* Contr*.* **1,** 227-293.

Kim, Y.J., Borsig, L., Varki, N.M. & Varki, A. (1998). P-selectin deficiency attenuates tumor growth and metastasis. *Proc Natl Acad Sci U S A*, **95,** 9325-30.

King, M.C. (1992). Breast cancer genes: how many, where and who are they? *Nat Genet*, **2,** 89-90.

Kohama, S.G., Goss, J.R., Finch, C.E. & McNeill, T.H. (1995). Increases of glial fibrillary acidic protein in the aging female mouse brain. *Neurobiol Aging*, **16,** 59-67.

Kohler, S., Ullrich, S., Richter, U. & Schumacher, U. E-/P-selectins and colon carcinoma metastasis: first in vivo evidence for their crucial role in a clinically relevant model of spontaneous metastasis formation in the lung. *Br J Cancer*, **102,** 602-9.

Kohler, W., Prokop, O. & Kuhnemund, O. (1973). Routine identification of group-C streptococci by means of an agglutinin (protectin) from the albumen gland of the edible snail, Helix pomatia. *J Med Microbiol*, **6,** 127-30.

Konecny, G.E., Meng, Y.G., Untch, M., Wang, H.J., Bauerfeind, I., Epstein, M., Stieber, P., Vernes, J.M., Gutierrez, J., Hong, K., Beryt, M., Hepp, H., Slamon, D.J. & Pegram, M.D. (2004). Association between HER-2/neu and vascular endothelial growth factor expression predicts clinical outcome in primary breast cancer patients. *Clin Cancer Res*, **10,** 1706-16.

Korczak, B., Goss, P., Fernandez, B., Baker, M. & Dennis, J.W. (1994). Branching N-linked oligosaccharides in breast cancer. *Adv Exp Med Biol*, **353,** 95-104.

Korczak, B., Le, T., Elowe, S., Datti, A. & Dennis, J.W. (2000). Minimal catalytic domain of N-acetylglucosaminyltransferase V. *Glycobiology*, **10,** 595-9.

Kramer, R.H., Vogel, K.G. & Nicolson, G.L. (1982). Tumor cell interactions with vascular endothelial cells and their extracellular matrix. *Prog Clin Biol Res*, **89,** 333-51.

Kreidberg, J.A. (2000). Functions of alpha3beta1 integrin. *Curr Opin Cell Biol*, **12,** 548-53.

Krishnamurthy,S. (2005). Pathology of Regional Lymph Nodes. In: Breast cancer. Ed/s. Roses, D. F. 2nd Edition.Elsevier Churchill Livingstone. Philadelphia.

Krueger, K. E., and Srivastava, S. (2009). Post-translational Protein Modifications. *Molecular & Cellular Proteomics*, vol **5** (10): 1799-1809.

Ktistakis, N.T., Kao, C.Y., Wang, R.H. & Roth, M.G. (1995). A fluorescent lipid analogue can be used to monitor secretory activity and for isolation of mammalian secretion mutants. *Mol Biol Cell*, **6,** 135-50.

Kuhl, C.K., Schrading, S., Bieling, H.B., Wardelmann, E., Leutner, C.C., Koenig, R., Kuhn, W. & Schild, H.H. (2007). MRI for diagnosis of pure ductal carcinoma in situ: a prospective observational study. *Lancet*, **370,** 485-92.

Kukuruzinska, M.A. & Robbins, P.W. (1987). Protein glycosylation in yeast: transcript heterogeneity of the ALG7 gene. *Proc Natl Acad Sci U S A*, **84,** 2145-9.

Kuwabara, H., Uda, H. & Takenaka, I. (1997). Immunohistochemical detection of sialosyl-Tn antigen in carcinoma of the prostate. *Br J Urol*, **80,** 456-9.

Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J. & Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, **393,** 648-59.

Laczy, B., Hill, B.G., Wang, K., Paterson, A.J., White, C.R., Xing, D., Chen, Y.F., Darley-Usmar, V., Oparil, S. & Chatham, J.C. (2009). Protein O-

GlcNAcylation: a new signaling paradigm for the cardiovascular system. *Am J Physiol Heart Circ Physiol*, **296,** H13-28.

Laemmli, I. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature, vol **227**: 680–685.

Laidler, P. & Litynska, A. (1997). Tumor cell N-glycans in metastasis. *Acta Biochim Pol*, **44,** 343-57.

Laitinen, L., Juusela, H. & Virtanen, I. (1990). Binding of the blood group-reactive lectins to human adult kidney specimens. *Anat Rec*, **226,** 10-7.

Lakhani, S.R. (1999). The transition from hyperplasia to invasive carcinoma of the breast. *J Pathol*, **187,** 272-8.

Langdon, R.H., Cuccui, J. & Wren, B.W. (2009). N-linked glycosylation in bacteria: an unexpected application. *Future Microbiol*, **4,** 401-12.

Lanigan, F., O'Connor, D., Martin, F. & Gallagher, W.M. (2007). Molecular links between mammary gland development and breast cancer. *Cell Mol Life Sci*, **64,** 3159-84.

LaRue, H., Parent-Vaugeois, C., Bergeron, A., Champetier, S. & Fradet, Y. (1997). Influence of spatial configuration on the expression of carcinoembryonic antigen and mucin antigens in human bladder cancer. *Int J Cancer*, **71,** 986-92.

Lasky, L.A. (1992). Selectins: interpreters of cell-specific carbohydrate information during inflammation. *Science*, **258,** 964-9.

Lasky, L.A. (1995). Selectin-carbohydrate interactions and the initiation of the inflammatory response. *Annu Rev Biochem*, **64,** 113-39.

Lasfargues, E.Y., Coutinho, W.G & Redfield, E.S. (1978). Isolation of two human tumour epithelial cell lines from solid breast carcinomas. *Journal of the National Cancer Insitute*, **61**, 967-978.

Lazarus, B.D., Love, D.C. & Hanover, J.A. (2009). O-GlcNAc cycling: implications for neurodegenerative disorders. *Int J Biochem Cell Biol*, **41,** 2134-46.

Lazarus, M.B., Nam, Y., Jiang, J., Sliz, P. & Walker, S. (2011). Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature*, **469,** 564-7.

Leathem, A.J. & Brooks, S.A. (1987). Predictive value of lectin binding on breast-cancer recurrence and survival. *Lancet*, **1,** 1054-6.

Lefebvre, T., Guinez, C., Dehennaut, V., Beseme-Dekeyser, O., Morelle, W. & Michalski, J.C. (2005). Does O-GlcNAc play a role in neurodegenerative diseases? *Expert Rev Proteomics*, **2,** 265-75.

Lefterova, M.I. & Lazar, M.A. (2009). New developments in adipogenesis. *Trends Endocrinol Metab*, **20,** 107-14.

Lehmann, F., Tiralongo, E. & Tiralongo, J. (2006). Sialic acid-specific lectins: occurrence, specificity and function. *Cell Mol Life Sci*, **63,** 1331-54.

Lehner, I., Niehof, M. & Borlak, J. (2003). An optimized method for the isolation and identification of membrane proteins. *Electrophoresis*, **24,** 1795-808.

Leivonen, M., Nordling, S., Lundin, J., von Boguslawski, K. & Haglund, C. (2001). STn and prognosis in breast cancer. *Oncology*, **61,** 299-305.

Lescar, J., Sanchez, J.F., Audfray, A., Coll, J.L., Breton, C., Mitchell, E.P. & Imberty, A. (2007). Structural basis for recognition of breast and colon cancer epitopes Tn antigen and Forssman disaccharide by Helix pomatia lectin. *Glycobiology*, **17,** 1077-83.

Levi, G., Tarrab-Hazdai, R. & Teichberg, V.I. (1983). Prevention and therapy with electrolectin of experimental autoimmune myasthenia gravis in rabbits. *Eur J Immunol*, **13,** 500-7.

Levy, L., Broad, S., Diekmann, D., Evans, R.D. & Watt, F.M. (2000). beta1 integrins regulate keratinocyte adhesion and differentiation by distinct mechanisms. *Mol Biol Cell*, **11,** 453-66.

Li, H., Chien, P.C., Jr., Tuen, M., Visciano, M.L., Cohen, S., Blais, S., Xu, C.F., Zhang, H.T. & Hioe, C.E. (2008). Identification of an N-linked glycosylation in the C4 region of HIV-1 envelope gp120 that is critical for recognition of neighboring CD4 T cell epitopes. *J Immunol*, **180,** 4011-21.

Li, Y., Luo, L., Rasool, N. & Kang, C.Y. (1993). Glycosylation is necessary for the correct folding of human immunodeficiency virus gp120 in CD4 binding. *J Virol*, **67,** 584-8.

Lidgren, A., Bergh, A., Grankvist, K., Rasmuson, T. & Ljungberg, B. (2008). Glucose transporter-1 expression in renal cell carcinoma and its correlation with hypoxia inducible factor-1 alpha. *BJU Int*, 101, 480-4.

Lis, H. & Sharon, N. (1993). Protein glycosylation. Structural and functional aspects. *Eur J Biochem*, **218,** 1-27.

Lohi, J. (2001). Laminin-5 in the progression of carcinomas. *Int J Cancer*, **94,** 763-7.

Lomax-Browne, H.J. (2009). Interactions between Cancer Cell Glycans and Endothelial Cells during Adhesion and Transmigration Events in Metastasis. PhD thesis, Oxford Brookes University, Oxford.

Lotan, R., Ito, H., Yasui, W., Yokozaki, H., Lotan, D. & Tahara, E. (1994). Expression of a 31-kDa lactoside-binding lectin in normal human gastric mucosa and in primary and metastatic gastric carcinomas. *Int J Cancer*, **56,** 474-80.

Lowe, J.B. & Marth, J.D. (2003). A genetic approach to Mammalian glycan function. *Annu Rev Biochem*, **72,** 643-91.

Ly, A. & Stamatatos, L. (2000). V2 loop glycosylation of the human immunodeficiency virus type 1 SF162 envelope facilitates interaction of this protein with CD4 and CCR5 receptors and protects the virus from neutralization by anti-V3 loop and anti-CD4 binding site antibodies. *J Virol*, **74,** 6769-76.

Macedo, N.M., Costa-Lotufo, L.V., Pessoa, C., Moraes, M.O., Bonfim, L.R. & Ramos, M.V. (2007). Inhibition of sea urchin fertilization by plant lectins. *Protein Pept Lett*, **14,** 886-93.

MacMahon, B., Cole, P. & Brown, J. (1973). Etiology of human breast cancer: a review. *J Natl Cancer Inst*, **50,** 21-42.

Magidovich, H. & Eichler, J. (2009). Glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of the N-glycosylation pathway in the third domain of life. *FEMS Microbiol Lett*, **300,** 122-30.

Malkin, D., Li, F.P., Strong, L.C., Fraumeni, J.F., Jr., Nelson, C.E., Kim, D.H., Kassel, J., Gryka, M.A., Bischoff, F.Z., Tainsky, M.A. & *et al*. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*, **250,** 1233-8.

Mandel, U., Hassan, H., Therkildsen, M.H., Rygaard, J., Jakobsen, M.H., Juhl, B.R., Dabelsteen, E. & Clausen, H. (1999). Expression of polypeptide GalNAc-transferases in stratified epithelia and squamous cell carcinomas: immunohistological evaluation using monoclonal antibodies to thee members of the GalNAc-transferase family. *Glycobiology*, **9,** 43-52.

Marcos, N.T., Pinho, S., Grandela, C., Cruz, A., Samyn-Petit, B., Harduin-Lepers, A., Almeida, R., Silva, F., Morais, V., Costa, J., Kihlberg, J., Clausen, H. & Reis, C.A. (2004). Role of the human ST6GalNAc-I and ST6GalNAc-II in the synthesis of the cancer-associated sialyl-Tn antigen. *Cancer Res*, **64,** 7050-7.

Marcum, J.A. & Rosenberg, R.D. (1985). Heparinlike molecules with anticoagulant activity are synthesized by cultured endothelial cells. *Biochem Biophys Res Commun*, **126,** 365-72.

Markiv, A., Peiris, D., Curley, G.P., Odell, M. & Dwek, M.V. (2011).Identification, cloning and characterization of two N-acetylgalactosamine binding lectins from the albumen gland of Helix pomatia. *J Biol Chem.*

Marquez, M., Nilsson, S., Lennartsson, L., Liu, Z., Tammela, T., Raitanen, M. & Holmberg, A.R. (2004). Charge-dependent targeting: results in six tumor cell lines. *Anticancer Res*, 24**,** 1347-51.

Mayrand, S.H., Dwen, P. & Pederson, T. (1993). Serine/threonine phosphorylation regulates binding of C hnRNP proteins to pre-mRNA. *Proc Natl Acad Sci U S A*, **90,** 7764-8.

McPherson, K., Steel, C.M. & Dixon, J.M. (2000). ABC of breast diseases. Breast cancer-epidemiology, risk factors, and genetics. *Bmj*, **321,** 624-8.

Mikecz, K., Brennan, F.R., Kim, J.H. & Glant, T.T. (1995). Anti-CD44 treatment abrogates tissue oedema and leukocyte infiltration in murine arthritis. *Nat Med*, **1,** 558-63.

Millis, R. R. (1984). Atlas of breast pathology. In current Histopathology series. Ed. Lancaster, Gresham, G. A. MTP Press. England, 11.

Mitchell, B.S., Vernon, K., Schumacher, U. & Habil, M. (1995). Ultrastructural localization of Helix pomatia agglutinin (HPA)-binding sites in human breast cancer cell lines and characterization of HPA-binding glycoproteins by western blotting. *Ultrastruct Pathol*, **19,** 51-9.

Mitchell, S.W. (1860). Researches upon the venom of rattlesnake. *Smithson. Contrib. Knowl*. XII, 89-90.

Miyamoto, S. (2006). Clinical applications of glycomic approaches for the detection of cancer and other diseases. *Curr Opin Mol Ther*, **8,** 507-13.

Molloy, M.P. (2000). Two-dimensional electrophoresis of membrane proteins using immobilized pH gradients. *Anal Biochem*, **280,** 1-10.

Moloney, D.J., Shair, L.H., Lu, F.M., Xia, J., Locke, R., Matta, K.L. & Haltiwanger, R.S. (2000). Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *J Biol Chem*, **275,** 9604-11.

Monaghan, P., Perusinghe, N.P., Cowen, P. & Gusterson, B.A. (1990). Peripubertal human breast development. *Anat Rec*, **226,** 501-8.

Montreuil, J. (1980). Primary structure of glycoprotein glycans: basis for the molecular biology of glycoproteins. *Adv Carbohydr Chem Biochem*, **37,** 157-223.

Morikawa, Y., Moore, J.P. & Jones, I.M. (1990). HIV-1 envelope protein gp120 expression by secretion in E. coli: assessment of CD4 binding and use in epitope mapping. *J Virol Methods*, **29,** 105-13.

Morimoto, K., Satoh-Yamaguchi, K., Hamaguchi, A., Inoue, Y., Takeuchi, M., Okada, M., Ikeda, W., Takai, Y. & Imai, T. (2008). Interaction of cancer cells with

platelets mediated by Necl-5/poliovirus receptor enhances cancer cell metastasis to the lungs. *Oncogene*, **27,** 264-73.

Mourali, N., Muenz, L.R., Tabbane, F., Belhassen, S., Bahi, J. & Levine, P.H. (1980). Epidemiologic features of rapidly progressing breast cancer in Tunisia. *Cancer*, **46,** 2741-6.

Murayama, T., Zuber, C., Seelentag, W.K., Li, W.P., Kemmner, W., Heitz, P.U. & Roth, J. (1997). Colon carcinoma glycoproteins carrying alpha 2,6-linked sialic acid reactive with Sambucus nigra agglutinin are not constitutively expressed in normal human colon mucosa and are distinct from sialyl-Tn antigen. *Int J Cancer*, **70,** 575-81.

Nadanaka, S., Sato, C., Kitajima, K., Katagiri, K., Irie, S. & Yamagata, T. (2001). Occurrence of oligosialic acids on integrin alpha 5 subunit and their involvement in cell adhesion to fibronectin. *J Biol Chem*, **276,** 33657-64.

Nakagawa, H., Zheng, M., Hakomori, S., Tsukamoto, Y., Kawamura, Y. & Takahashi, N. (1996). Detailed oligosaccharide structures of human integrin alpha 5 beta 1 analyzed by a three-dimensional mapping technique. *Eur J Biochem*, **237,** 76-85.

Nakano, S., Tsuruta, J. & Iyama, K. (2000). [Remodeling of basement membrane in association with cancer invasion]. *Rinsho Byori*, **48,** 451-7.

Nash, G.F., Turner, L.F., Scully, M.F. & Kakkar, A.K. (2002). Platelets and cancer. *Lancet Oncol*, **3,** 425-30.

Natali, P.G., Nicotra, M.R., Botti, C., Mottolese, M., Bigotti, A. & Segatto, O. (1992). Changes in expression of alpha 6/beta 4 integrin heterodimer in primary and metastatic breast cancer. *Br J Cancer*, **66,** 318-22.

Nicolson, G.L., Dulski, K., Basson, C. & Welch, D.R. (1985). Preferential organ attachment and invasion in vitro by B16 melanoma cells selected for differing metastatic colonization and invasive properties. *Invasion Metastasis*, **5,** 144-58.

Nie, L., Wu, G. & Zhang, W. (2006). Correlation of mRNA expression and protein abundance affected by multiple sequence features related to translational efficiency in Desulfovibrio vulgaris: a quantitative analysis. *Genetics*, 174, 2229-43.

Nishimura, H., Kawabata, S., Kisiel, W., Hase, S., Ikenaka, T., Takao, T., Shimonishi, Y. & Iwanaga, S. (1989). Identification of a disaccharide (Xyl-Glc) and a trisaccharide (Xyl2-Glc) O-glycosidically linked to a serine residue in the first epidermal growth factor-like domain of human factors VII and IX and protein Z and bovine protein Z. *J Biol Chem*, **264,** 20320-5.

O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, 250, 4007-21.

Offner, H., Celnik, B., Bringman, T.S., Casentini-Borocz, D., Nedwin, G.E. & Vandenbark, A.A. (1990). Recombinant human beta-galactoside binding lectin suppresses clinical and histological signs of experimental autoimmune encephalomyelitis. *J Neuroimmunol*, **28,** 177-84.

Ogden, R.C. & Adams, D.A. (1987). Electrophoresis in agarose and acrylamide gels. *Methods Enzymol*, **152,** 61-87.

Oku, N., Tokudome, Y., Koike, C., Nishikawa, N., Mori, H., Saiki, I. & Okada, S. (1996). Liposomal Arg-Gly-Asp analogs effectively inhibit metastatic B16 melanoma colonization in murine lungs. *Life Sci*, **58,** 2263-70.

Opdenakker, G., Rudd, P.M., Ponting, C.P. & Dwek, R.A. (1993). Concepts and principles of glycobiology. *Faseb J*, **7,** 1330-7.

Orntoft, T.F., Holmes, E.H., Johnson, P., Hakomori, S. & Clausen, H. (1991). Differential tissue expression of the Lewis blood group antigens: enzymatic, immunohistologic, and immunochemical evidence for Lewis a and b antigen expression in Le(a-b-) individuals. *Blood*, **77,** 1389-96.

Orntoft, T.F. & Vestergaard, E.M. (1999). Clinical aspects of altered glycosylation of glycoproteins in cancer. *Electrophoresis*, **20,** 362-71.

Osborne, C.K., Hamilton, B., Titus, G. & Livingston, R.B. (1980). Epidermal growth factor stimulation of human breast cancer cells in culture. *Cancer Res*, **40,** 2361-6.

Pannell. P., Kotasek. D. (1997). Cancer and Clinical Biochemistry. Clinical Biochemistry in Medicine Series. Association of Clinical Biochemists, ACB Venture Publications. London.

Papaspyridonos, M., McNeill, E., de Bono, J.P., Smith, A., Burnand, K.G., Channon, K.M. & Greaves, D.R. (2008). Galectin-3 is an amplifier of inflammation in atherosclerotic plaque progression through macrophage activation and monocyte chemoattraction. *Arterioscler Thromb Vasc Biol*, **28,** 433-40.

Pappin, D.J., Hojrup, P. & Bleasby, A.J. (1993). Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol*, **3,** 327-32.

Parker, R., Flowers, T.J., Moore, A.L. & Harpham, N.V. (2006). An accurate and reproducible method for proteome profiling of the effects of salt stress in the rice leaf lamina. *J Exp Bot*, **57,** 1109-18.

Parks, A.G. (1959). The micro-anatomy of the breast. *Ann R Coll Surg Engl*, **25,** 235-51.

Patchett, R.A., Kelly, A.F. & Kroll, R.G. (1991). The adsorption of bacteria to immobilized lectins. *J Appl Bacteriol*, **71,** 277-84.

Perone, M.J., Bertera, S., Shufesky, W.J., Divito, S.J., Montecalvo, A., Mathers, A.R., Larregina, A.T., Pang, M., Seth, N., Wucherpfennig, K.W., Trucco, M., Baum, L.G. & Morelli, A.E. (2009). Suppression of autoimmune diabetes by soluble galectin-1. *J Immunol*, **182,** 2641-53.

Petersen, H.V., Jensen, J.N., Stein, R. & Serup, P. (2002). Glucose induced MAPK signalling influences NeuroD1-mediated activation and nuclear localization. *FEBS Lett*, **528,** 241-5.

Peto, J., Collins, N., Barfoot, R., Seal, S., Warren, W., Rahman, N., Easton, D.F., Evans, C., Deacon, J. & Stratton, M.R. (1999). Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *J Natl Cancer Inst*, **91,** 943-9.

Petrescu, A.J., Milac, A.L., Petrescu, S.M., Dwek, R.A. & Wormald, M.R. (2004). Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology*, **14,** 103-14.

Pilatte, Y., Bignon, J. & Lambre, C.R. (1993). Sialic acids as important molecules in the regulation of the immune system: pathophysiological implications of sialidases in immunity. *Glycobiology*, **3,** 201-18.

Piller, V., Piller, F. & Cartron, J.P. (1990). Comparison of the carbohydrate-binding specificities of seven N-acetyl-D-galactosamine-recognizing lectins. *Eur J Biochem*, 191, 461-6.

Pinho, S., Marcos, N.T., Ferreira, B., Carvalho, A.S., Oliveira, M.J., Santos-Silva, F., Harduin-Lepers, A. & Reis, C.A. (2007). Biological significance of cancer-associated sialyl-Tn antigen: modulation of malignant phenotype in gastric carcinoma cells. *Cancer Lett*, **249,** 157-70.

Pinho, S.S., Reis, C.A., Paredes, J., Magalhaes, A.M., Ferreira, A.C., Figueiredo, J., Xiaogang, W., Carneiro, F., Gartner, F. & Seruca, R. (2009). The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. *Hum Mol Genet*, **18,** 2599-608.

Polyak, K. (2001). On the birth of breast cancer. *Biochim Biophys Acta*, **1552,** 1-13.

Polyak, K. (2007). Breast cancer stem cells: a case of mistaken identity? *Stem Cell Rev*, **3,** 107-9.

Pouliot, N., Nice, E.C. & Burgess, A.W. (2001). Laminin-10 mediates basal and EGF-stimulated motility of human colon carcinoma cells via alpha(3)beta(1) and alpha(6)beta(4) integrins. *Exp Cell Res*, **266,** 1-10.

Prokop, O., Rackwitz, A. & Schlesinger, D. (1965). A "new" human blood group receptor Ahel tested with saline extracts from Helix hortensis (garden snail). *J Forensic Med*, **12,** 108-10

Prokopishyn, N.L., Puzon-McLaughlin, W., Takada, Y. & Laferte, S. (1999). Integrin alpha3beta1 expressed by human colon cancer cells is a major carrier of oncodevelopmental carbohydrate epitopes. *J Cell Biochem*, 72, 189-209.

Prosper, F. & Verfaillie, C.M. (2001). Regulation of hematopoiesis through adhesion receptors. *J Leukoc Biol*, **69,** 307-16.

Pure, E. & Cuff, C.A. (2001). A crucial role for CD44 in inflammation. *Trends Mol Med*, **7,** 213-21.

Rabilloud, T. (1998). Use of thiourea to increase the solubility of membrane proteins in two-dimensional electrophoresis. *Electrophoresis*, **19,** 758-60.

Rabilloud, T., Adessi, C., Giraudel, A. & Lunardi, J. (1997). Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis*, 18, 307-16.

Rabinovich, G.A., Liu, F.T., Hirashima, M. & Anderson, A. (2007). An emerging role for galectins in tuning the immune response: lessons from experimental models of inflammatory disease, autoimmunity and cancer. *Scand J Immunol*, **66,** 143-58.

Rambaruth, N.D. & Dwek, M.V. (2011) Cell surface glycan-lectin interactions in tumor metastasis. *Acta Histochem*, **113,** 591-600.

Rane, M.J., Pan, Y., Singh, S., Powell, D.W., Wu, R., Cummins, T., Chen, Q., McLeish, K.R. & Klein, J.B. (2003). Heat shock protein 27 controls apoptosis by regulating Akt activation. *J Biol Chem*, **278,** 27828-35.

Redondo, D.A.G.P., Nakamura, C.V., de Souza, W. & Morgado-Diaz, J.A. (2004). Differential expression of sialic acid and N-acetylgalactosamine residues on the cell surface of intestinal epithelial cells according to normal or metastatic potential. *J Histochem Cytochem*, **52,** 629-40.

Reynolds, L.P. & Redmer, D.A. (2001). Angiogenesis in the placenta. *Biol Reprod*, **64,** 1033-40.

Rivinoja, A., Kokkonen, N., Kellokumpu, I. & Kellokumpu, S. (2006). Elevated Golgi pH in breast and colorectal cancer cells correlates with the expression of oncofetal carbohydrate T-antigen. *J Cell Physiol*, 208, 167-74.

Robinson, C.J. & Stringer, S.E. (2001). The splice variants of vascular endothelial growth factor (VEGF) and their receptors. *J Cell Sci*, **114,** 853-65.

Ronnov-Jessen, L., Petersen, O.W. & Bissell, M.J. (1996). Cellular changes involved in conversion of normal to malignant breast: importance of the stromal reaction. *Physiol Rev*, **76,** 69-125.

Roth, J. (1984). Cytochemical localization of terminal N-acetyl-D-galactosamine residues in cellular compartments of intestinal goblet cells: implications for the topology of O-glycosylation. *J Cell Biol*, **98,** 399-406.

Rudiger, H. & Gabius, H.J. (2001). Plant lectins: occurrence, biochemistry, functions and applications. *Glycoconj J*, **18,** 589-613.

Ruoslahti, E. (1996). Integrin signaling and matrix assembly. *Tumour Biol*, **17,** 117-24.

Rush, J.S., Gao, N., Lehrman, M.A. & Waechter, C.J. (2008). Recycling of dolichyl monophosphate to the cytoplasmic leaflet of the endoplasmic reticulum after the cleavage of dolichyl pyrophosphate on the lumenal monolayer. *J Biol Chem*, **283,** 4087-93.

Russo, J., Calaf, G., Roi, L. & Russo, I.H. (1987). Influence of age and gland topography on cell kinetics of normal human breast tissue. *J Natl Cancer Inst*, **78,** 413-8.

Ryder, S.D., Smith, J.A. & Rhodes, J.M. (1992). Peanut lectin: a mitogen for normal human colonic epithelium and human HT29 colorectal cancer cells. *J Natl Cancer Inst*, **84,** 1410-6.

Saint-Guirons, J., Zeqiraj, E., Schumacher, U., Greenwell, P. & Dwek, M. (2007). Proteome analysis of metastatic colorectal cancer cells recognized by the lectin Helix pomatia agglutinin (HPA). *Proteomics*, **7,** 4082-9.

Saito, H., Gu, J., Nishikawa, A., Ihara, Y., Fujii, J., Kohgo, Y. & Taniguchi, N. (1995). Organization of the human N-acetylglucosaminyltransferase V gene. *Eur J Biochem*, **233,** 18-26.

Salmi, M., Kalimo, K. & Jalkanen, S. (1993). Induction and function of vascular adhesion protein-1 at sites of inflammation. *J Exp Med*, **178,** 2255-60.

Sanchez, J.F., Lescar, J., Chazalet, V., Audfray, A., Gagnon, J., Alvarez, R., Breton, C., Imberty, A. & Mitchell, E.P. (2006). Biochemical and structural analysis of Helix pomatia agglutinin. A hexameric lectin with a novel fold. *J Biol Chem*, **281,** 20171-80.

Santoni, V., Molloy, M. & Rabilloud, T. (2000). Membrane proteins and proteomics: un amour impossible? *Electrophoresis*, **21,** 1054-70.

Sata, T., Roth, J., Zuber, C., Stamm, B. & Heitz, P.U. (1991). Expression of alpha 2,6-linked sialic acid residues in neoplastic but not in normal human colonic mucosa. A lectin-gold cytochemical study with Sambucus nigra and Maackia amurensis lectins. *Am J Pathol*, **139,** 1435-48.

Schenk, B., Fernandez, F. & Waechter, C.J. (2001). The ins(ide) and out(side) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum. *Glycobiology*, **11,** 61R-70R.

Schmidt-Rhode, P., Schulz, K.D., Sturm, G., Raab-Frick, A. & Prinz, H. (1987). CA 15.3 as a tumour marker in breast cancer. *Int J Biol Markers*, **2,** 135-42.

Schuessler, M.H., Pintado, S., Welt, S., Real, F.X., Xu, M., Melamed, M.R., Lloyd, K.O. & Oettgen, H.F. (1991). Blood group and blood-group-related antigens in normal pancreas and pancreas cancer: enhanced expression of precursor type 1, Tn and sialyl-Tn in pancreas cancer. *Int J Cancer*, **47,** 180-7.

Schumacher, U. & Adam, E. (1997). Lectin histochemical HPA-binding pattern of human breast and colon cancers is associated with metastases formation in severe combined immunodeficient mice. *Histochem J*, 29, 677-84.

Schumacher, U., Adam, E., Brooks, S.A. & Leathem, A.J. (1995). Lectin-binding properties of human breast cancer cell lines and human milk with particular reference to Helix pomatia agglutinin. *J Histochem Cytochem*, 43, 275-81.

Schumacher, U., Higgs, D., Loizidou, M., Pickering, R., Leathem, A. & Taylor, I. (1994). Helix pomatia agglutinin binding is a useful prognostic indicator in colorectal carcinoma. *Cancer*, **74,** 3104-7.

Schumacher, U., Kretzschmar, H., Brooks, S. & Leathem, A. (1992). Helix pomatia lectin binding pattern of brain metastases originating from breast cancers. *Pathol Res Pract*, **188,** 284-6.

Seidenfaden, R., Krauter, A., Schertzinger, F., Gerardy-Schahn, R. & Hildebrandt, H. (2003). Polysialic acid directs tumor cell growth by controlling heterophilic neural cell adhesion molecule interactions. *Mol Cell Biol*, **23,** 5908-18.

Semmes, O.J., Malik, G. & Ward, M. (2006). Application of mass spectrometry to the discovery of biomarkers for detection of prostate cancer. *J Cell Biochem*, **98,** 496-503.

Senda, M., Ito, A., Tsuchida, A., Hagiwara, T., Kaneda, T., Nakamura, Y., Kasama, K., Kiso, M., Yoshikawa, K., Katagiri, Y., Ono, Y., Ogiso, M., Urano, T., Furukawa, K., Oshima, S. & Furukawa, K. (2007). Identification and expression of a sialyltransferase responsible for the synthesis of disialylgalactosylgloboside in normal and malignant kidney cells: downregulation of ST6GalNAc VI in renal cancers. *Biochem J*, **402,** 459-70.

Sewell, R., Backstrom, M., Dalziel, M., Gschmeissner, S., Karlsson, H., Noll, T., Gatgens, J., Clausen, H., Hansson, G.C., Burchell, J. & Taylor-Papadimitriou, J. (2006). The ST6GalNAc-I sialyltransferase localizes throughout the Golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer. *J Biol Chem*, **281,** 3586-94.

Shafi, R., Iyer, S.P., Ellies, L.G., O'Donnell, N., Marek, K.W., Chui, D., Hart, G.W. & Marth, J.D. (2000). The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc Natl Acad Sci U S A*, **97,** 5735-9

Shao, L. & Haltiwanger, R.S. (2003). O-fucose modifications of epidermal growth factor-like repeats and thrombospondin type 1 repeats: unusual modifications in unusual places. *Cell Mol Life Sci*, **60,** 241-50.

Shapiro, S. (1989). Determining the efficacy of breast cancer screening. *Cancer*, **63,** 1873-80.

Sharma, V. & Surolia, A. (1997). Analyses of carbohydrate recognition by legume lectins: size of the combining site loops and their primary specificity. *J Mol Biol*, **267,** 433-45.

Shiraishi, T., Atsumi, S. & Yatani, R. (1992). Comparative study of prostatic carcinoma bone metastasis among Japanese in Japan and Japanese Americans and whites in Hawaii. *Adv Exp Med Biol*, **324,** 7-16.

Shiraishi, T., Nakayama, T., Fukutome, K., Watanabe, M. & Murata, T. (1999). Malignant myoepithelioma of the breast metastasizing to the jaw. *Virchows Arch*, **435,** 520-3.

Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A & McGuire, W.L. (1987). Human breast cancer: Correlation of Relapse and Survival with Amplification of HER-2/neu oncogene. *Science*, **235**, 177-182.

Slawson, C., Copeland, R.J. & Hart, G.W. O-GlcNAc signaling: a metabolic link between diabetes and cancer? *Trends Biochem Sci*, **35,** 547-55.

Slawson, C., Lakshmanan, T., Knapp, S. & Hart, G.W. (2008). A mitotic GlcNAcylation/phosphorylation signaling complex alters the posttranslational state of the cytoskeletal protein vimentin. *Mol Biol Cell*, **19,** 4130-40.

Slifkin, M. & Cumbie, R. (1989). Rapid detection of herpes simplex virus with fluorescein-labeled Helix pomatia lectin. *J Clin Microbiol*, **27,** 1036-9.

Slovin, S.F., Ragupathi, G., Musselli, C., Fernandez, C., Diani, M., Verbel, D., Danishefsky, S., Livingston, P. & Scher, H.I. (2005). Thomsen-Friedenreich (TF) antigen as a target for prostate cancer vaccine: clinical trial results with TF cluster (c)-KLH plus QS21 conjugate vaccine in patients with biochemically relapsed prostate cancer. *Cancer Immunol Immunother*, **54,** 694-702.

Smets, L.A. & Van Beek, W.P. (1984). Carbohydrates of the tumor cell surface. *Biochim Biophys Acta*, **738**, 237-49.

Soule, H.D., Vazguez, J., Long, A., Albert, S & Brennan, M. (1973). A human cell line from a pleural effusion derived fron a breast carcinoma. Journal of the National Cancer Institute. **51,** 1409-1416.

Spiro, R.G. (1973). Glycoproteins. *Adv Protein Chem*, **27,** 349-467.

Spiro, R.G. (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology*, **12,** 43R-56R.

Springer, G.F. (1989). Tn epitope (N-acetyl-D-galactosamine alpha-O-serine/theonine) density in primary breast carcinoma: a functional predictor of aggressiveness. *Mol Immunol*, **26,** 1-5.

Springer, G.F. (1997). Immunoreactive T and Tn epitopes in cancer diagnosis, prognosis, and immunotherapy. *J Mol Med*, **75,** 594-602.

Springer, G.F., Desai, P.R., Wise, W., Carlstedt, S.C., Tegtmeyer, H., Stein, R. & Scanlon, E.F. (1990). Pancarcinoma T and Tn epitopes: autoimmunogens and diagnostic markers that reveal incipient carcinomas and help establish prognosis. *Immunol Ser*, **53,** 587-612.

Springer, T.A. (1990). Adhesion receptors of the immune system. *Nature*, **346,** 425-34.

Springer, T.A. (1995). Traffic signals on endothelium for lymphocyte recirculation and leukocyte emigration. *Annu Rev Physiol*, **57,** 827-72.

Srinivas, U., Pahlsson, P. & Lundblad, A. (1996). E-selectin: sialyl Lewis, a dependent adhesion of colon cancer cells, is inhibited differently by antibodies against E-selectin ligands. *Scand J Immunol*, **44,** 197-203.

Stamenkovic, I., Amiot, M., Pesando, J.M. & Seed, B. (1989). A lymphocyte molecule implicated in lymph node homing is a member of the cartilage link protein family. *Cell*, **56,** 1057-62.

Steeg, P.S. (2006). Tumor metastasis: mechanistic insights and clinical challenges. *Nat Med*, **12,** 895-904.

Steeg, P.S., de la Rosa, A., Flatow, U., MacDonald, N.J., Benedict, M. & Leone, A. (1993). Nm23 and breast cancer metastasis. *Breast Cancer Res Treat*, **25,** 175-87.

Stellwagen, E. & Stellwagen, N.C. (2002). The free solution mobility of DNA in Tris-acetate-EDTA buffers of different concentrations, with and without added NaCl. *Electrophoresis*, **23,** 1935-41.

Stenman, U.H., Abrahamsson, P.A., Aus, G., Lilja, H., Bangma, C., Hamdy, F.C., Boccon-Gibod, L. & Ekman, P. (2005). Prognostic value of serum markers for prostate cancer. *Scand J Urol Nephrol Suppl*, 64-81.

Stillmark, H. (1888). Uber Ricin, ein giftiges Ferment aus den Samen von *Ricinus communis* L. und einigen anderen Euphorbiaceen. Inaug. Diss Dorpat.

Stoll, B.A. (1997). Breast cancer: further metabolic-endocrine risk markers? *Br J Cancer*, **76,** 1652-4.

Storm, F.K., Mahvi, D.M. & Gilchist, K.W. (1996). Lack of association between tumor necrosis and hsp-27 expression in primary breast cancer. *J Surg Oncol*, **61,** 14-6.

Sullivan, R.P., Mortimer, G & Muircheartaigh, I.O. (1993). Cell proliferation in breast tumours: analysis of histological parameters Ki67 and PCNA expression. *Irish Journal of Medical Science,* **162,** 343-347.

Sumida, H., Nakamura, H., Thompson, R.P. & Yasuda, M. (1997). Binding of lectins to novel migration promoters on cardiac mesenchymal cells in the chick. *Cell Struct Funct*, **22,** 413-20.

Suzuki, M., Suzuki, M., Nakayama, J., Suzuki, A., Angata, K., Chen, S., Sakai, K., Hagihara, K., Yamaguchi, Y. & Fukuda, M. (2005). Polysialic acid facilitates tumor invasion by glioma cells. *Glycobiology*, **15,** 887-94.

Szymanski, C.M. & Wren, B.W. (2005). Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol*, **3,** 225-37.

Takahashi, H., Oyaizu, T., Fujita, Y. & Tsubura, A. (1994). Lectin-binding profiles in MNNG-induced shrew esophageal carcinomas. *Anticancer Res*, **14,** 1569-72.

Takano, Y., Teranishi, Y., Terashima, S., Motoki, R. & Kawaguchi, T. (2000). Lymph node metastasis-related carbohydrate epitopes of gastric cancer with submucosal invasion. *Surg Today*, **30,** 1073-82.

Takeichi, M. (1990). Cadherins: a molecular family important in selective cell-cell adhesion. *Annu Rev Biochem*, **59,** 237-52.

Takenaka, Y., Fukumori, T. & Raz, A. (2004). Galectin-3 and metastasis. *Glycoconj J*, **19,** 543-9.

Tan, S., Seow, T.K., Liang, R.C., Koh, S., Lee, C.P., Chung, M.C. & Hooi, S.C. (2002). Proteome analysis of butyrate-treated human colon cancer cells (HT-29). *Int J Cancer*, **98,** 523-31.

Tanabe, K.K., Ellis, L.M. & Saya, H. (1993). Expression of CD44R1 adhesion molecule in colon carcinomas and metastases. *Lancet*, **341,** 725-6.

Tarone, G., Hirsch, E., Brancaccio, M., De Acetis, M., Barberis, L., Balzac, F., Retta, S.F., Botta, C., Altruda, F. & Silengo, L. (2000). Integrin function and regulation in development. *Int J Dev Biol*, **44,** 725-31.

Taylor-Papadimitriou, J. and Lane, E. B. (1987). Keratin expression in the mammary gland. In The Mammary Gland: Development, Regulation and Function. Ed. Neville M. C. and Daniel C. W. Plenum Press, New York, 181-215.

Ten Hagen, K.G., Bedi, G.S., Tetaert, D., Kingsley, P.D., Hagen, F.K., Balys, M.M., Beres, T.M., Degand, P. & Tabak, L.A. (2001). Cloning and characterization of a ninth member of the UDP-GalNAc:polypeptide N-

acetylgalactosaminyltransferase family, ppGaNTase-T9. *J Biol Chem*, **276,** 17395-404.

Ten Hagen, K.G., Fritz, T.A. & Tabak, L.A. (2003). All in the family: the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferases. *Glycobiology*, **13,** 1R-16R.

Ten Hagen, K.G., Hagen, F.K., Balys, M.M., Beres, T.M., Van Wuyckhuyse, B. & Tabak, L.A. (1998). Cloning and expression of a novel, tissue specifically expressed member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family. *J Biol Chem*, **273,** 27749-54.

Thomas, H.V., Key, T.J., Allen, D.S., Moore, J.W., Dowsett, M., Fentiman, I.S. & Wang, D.Y. (1997). A prospective study of endogenous serum hormone concentrations and breast cancer risk in post-menopausal women on the island of Guernsey. *Br J Cancer*, **76,** 401-5.

Thompson, D., Szabo, C.I., Mangion, J., Oldenburg, R.A., Odefrey, F., Seal, S., Barfoot, R., Kroeze-Jansema, K., Teare, D., Rahman, N., Renard, H., Mann, G., Hopper, J.L., Buys, S.S., Andrulis, I.L., Senie, R., Daly, M.B., West, D., Ostrander, E.A., Offit, K., Peretz, T., Osorio, A., Benitez, J., Nathanson, K.L., Sinilnikova, O.M., Olah, E., Bignon, Y.J., Ruiz, P., Badzioch, M.D., Vasen, H.F., Futreal, A.P., Phelan, C.M., Narod, S.A., Lynch, H.T., Ponder, B.A., Eeles, R.A., Meijers-Heijboer, H., Stoppa-Lyonnet, D., Couch, F.J., Eccles, D.M., Evans, D.G., Chang-Claude, J., Lenoir, G., Weber, B.L., Devilee, P., Easton, D.F., Goldgar, D.E. & Stratton, M.R. (2002). Evaluation of linkage of breast cancer to the putative BRCA3 locus on chromosome 13q21 in 128 multiple case families from the Breast Cancer Linkage Consortium. *Proc Natl Acad Sci U S A*, **99,** 827-31.

Toniolo, P.G., Levitz, M., Zeleniuch-Jacquotte, A., Banerjee, S., Koenig, K.L., Shore, R.E., Strax, P. & Pasternack, B.S. (1995). A prospective study of endogenous estrogens and breast cancer in postmenopausal women. *J Natl Cancer Inst*, **87,** 190-7.

Tortora, G., Grabowski, S. (1993). *Principles of Anatomy and Physiology*. Seventh Edition. Harper Collins. New York.

Trojanowicz, B., Winkler, A., Hammje, K., Chen, Z., Sekulla, C., Glanz, D., Schmutzler, C., Mentrup, B., Hombach-Klonisch, S., Klonisch, T., Finke, R., Kohrle, J., Dralle, H. & Hoang-Vu, C. (2009). Retinoic acid-mediated down-regulation of ENO1/MBP-1 gene products caused decreased invasiveness of the follicular thyroid carcinoma cell lines. *J Mol Endocrinol*, **42,** 249-60.

Tsai, S.T., Chien, I.H., Shen, W.H., Kuo, Y.Z., Jin, Y.T., Wong, T.Y., Hsiao, J.R., Wang, H.P., Shih, N.Y. & Wu, L.W. ENO1, a potential prognostic head and neck cancer marker, promotes transformation partly via chemokine CCL20 induction. *Eur J Cancer*, **46,** 1712-23.

Uhlenbruck, G. & Prokop, O. (1966). An agglutinin from Helix pomatia, which reacts with terminal N-acetyl-D-galactosamine. *Vox Sang*, **11,** 519-20.

Valcu, C.M. & Valcu, M. (2007). Reproducibility of two-dimensional gel electrophoresis at different replication levels. *J Proteome Res*, **6,** 4677-83.

Van den Steen, P., Rudd, P.M., Dwek, R.A. & Opdenakker, G. (1998). Concepts and principles of O-linked glycosylation. *Crit Rev Biochem Mol Biol*, **33,** 151-208.

Varki, A. (1993). Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology*, 3, 97-130.

Varki, A. (1994). Selectin ligands. *Proc Natl Acad Sci U S A*, **91,** 7390-7.

Varki, A. (1997). Selectin ligands: will the real ones please stand up? *J Clin Invest*, **100,** S31-5.

Varki, A., Cummings, R ., Esko, J., Freeze, H., Hart, G. and Marth, J. (1991). *Essential of Glycobiology*. Cold Spring Harbor Laboratory Press. New York.

Vazquez-Martin, C., Cuevas, E., Gil-Martin, E. & Fernandez-Briera, A. (2004). Correlation analysis between tumor-associated antigen sialyl-Tn expression and ST6GalNAc I activity in human colon adenocarcinoma. *Oncology*, **67,** 159-65.

Vogel, C.L., Gobleigh, M.A., Tripathy, D.,Gutheil, J.C & Stewart S.J. (2002). Efficacy and Safety of Trastuzumab as a single agent in first-line treatment of HER2-overexpressing Metastatic Breast cancer . *Journal of clinical oncology*, **20**, 719-726.

Vorherr , H. (1974). *The Breast: Morphology, Physiology, and Lactation*. Academic Press. New York, NY.

Vretblad, P., Hjorth, R. & Laas, T. (1979). The isolectins of Helix pomatia. Separation by isoelectric focusing and preliminary characterization. *Biochim Biophys Acta*, **579,** 52-61.

Walsh, M.T., Watzlawick, H., Putnam, F.W., Schmid, K. & Brossmer, R. (1990). Effect of the carbohydrate moiety on the secondary structure of beta 2-glycoprotein. I. Implications for the biosynthesis and folding of glycoproteins. *Biochemistry*, **29,** 6250-7.

Wang, F., Goto, M., Kim, Y.S., Higashi, M., Imai, K., Sato, E. & Yonezawa, S. (2001). Altered GalNAc-alpha-2,6-sialylation compartments for mucin-associated sialyl-Tn antigen in colorectal adenoma and adenocarcinoma. *J Histochem Cytochem*, **49,** 1581-92.

Wassarman, P.M. (1990). Profile of a mammalian sperm receptor. *Development*, **108,** 1-17.

Wasylnka, J.A., Simmer, M.I. & Moore, M.M. (2001). Differences in sialic acid density in pathogenic and non-pathogenic Aspergillus species. *Microbiology*, **147,** 869-77.

Weerapana, E. & Imperiali, B. (2006). Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology*, **16,** 91R-101R.

Weiderpass, E., Gridley, G., Persson, I., Nyren, O., Ekbom, A. & Adami, H.O. (1997). Risk of endometrial and breast cancer in patients with diabetes mellitus. *Int J Cancer*, **71,** 360-3.

Weidner, N. (2002). New paradigm for vessel intravasation by tumor cells. *Am J Pathol*, **160,** 1937-9.

Weigand, R.A., Isenberg, W.M., Russo, J., Brennan, M.J. & Rich, M.A. (1982). Blood vessel invasion and axillary lymph node involvement as prognostic indicators for human breast cancer. *Cancer*, **50,** 962-9.

Weiss, L. (2000). The Prediction of Invasion, Metastasis and Progression by Grading. *Cancer and Metastasis Reviews*, **19,** 315-325.

Weiss, W. & Gorg, A. (2008). Sample solublization buffers for two-dimensional electrophoresis. *Methods Mol Biol*, **424,** 35-42.

Welinder, C., Jansson, B., Ferno, M., Olson, H. & Baldetorp, B. (2009). Expression of Helix pomatia lectin binding glycoproteins in women with breast cancer in relationship to their blood group phenotypes. *J Proteome Res*, **8,** 782-7.

Wells, L., Vosseller, K. & Hart, G.W. (2001). Glycosylation of nucleocytoplasmic proteins: signal transduction and O-GlcNAc. *Science*, **291**, 2376-8.

Welsh, J.B., Sapinoso, L.M., Kern, S.G., Brown, D.A., Liu, T., Bauskin, A.R., Ward, R.L., Hawkins, N.J., Quinn, D.I., Russell, P.J., Sutherland, R.L., Breit, S.N., Moskaluk, C.A., Frierson, H.F., Jr. & Hampton, G.M. (2003). Large-scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. *Proc Natl Acad Sci U S A*, **100,** 3410-5.

Wenger, C.r., Beardslee, S., Owens,M.A Pounds, G., Oldaker, T., Vendley, P & McGuire, W.L. (1993). DNA ploidy, S-phase, and steriod receptors in more than 127,000 breast cancer patients. *Breast cancer research and treatment.* **28,** 9-20.

Wheelock, M.J. & Johnson, K.R. (2003). Cadherins as modulators of cellular phenotype. *Annu Rev Cell Dev Biol*, **19,** 207-35.

Wolff, E.A., Greenfield, B., Taub, D.D., Murphy, W.J., Bennett, K.L. & Aruffo, A. (1999). Generation of artificial proteoglycans containing glycosaminoglycan-modified CD44. Demonstration of the interaction between rantes and chondroitin sulfate. *J Biol Chem*, **274,** 2518-24.

Wooster, R., Bignell, G., Lancaster, J., Swift, S., Seal, S., Mangion, J., Collins, N., Gregory, S., Gumbs, C. & Micklem, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature*, **378,** 789-92.

Wopereis, S., Lefeber, D.J., Morava, E. & Wevers, R.A. (2006). Mechanisms in protein O-glycan biosynthesis and clinical and molecular aspects of protein O-glycan biosynthesis defects: a review. *Clin Chem*, **52,** 574-600.

Wygrecka, M., Marsh, L.M., Morty, R.E., Henneke, I., Guenther, A., Lohmeyer, J., Markart, P. & Preissner, K.T. (2009). Enolase-1 promotes plasminogen-mediated recruitment of monocytes to the acutely inflamed lung. *Blood*, **113,** 5588-98.

Wu, A.M. & Sugii, S. (1991). Coding and classification of d-galactose, N-acetyl-d-galactosamine, and [beta]-d-Galp-[1-->3(4)]-[beta]-d-GlcpNAc, specificities of applied lectins. *Carbohydrate Research*, **213,** 127-143.

Xiong, L., Andrews, D. & Regnier, F. (2003). Comparative proteomics of glycoproteins based on lectin selection and isotope coding. *J Proteome Res*, **2,** 618-25.

Xu, H., Manivannan, A., Liversidge, J., Sharp, P.F., Forrester, J.V. & Crane, I.J. (2002). Involvement of CD44 in leukocyte trafficking at the blood-retinal barrier. *J Leukoc Biol*, **72,** 1133-41.

Yang, R.Y. & Liu, F.T. (2003). Galectins in cell growth and apoptosis. *Cell Mol Life Sci*, **60,** 267-76.

Yonezawa, S. (1994). Application of immunohistochemistry for diagnosis of neoplasms - mucin antigens expression and biological behaviour of neoplasms. *Acta Histochemica and Cytochemica*, **27**, 561-566.

Yoshida, Y., Okamura, T., Yano, K. & Ezaki, T. (1994). Silver stained nucleolar organizer region proteins and Helix pomatia agglutinin immunostaining in esophageal carcinoma: correlated prognostic factors. *J Surg Oncol*, **56,** 116-21.

Yoshimura, M., Nishikawa, A., Ihara, Y., Taniguchi, S. & Taniguchi, N. (1995). Suppression of lung metastasis of B16 mouse melanoma by N-acetylglucosaminyltransferase III gene transfection. *Proc Natl Acad Sci U S A*, **92,** 8754-8.

Young, W.W., Jr., Holcomb, D.R., Ten Hagen, K.G. & Tabak, L.A. (2003). Expression of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase isoforms in murine tissues determined by real-time PCR: a new view of a large family. *Glycobiology*, **13,** 549-57.

Yu, C.J., Shih, J.Y., Lee, Y.C., Shun, C.T., Yuan, A. & Yang, P.C. (2005). Sialyl Lewis antigens: association with MUC5AC protein and correlation with post-operative recurrence of non-small cell lung cancer. *Lung Cancer*, **47,** 59-67.

Yu, Q. & Stamenkovic, I. (1999). Localization of matrix metalloproteinase 9 to the cell surface provides a mechanism for CD44-mediated tumor invasion. *Genes Dev*, **13,** 35-48.

Zhang, D., Tai, L.K., Wong, L.L., Putti, T.C., Sethi, S.K., Teh, M. & Koay, E.S. (2008). Proteomic characterization of differentially expressed proteins in breast cancer: Expression of hnRNP H1, RKIP and GRP78 is strongly associated with HER-2/neu status. *Proteomics Clin Appl*, **2,** 99-107.

Zhang, M., Gaschen, B., Blay, W., Foley, B., Haigwood, N., Kuiken, C. & Korber, B. (2004). Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin. *Glycobiology*, **14,** 1229-46.

Zhou, J., Nong, L., Wloch, M., Cantor, A., Mulshine, J.L. & Tockman, M.S. (2001). Expression of early lung cancer detection marker: hnRNP-A2/B1 and its relation to microsatellite alteration in non-small cell lung cancer. *Lung Cancer*, **34,** 341-50.

Zhou, Z., Wang, J., Cao, R., Morita, H., Soininen, R., Chan, K.M., Liu, B., Cao, Y. & Tryggvason, K. (2004). Impaired angiogenesis, delayed wound healing and retarded tumor growth in perlecan heparan sulfate-deficient mice. *Cancer Res*, **64,** 4699-702.

Zutter, M.M., Krigman, H.R. & Santoro, S.A. (1993). Altered integrin expression in adenocarcinoma of the breast. Analysis by in situ hybridization. *Am J Pathol*, **142,** 1439-48.

Appendices

Appendix 1

Recipes for buffer used in 1DE and 2DE experiments

Resolving gel buffer (1M Tris-HCl pH 8.8)	
Tris (FW 121.1)	1M
dH_2O	150ml
HCl	to pH 8.8
dH_2O	To 200 ml

Stacking gel buffer (0.5M Tris-HCl pH 6.8)	
Tris (FW 121.1)	1M
dH_2O	150ml
HCl	to pH 8.8
dH_2O	To 200 ml

Water-saturated n-butanol	
n-butanol	50 ml
dH_2O	50 ml

15 % Ammonium persulfate	
Ammonium persulfate (FW 228.2)	15 % w/v
dH_2O	1 ml

10 % SDS	
SDS (FW 288.38)	10 %
dH_2O	to 100 ml

Resolving gel (5 ml per Minigel)	10%	12%
Acrylamide/bisacrylamide 40%	1.25 ml	1.5 ml
Resolving buffer	1.9 ml	1.9 ml
dH_2O	1.8 ml	1.55 ml
10% SDS	50 µl	50 µl
15% APS	25 µl	25 µl
TEMED	2.5 µl	2.5 µl

Stacking gel (2.5 ml) 4%	
Acrylamide/bisacrylamide 40%	0.25 ml
Stacking buffer	0.63 ml
dH_2O	1.5 ml
10% SDS	25 µl
15% APS	1.25 µl
TEMED	2.5 µl

Running buffer	
Tris (FW 121.1)	0.025M
Glycine(FW 75.07)	150ml
SDS(FW 288.4)	0.1%w/v
dH_2O	To 1L

Recipes for buffer used in agarose gel electrophoresis

10 x TAE buffer

Tris-base	0.4M
Glacial acetic acid	0.2M
EDTA –Na_2-salt	0.01M
Distilled water	To final volume of 1000 ml

Appendix 2

Average normalised volume, fold difference and p values of protein 1-18

Spot No	Average normalised Volume				Fold difference	Anova(p)
	HMT3522	BT474	MCF-7	T47D		
1	1.96E+05	2.45E+05	2.79E+05	7.05E+05	3.6	7.27E-08
2	1.56E+04	2.37E+04	2.29E+04	5.47E+04	3.5	1.98E-07
3	1.82E+04	1.21E+04	2.68E+04	2.66E+04	2.2	0.00E+00
4	1.17E+04	1.13E+04	2.11E+04	1.78E+04	6.6	0.00E+00
5	1.62E+05	2.18E+05	2.21E+05	3.08E+05	1.9	4.56E-07
6	1.31E+06	1.04E+05	4.08E+05	7.14E+05	3.2	1.97E-06
7	1.41E+05	1.55E+05	1.76E+05	3.51E+05	2.5	1.46E-05
8	1.00E+05	1.29E+05	3.84E+04	5.17E+04	3.4	3.56E-06
9	1.96E+05	1.96E+05	2.41E+04	4.12E+04	2.1	2.31E-14
10	2.83E+05	1.72E+05	7.60E+05	9.17E+05	6.1	3.10E-07
11	1.83E+04	1.48E+04	2.46E+04	8.41E+04	5.7	3.34E-07
12	4.89E+05	3.69E+05	2.95E+05	1.96E+05	2.5	1.18E-11
13	2.95E+05	3.81E+05	4.28E+05	4.57E+05	1.6	5.85E-10
14	4.04E+05	1.33E+05	5.94E+05	7.27E+05	5.4	1.33E-06
15	1.42E+05	1.41E+05	3.37E+05	6.04E+05	4.3	1.35E-08
16	5.28E+03	1.01E+04	3.17E+04	5.68E+04	10.8	6.77E-08
17	1.18E+04	1.85E+05	2.22E+05	3.58E+05	30.3	6.40E-09
18	3.16E+04	1.98E+05	3.73E+05	5.75E+05	2.9	5.77E-15

Appendix 3

Sequences and percentage coverage of protein spots 1-18

Protein spot 1

Glutamate sunthetase enoyl-CoA hydratase-like protein [Homo sapiens] Sequence Coverage: 7%

```
  1 MAAGIVASRR LRDLLTRRLT GSNYPGLSIS LRLTGSSAQE EASGVALGEA
 51 PDHSYESLRV TSAQKHVLHV QLNRPNKRNA MNKVFWREMV ECFNKISRDA
101 DCRAVVISGA GKMFTAGIDL MDMASDILQP KGDDVARISW YLRDIITRYQ
151 ETFNVIERCP KPVIAAVHGG CIGGGVDLVT ACDIRYCAQD AFFQVKEVDV
201 GLAADVGTLE RLPKVIGNQS LVNELAFTAH KMMADEALDS GLVSRVFPDK
251 EVMLDAALPL APEISSKTTV LVQSTKVNLL YSRDHSVAES LNYVASWNMS
301 MLQTQDLVKS VQPTTENKEL KTVTFSKL
```

Protein spot 2

heterogeneous nuclear ribonucleoprotein H1 [Homo sapiens]
Sequence Coverage: 5%

```
  1 MMLGTEGGEG FVVKVRGLPW SCSADEVQRF FSDCKIQNGA QGIRFIYTRE
 51 GRPSGEAFVE LESEDEVKLA LKKDRETMGH RYVEVFKSNN VEMDWVLKHT
101 GPNSPDTAND GFVRLRGLPF GCSKEEIVQF FSGLEIVPNG ITLPVDFQGR
151 STGEAFVQFA SQEIAEKALK KHKERIGHRY IEIFKSSRAE VRTHYDPPRK
201 LMAMQRPGPY DRPGAGRGYN SIGRGAGFER MRRGAYGGGY GGYDDYNGYN
251 DGYGFGSDRF GRDLNYCFSG MSDHRYGDGG STFQSTTGHC VHMRGLPYRA
301 TENDIYNFFS PLNPVRVHIE IGPDGRVTGE ADVEFATHED AVAAMSKDKA
351 NMQHRYVELF LNSTAGASGG AYEHRYVELF LNSTAGASGG AYGSQMMGGM
401 GLSNQSSYGG PASQQLSGGY GGGYGGQSSM SGYDQVLQEN SSDFQSNIA
```

Protein spot 3

heterogeneous nuclear ribonucleoprotein H1 [Homo sapiens]
Sequence Coverage: 9%

```
  1 MMLGTEGGEG FVVKVRGLPW SCSADEVQRF FSDCKIQNGA QGIRFIYTRE
 51 GRPSGEAFVE LESEDEVKLA LKKDRETMGH RYVEVFKSNN VEMDWVLKHT
101 GPNSPDTAND GFVRLRGLPF GCSKEEIVQF FSGLEIVPNG ITLPVDFQGR
151 STGEAFVQFA SQEIAEKALK KHKERIGHRY IEIFKSSRAE VRTHYDPPRK
201 LMAMQRPGPY DRPGAGRGYN SIGRGAGFER MRRGAYGGGY GGYDDYNGYN
251 DGYGFGSDRF GRDLNYCFSG MSDHRYGDGG STFQSTTGHC VHMRGLPYRA
301 TENDIYNFFS PLNPVRVHIE IGPDGRVTGE ADVEFATHED AVAAMSKDKA
351 NMQHRYVELF LNSTAGASGG AYEHRYVELF LNSTAGASGG AYGSQMMGGM
401 GLSNQSSYGG PASQQLSGGY GGGYGGQSSM SGYDQVLQEN SSDFQSNIA
```

Protein spot 4

heterogeneous nuclear ribonucleoprotein H1 [Homo sapiens]

Sequence Coverage: 7%

```
1 MMLGTEGGEG FVVKVRGLPW SCSADEVQRF FSDCKIQNGA QGIRFIYTRE
   51 GRPSGEAFVE LESEDEVKLA LKKDRETMGH RYVEVFKSNN VEMDWVLKHT
  101 GPNSPDTAND GFVRLRGLPF GCSKEEIVQF FSGLEIVPNG ITLPVDFQGR
  151 STGEAFVQFA SQEIAEKALK KHKERIGHRY IEIFKSSRAE VRTHYDPPRK
  201 LMAMQRPGPY DRPGAGRGYN SIGRGAGFER MRRGAYGGGY GGYDDYNGYN
  251 DGYGFGSDRF GRDLNYCFSG MSDHRYGDGG STFQSTTGHC VHMRGLPYRA
  301 TENDIYNFFS PLNPVRVHIE IGPDGRVTGE ADVEFATHED AVAAMSKDKA
  351 NMQHRYVELF LNSTAGASGG AYEHRYVELF LNSTAGASGG AYGSQMMGGM
  401 GLSNQSSYGG PASQQLSGGY GGGYGGQSSM SGYDQVLQEN SSDFQSNIA
```

Protein spot 5

elongation factor Tu

Sequence Coverage: 8%

```
  1 MAAATLLRAT PHFSGLAAGR TFLLQGLLRL LKAPALPLLC RGLAVEAKKT
   51 YVRDKPHVNV GTIGHVDHGK TTLTAAITKI LAEGGGAKFK KYEEIDNAPE
  101 ERARGITINA AHVEYSTAAR HYAHTDCPGH ADYVKNMITG TAPLDGCILV
  151 VAANDGPMPQ TREHLLLARQ IGVEHVVVYV NKADAVQDSE MVELVELEIR
  201 ELLTEFGYKG EETPVIVGSA LCALEGRDPE LGLKSVQKLL DAVDTYIPVP
  251 ARDLEKPFLL PVEAVYSVPG RGTVVTGTLE RGILKKGDEC ELLGHSKNIR
  301 TVVTGIEMFH KSLERAEAGD NLGALVRGLK REDLRRGLVM VKPGSIKPHQ
  351 KVEAQVYILS KEEGGRHKPF VSHFMPVMFS LTWNMACRII LPPEKELAMP
  401 GEDLKFNLIL RQPMILEKGQ RFTLRDGNRT IGTGLVTNTL AMTEEEKNIK
  451 WG
```

Protein spot 6

enolase 1 [Homo sapiens]

Sequence Coverage: 7%

```
1 MSILKIHARE IFDSRGNPTV EVDLFTSKGL FRAAVPSGAS TGIYEALELR
   51 DNDKTRYMGK GVSKAVEHIN KTIAPALVSK KLNVTEQEKI DKLMIEMDGT
  101 ENKSKFGANA ILGVSLAVCK AGAVEKGVPL YRHIADLAGN SEVILPVPAF
  151 NVINGGSHAG NKLAMQEFMI LPVGAANFRE AMRIGAEVYH NLKNVIKEKY
  201 GKDATNVGDE GGFAPNILEN KEGLELLKTA IGKAGYTDKV VIGMDVAASE
  251 FFRSGKYDLD FKSPDDPSRY ISPDQLADLY KSFIKDYPVV SIEDPFDQDD
  301 WGAWQKFTAS AGIQVVGDDL TVTNPKRIAK AVNEKSCNCL LLKVNQIGSV
  351 TESLQACKLA QANGWGVMVS HRSGETEDTF IADLVVGLCT GQIKTGAPCR
  401 SERLAKYNQL LRIEEELGSK AKFAGRNFRN PLAK
```

Protein spot 7

elongation factor Tu

Sequence Coverage: 12%

```
1 MAAATLLRAT PHFSGLAAGR TFLLQGLLRL LKAPALPLLC RGLAVEAKKT
   51 YVRDKPHVNV GTIGHVDHGK TTLTAAITKI LAEGGGAKFK KYEEIDNAPE
  101 ERARGITINA AHVEYSTAAR HYAHTDCPGH ADYVKNMITG TAPLDGCILV
  151 VAANDGPMPQ TREHLLLARQ IGVEHVVVYV NKADAVQDSE MVELVELEIR
  201 ELLTEFGYKG EETPVIVGSA LCALEGRDPE LGLKSVQKLL DAVDTYIPVP
  251 ARDLEKPFLL PVEAVYSVPG RGTVVTGTLE RGILKKGDEC ELLGHSKNIR
  301 TVVTGIEMFH KSLERAEAGD NLGALVRGLK REDLRRGLVM VKPGSIKPHQ
  351 KVEAQVYILS KEEGGRHKPF VSHFMPVMFS LTWNMACRII LPPEKELAMP
  401 GEDLKFNLIL RQPMILEKGQ RFTLRDGNRT IGTGLVTNTL AMTEEEKNIK
  451 WG
```

Protein spot 8

macropain subunit delta [Homo sapiens]
Sequence Coverage: 10%

```
1 IANRVTDKLT PIHDRIFCCR SGSAADTQAV ADAVTYQLGF HSIELNEPPL
   51 VHTAASLFKE MCYRYREDLM AGIIIAGWDP QEGGQVYSVP MGGMMVRQSF
  101 AIGGSGSSYI YGYVDATYRE GMTKEECLQF TANALALAME RDGSSGGVIR
  151 LAAIAESGVE RQVLLGDQIP KFAVATLPPA
```

Protein spot 9

proteasome subunit alpha type 2 [Homo sapiens]
Sequence Coverage: 5%

```
1 MAERGYSFSL TTFSPSGKLV QIEYALAAVA GGAPSVGIKA ANGVVLATEK
   51 KQKSILYDER SVHKVEPITK HIGLVYSGMG PDYRVLVHRA RKLAQQYYLV
  101 YQEPIPTAQL VQRVASVMQE YTQSGGVRPF GVSLLICGWN EGRPYLFQSD
  151 PSGAYFAWKA TAMGKNYVNG KTFLEKRYNE DLELEDAIHT AILTLKESFE
  201 GQMTEDNIEV GICNEAGFRR LTPTEVKDYL AAIA
```

Protein spot 10

heat shock protein 27 [Homo sapiens]
Sequence Coverage: 17%

```
1 MTERRVPFSL LRGPSWDPFR DWYPHSRLFD QAFGLPRLPE EWSQWLGGSS
   51 WPGYVRPLPP AAIESPAVAA PAYSRALSRQ LSSGVSEIRH TADRWRVSLD
  101 VNHFAPDELT VKTKDGVVEI TGKHEERQDE HGYISRCFTR KYTLPPGVDP
  151 TQVSSSLSPE GTLTVEAPMP KLATQSNEIT IPVTFESRAQ LGGRSCKIR
```

Protein spot 11

Enoyl Coenzyme A hydratase 1, peroxisomal [Homo sapiens]
Sequence Coverage: 7%

```
1 MAAGIVASRR LRDLLTRRLT GSNYPGLSIS LRLTGSSAQE AASGVALGEA
   51 PDHSYESLRV TSAQKHVLHV QLNRPNKRNA MNKVFWREMV ECFNKISRDA
  101 DCRAVVISGA GKMFTAGIDL MDMASDILQP KGDDVARISW YLRDIITRYQ
  151 ETFNVIERCP KPVIAAVHGG CIGGGVDLVT ACDIRYCAQD AFFQVKEVDV
  201 GLAADVGTLQ RLPKVIGNQS LVNELAFTAR KMMADEALGS GLVSRVFPDK
  251 EVMLDAALAL AAEISSKSPV AVQSTKVNLL YSRDHSVAES LNYVASWNMS
  301 MLQTQDLVKS VQATTENKEL KTVTFSKL
```

Protein spot 12

glial fibrillary acidic protein isoform 1 [Homo sapiens]
Sequence Coverage: 5%

```
1 MERRRITSAA RRSYVSSGEM MVGGLAPGRR LGPGTRLSLA RMPPPLPTRV
   51 DFSLAGALNA GFKETRASER AEMMELNDRF ASYIEKVRFL EQQNKALAAE
  101 LNQLRAKEPT KLADVYQAEL RELRLRLDQL TANSARLEVE RDNLAQDLAT
  151 VRQKLQDETN LRLEAENNLA AYRQEADEAT LARLDLERKI ESLEEEIRFL
  201 RKIHEEEVRE LQEQLARQQV HVELDVAKPD LTAALKEIRT QYEAMASSNM
  251 HEAEEWYRSK FADLTDAAAR NAELLRQAKH EANDYRRQLQ SLTCDLESLR
  301 GTNESLERQM REQEERHVRE AASYQEALAR LEEEGQSLKD EMARHLQEYQ
  351 DLLNVKLALD IEIATYRKLL EGEENRITIP VQTFSNLQIR ETSLDTKSVS
  401 EGHLKRNIVV KTVEMRDGEV IKESKQEHKD VM
```

Protein spot 13

prosomal protein P30-33K

Sequence Coverage: 11%

```
1 MQLSKVKFRN QYDNDVTVWT AQGRIHQIEY AMEAVKQGSA TVGLKSKTHA
   51 VLVALKRAQS ELAAHQKKIL HVDNHIGISI AGLTADARLL CNFMRQECLD
  101 SRFVFDRPLP VSRLVSLIGS KTQIPTQRYG RRPYGVGLLI AGYDDMGPHI
  151 FQTCPSANYF DCRAMSIGAR SQSARTYLER HMSEFMECNL NELVKHGLRA
  201 LRETLPAEQD LTTKNVSIGI VGKDLEFTIY DDDDVSPFLE GLEERPQRKA
  251 QPAQPADEPA EKADEPMEH
```

Protein spot 14

heterogeneous nuclear ribonucleoprotein D-like [Homo sapiens]

Sequence Coverage: 7%

```
1 MEVPPRLSHV PPPLFPSAPA TLASRSLSHW RPRPPRQLAP LLPSLAPSSA
   51 RQGARRAQRH VTAQQPSRLA GGAAIKGGRR RRPDLFRRHF KSSSIQRSAA
  101 AAAATRTARQ HPPADSSVTM EDMNEYSNIE EFAEGSKINA SKNQQDDGKM
  151 FIGGLSWDTS KKDLTEYLSR FGEVVDCTIK TDPVTGRSRG FGFVLFKDAA
  201 SVDKVLELKE HKLDGKLIDP KRAKALKGKE PPKKVFVGGL SPDTSEEQIK
  251 EYFGAFGEIE NIELPMDTKT NERRGFCFIT YTDEEPVKKL LESRYHQIGS
  301 GKCEIKVAQP KEVYRQQQQQ QKGGRGAAAG GRGGTRGRGR GQGQNWNQGF
  351 NNYYDQGYGN YNSAYGGDQN YSGYGGYDYT GYNYGNYGYG QGYADYSGQQ
  401 STYGKASRGG GNHQNNYQPY
```

Protein spot 15

heterogeneous nuclear ribonucleoprotein A2/B1 isoform A2 [Homo sapiens] Sequence Coverage: 12%

```
1 MEREKEQFRK LFIGGLSFET TEESLRNYYE QWGKLTDCVV MRDPASKRSR
   51 GFGFVTFSSM AEVDAAMAAR PHSIDGRVVE PKRAVAREES GKPGAHVTVK
  101 KLFVGGIKED TEEHHLRDYF EEYGKIDTIE IITDRQSGKK RGFGFVTFDD
  151 HDPVDKIVLQ KYHTINGHNA EVRKALSRQE MQEVQSSRSG RGGNFGFGDS
  201 RGGGGNFGPG PGSNFRGGSD GYGSGRGFGD GYNGYGGGPG GGNFGGSPGY
  251 GGGRGGYGGG GPGYGNQGGG YGGGYDNYGG GNYGSGNYND FGNYNQQPSN
  301 YGPMKSGNFG GSRNMGGPYG GGNYGPGGSG GSGGYGGRSR Y
```

Protein spot 16

chaperonin containing TCP1, subunit 2 [Homo sapiens]

Sequence Coverage: 11%

```
1 MASLSLAPVN IFKAGADEER AETARLTSFI GAIAIGDLVK STLGPKGMDK
   51 ILLSSGRDAS LMVTNDGATI LKNIGVDNPA AKVLVDMSRV QDDEVGDGTT
  101 SVTVLAAELL REAESLIAKK IHPQTIIAGW REATKAAREA LLSSAVDHGS
  151 DEVKFRQDLM NIAGTTLSSK LLTHHKDHFT KLAVEAVLRL KGSGNLEAIH
  201 IIKKLGGSLA DSYLDEGFLL DKKIGVNQPK RIENAKILIA NTGMDTDKIK
  251 IFGSRVRVDS TAKVAEIEHA EKEKMKEKVE RILKHGINCF INRQLIYNYP
  301 EQLFGAAGVM AIEHADFAGV ERLALVTGGE IASTFDHPEL VKLGSCKLIE
  351 EVMIGEDKLI HFSGVALGEA CTIVLRGATQ QILDEAERSL HDALCVLAQT
  401 VKDSRTVYGG GCSEMLMAHA VTQLANRTPG KEAVAMESYA KALRMLPTII
  451 ADNAGYDSAD LVAQLRAAHS EGNTTAGLDM REGTIGDMAI LGITESFQVK
  501 RQVLLSAAEA AEVILRVDNI IKAAPRKRVP DHHPC
```

Protein spot 17

proteasome beta 7 subunit proprotein [Homo sapiens]
Sequence Coverage: 7%

```
1 MAAVSVYAPP VGGFSFDNCR RNAVLEADFA KRGYKLPKVR KTGTTIAGVV
   51 YKDGIVLGAD TRATEGMVVA DKNCSKIHFI SPNIYCCGAG TAADTDMTTQ
  101 LISSNLELHS LSTGRLPRVV TANRMLKQML FRYQGYIGAA LVLGGVDVTG
  151 PHLYSIYPHG STDKLPYVTM GSGSLAAMAV FEDKFRPDME EEEAKNLVSE
  201 AIAAGIFNDL GSGSNIDLCV ISKNKLDFLR PYTVPNKKGT RLGRYRCEKG
  251 TTAVLTEKIT PLEIEVLEET VQTMDTS
```

Protein spot 18

Nm23 protein [Homo sapiens]
Sequence Coverage: 12%

```
1 CCEPRGSRAR FGCWRLQPEF KPKQLEGTMA NCERTFIAIK PDGVQRGLVG
   51 EIIKRFEQKG FRLVGLKFMQ ASEDLLKEHY VDLKDRPFFA GLVKYMHSGP
  101 VVAMVWEGLN VVKTGRVMLG ETNPADSKPG TIRGDFCIQV GRNIIHGSDS
  151 VESAEKEIGL WFHPEELVDY TSCAQNWIYE
```

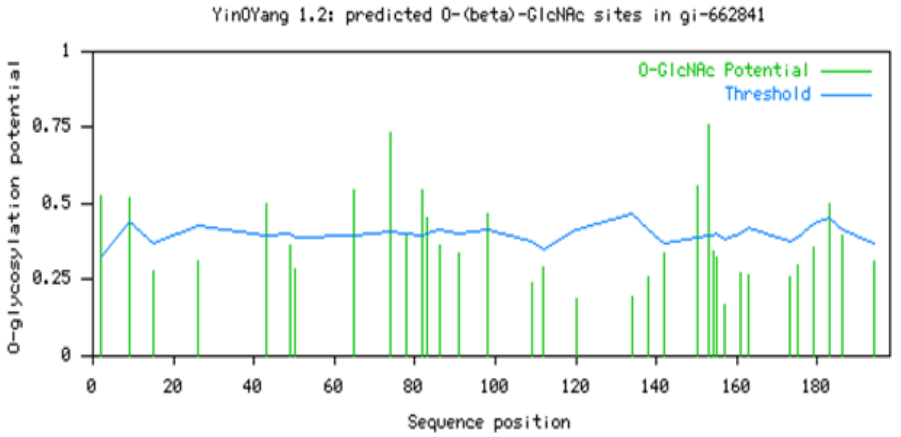
Appendix 4

Number of potential glycosylation and phosphorylation sites within HPA binding glycoproteins

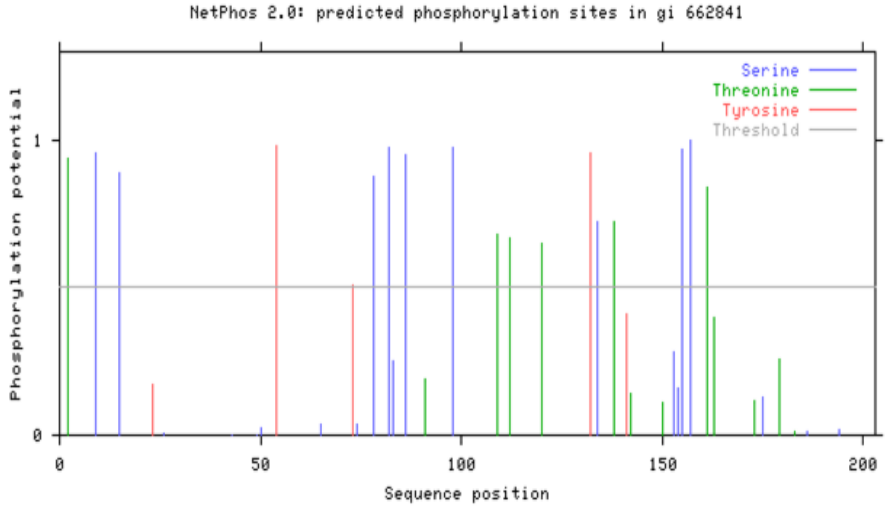
	Number of potential glycosylation and phosphorylation sites			
	N-Linked	**O-GalNAC**	**O-GlcNAC**	**O-Phosphate**
HnRNP-D-like	0	4	12	22
GFAP	0	0	4	27
Hsp 27	0	8	10	12
ENO I	0	0	4	16
HnRNP-HI	0	0	6	23
HnRNP-A2/B1	0	0	4	24

Appendix 5

Number of potential glycosylation and phosphorylation sites within Hsp27



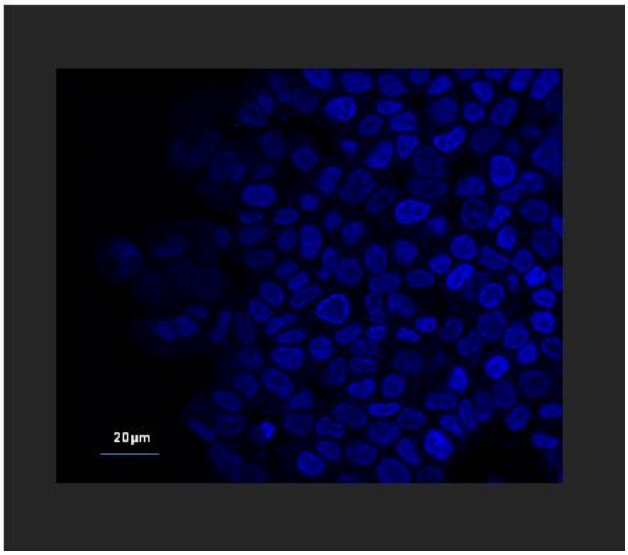
10 sites



12 sites

Appendix 6

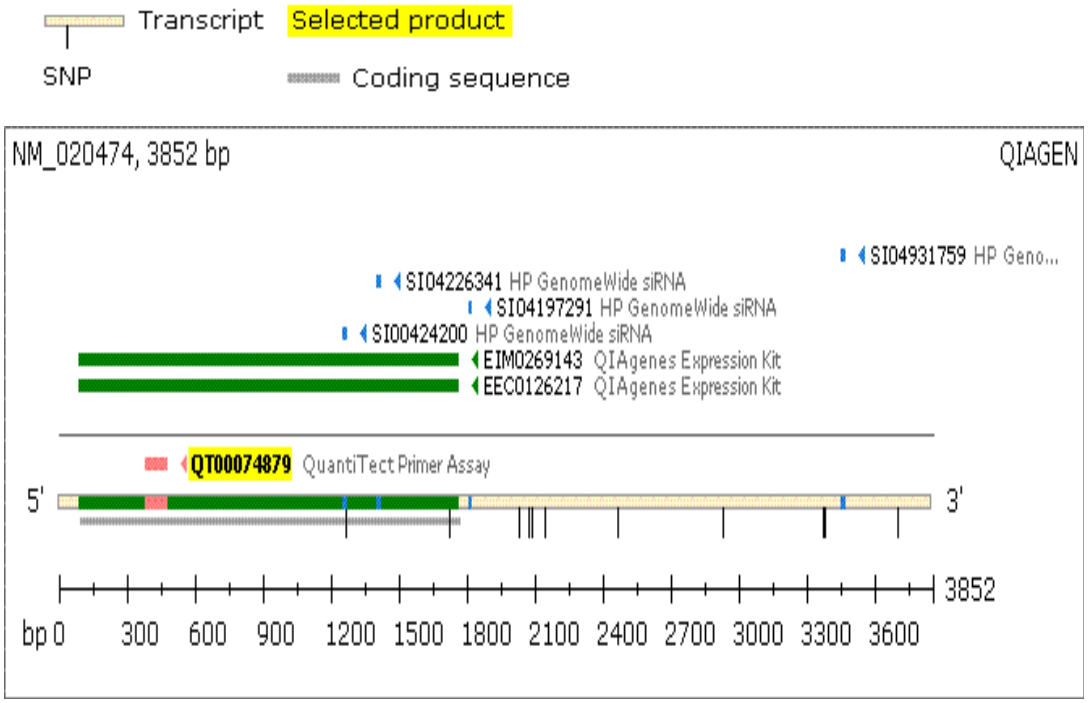
Negative control for O-glcNAc staining; staining of IgG-FITC only



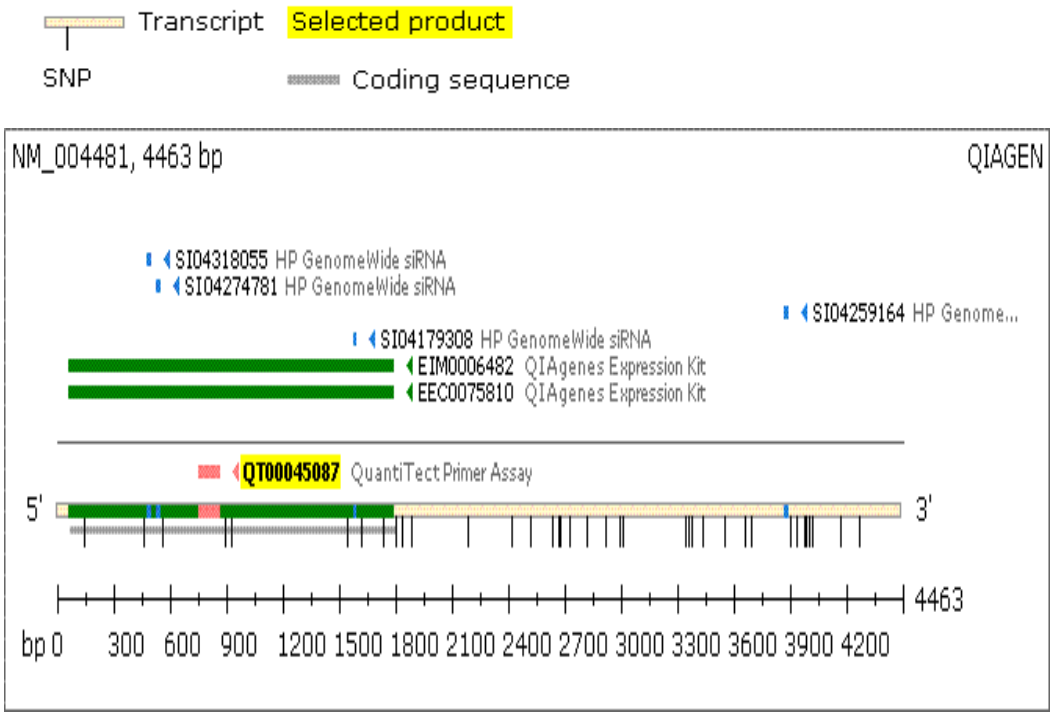
Appendix 7

Target sequences for ppGalNAc Ts

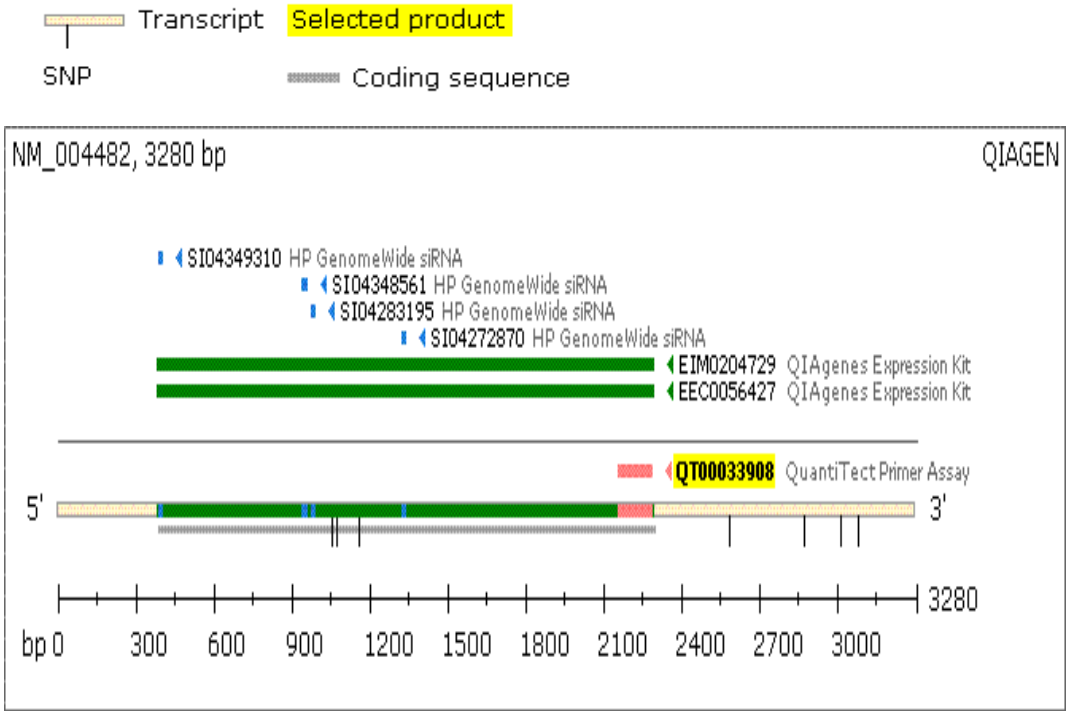
GalNAc T 1
Hs_GALNT1_1_SG QuantiTect Primer Assay (200) (QT00074879)



GalNAc T 2
Hs_GALNT2_1_SG QuantiTect Primer Assay (200) (QT00045087)

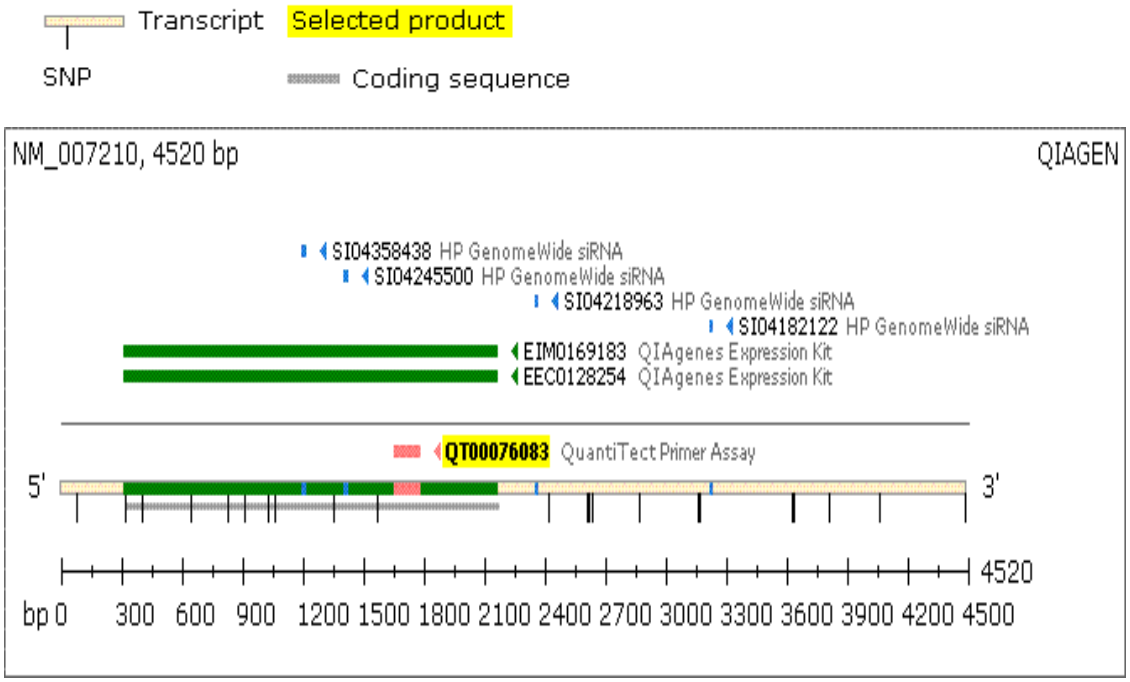


GalNAc T 3
Hs_GALNT3_1_SG QuantiTect Primer Assay (200) (QT00033908)



GalNAc T 6

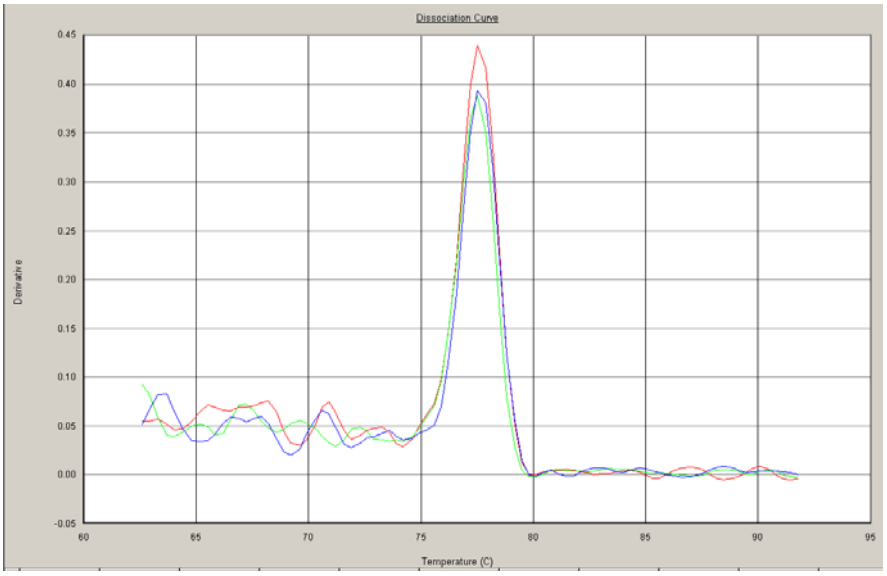
Hs_GALNT6_1_SG QuantiTect Primer Assay (200) (QT00076083)



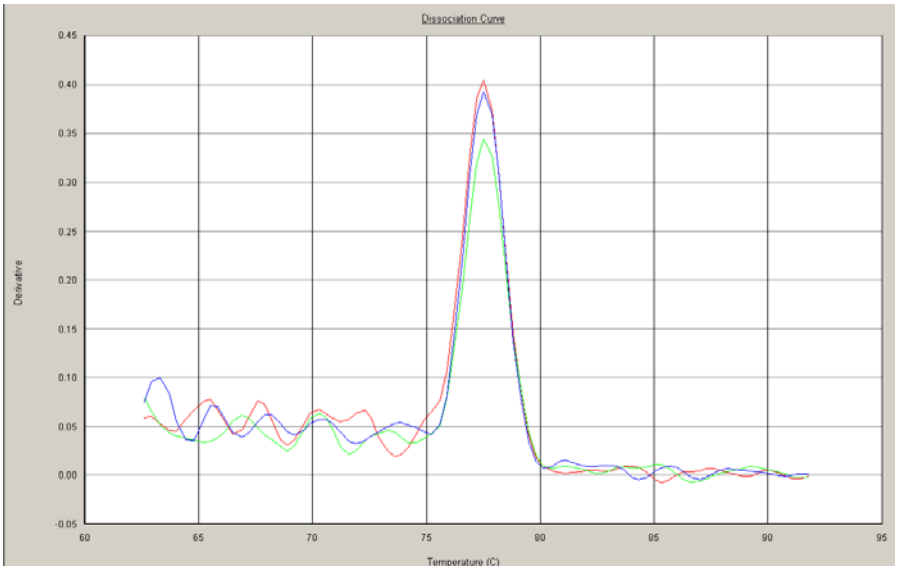
Appendix 8

Representative melting curves of *beta-actin*, ppGalNAc *Ts* and ST6GalNA*cs* genes in T47D cell lines

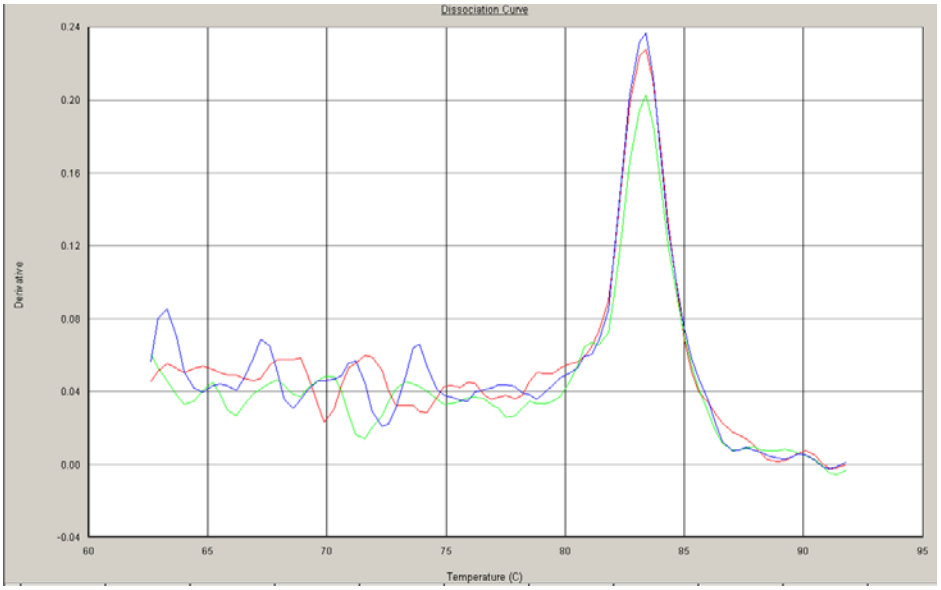
beta-actin



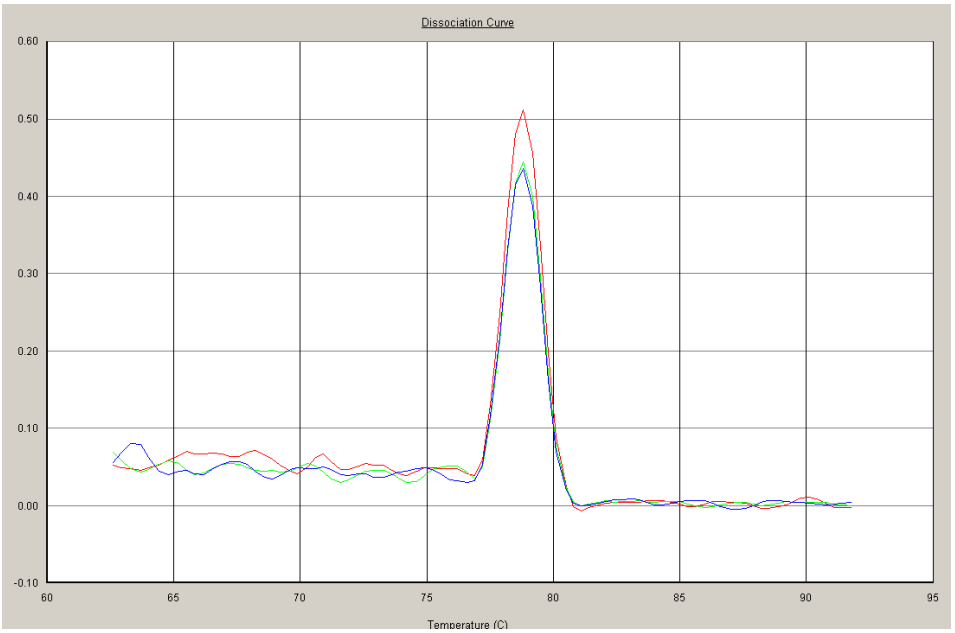
ppGalNAc T1



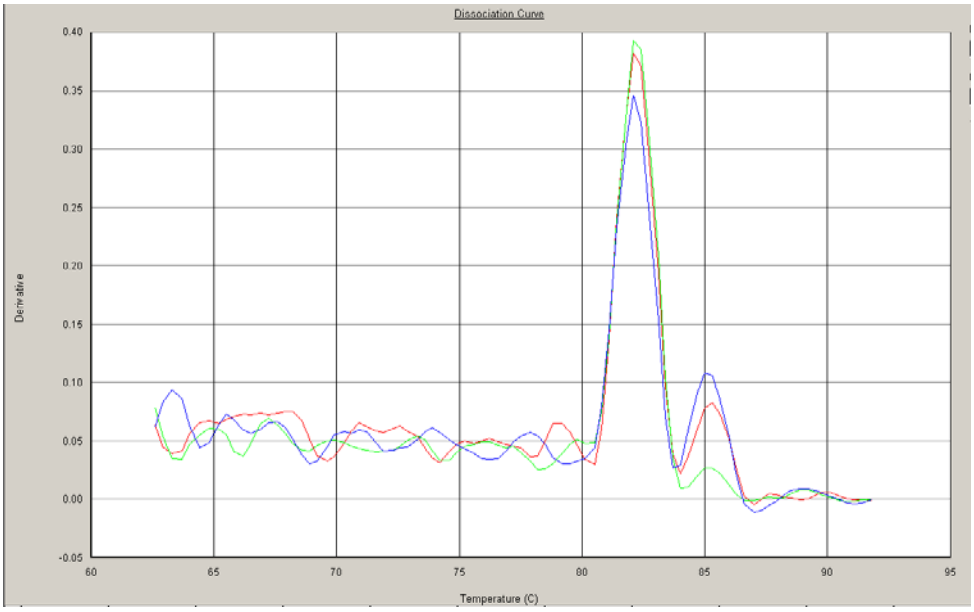
ppGalNAc T2



ppGalNAc T3



ppGalNAc T6



ST6GalNAc II