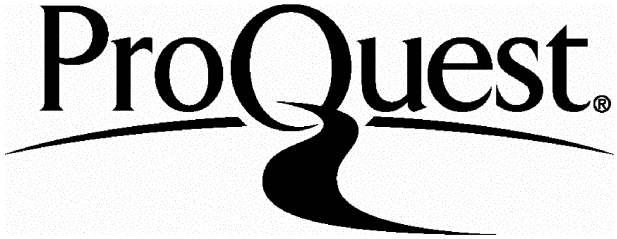
PROTEOMIC ANALYSIS OF ErbB-2 OVEREXPRESSION IN HUMAN MAMMARY LUMINAL EPITHELIAL CELL LINES

Severine GHARBI

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY
UNIVERSITY COLLEGE LONDON
GOWER STREET, LONDON

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy from the University of London



ABSTRACT

In the fight against cancer, constant effort is being made to find new markers of the genetic alterations associated with tumourigenesis in order to generate targeted therapies. An important molecular marker of breast cancers is ErbB-2. This receptor tyrosine kinase is found to be over expressed in 20 to 30% of human breast cancers and is now a target for breast cancer therapy. Understanding the role of ErbB-2 and other members of the ErbB receptor family has been the subject of intense effort and much is now known about the signalling pathways activated downstream of the ErbB family members. Despite these efforts, the mechanisms involved in ErbB-2-mediated tumourigenesis are still not fully understood.

In order to further understand these mechanisms at the level of protein expression, a global proteomic analysis of ErbB-2 overexpression was carried out using a newly developed gel-based technique, two-dimensional difference gel electrophoresis (2D-DIGE). This approach, which uses differential covalent labelling of protein lysates with fluorphores, was developed and further optimised for sensitive, reproducible and quantitative comparison of protein expression in multiple samples. The differential protein expression occurring in response to ErbB-2 overexpression and growth factor treatment was then examined in a model human mammary luminal epithelial cell system. 2D-DIGE analysis and mass spectrometry were combined for the identification of molecular markers of ErbB-2 overexpression, some of which have been previously implicated in transformation or found to be up-regulated in other types of cancer. Protein identities as well as changes in expression were validated by immunoblotting. Subsequently, several of these proteins were further characterised and their connection with ErbB-2 overexpression investigated.

Thus this study showed that 2D-DIGE analysis combined with mass spectrometry represents a powerful approach for the sensitive, reproducible and accurate identification of targets of ErbB-2-mediated cellular transformation. Together with post-genomic studies, this work has provided novel and important information on the regulation of the molecular processes involved in breast carcinogenesis.

ACKNOWLEDGEMENTS

There are many people I would like to acknowledge for the realisation of this thesis.
I would first like to thank my supervisor, Pr. M.D. Waterfield, for providing me with this great opportunity to carry out this work in the Cancer Proteomics laboratory, for trusting me and giving me confidence in my work.
I am extremely grateful of the tremendous help and day-to-day supervision provided by Dr. John Timms, guiding me throughout my experimental work and giving me support during the elaboration of this thesis. I thank his extreme patience and help.
I would also like to thank Dr. Rainer Cramer (and the Bioanalytical Chemistry group) and Dr. Soren Naaby Hansen for their instruction and expertise both in mass spectrometry and 2D-gel-based proteomics respectively.

Many members of the laboratory have been there for me when I needed them the most. I would like to thank most specially Dr. Sarah White for her moral support, but also my partner in writing, Sarah Hart, for treating me with goodies. Furthermore, there is no productive work without a good group of friends to keep a happy and enjoyable atmosphere at work and I would thus like to thank all of you in the Ludwig institute present and gone for having been there, for sharing this experience of my life with you. Just to name some of you amongst all the others Emma, Clemens, Craig, Sarah B, Pedro, Mark, Akunna, Mariana, Karin, Hong-Lin, Bertran, Steve, Malcolm, Denis, Carol, Marketa and Alice...

However, most of all, I would like to give a big thank you to Jaime. You cannot imagine the major impact you had at this stage of my life, you have given me strength and courage to carry this thesis to an end not only for your great mind and extremely useful discussions but also for your needed support and care.
Finally, I would like to thank my parents and Akiko for their support and for the unconditional trust they have given me throughout the years.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

CHAPTER 5

CHAPTER 6

ABBREVIATIONS

Chemicals and reagents

AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride
CBB	coomassie brilliant blue
CHAPS	(3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate)
CHCA	α-cyano-4-hydroxycinnamic acid
Cy2	3-(4-carboxymethyl)phenylmethyl)-3'-ethyloxacarbocyanine halide *N*-hydroxy-succinimidyl ester
Cy3	1-(5-carboxypentyl)-1'-propylindocarbocyanine halide *N*-hydroxysuccinimidyl ester
Cy5	1-(5-carboxypentyl)-1'-methylindodicarbocyanine halide *N*-hydroxysuccinimidyl
ddH2O	double deionised water
DHB	2,5-dihydroxybenzoic acid
DMSO	dimethyl sulfoxide
DTT	ditthiothreitol
EDTA	ethylenediaminetetraacetic acid
FCS	foetal calf serum
HEPES	N-[2-hydroxyethyl]piperazine-N'[2-ethanesulphonic acid]
IA	iodoacetamide
NaCl	sodium chloride
NHS	N-hydroxy succinimidyl
PBS	phosphate-buffered saline
SDS	sodium dodecyl sulphate
TBS	Tris buffered saline
TFA	trifluoroacetic acid
Tris	Tris-(hydroxymethyl)aminomethane

Software and instrumentation

BVA	biological variation analysis
CCD	charge-coupled device

DIA	difference in-gel analysis
ESI	electrospray ionisation
FACS	fluorescence activated cell sorting
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MALDI	matrix assisted laser desorption/ionisation
TOF	time of flight
General	
1DE	one-dimensional gel electrophoresis
2DE	two-dimensional gel electrophoresis
Ab	antibody (mAb: monoclonal Ab; pAb: polyclonal Ab)
DIGE	difference gel electrophoresis
DNA	deoxyribonucleic acid
ECL	enhanced chemiluminescence
FDA	federal drug agency
FITC	fluorescein isothiocyanate
HMLEC	human mammary luminal epithelial cell
HPLC	high performance liquid chromatography
HRP	horseradish peroxidase
IEF	isoelectrofocusing
IP	immunoprecipitation
IPG	immobilised pH gradient
kDa	kilodalton
kVh	kilovolt hours
m/z	mass-to-charge ratio
PAGE	polyacrylamide gel electrophoresis
PEM	protein expression map
ppm	parts per million
PVDF	polyvinylidene fluoride
RNA	ribonucleic acid

SD	standard deviation
TRITC	tetramethylrhodamine isothiocyanate
Vhr	volt hours

Enzymes, growth factors and inhibitors

AP-1	activated protein-1
BSA	bovine serum albumin
BpVphen	potassium bisperoxo(1,10-phenanthroline)oxovanadate
CIII	copine III
CHX	cycloheximide
CK17	cytokeratin 17
CPS1	carbamoyl-phosphate synthetase-1
EGF	epidermal growth factor
EGFR	epidermal growth factor receptor
ErbB-2	erythroblastosis B-2
ErbB-3	erythroblastosis B-3
ErbB-4	erythroblastosis B-4
hTERT	human telomerase reverse transcriptase
Hrgβ1	heregulin beta 1
IFN	interferon
LT	SV40 large T antigen
MAPK	mitogen activated protein kinase
MBP	myelin basic protein
MnSOD	manganese superoxyde dismutase
MVB	multivesicular body
MxA	interferon induced resistance protein
PGP9.5	protein gene product 9.5 (also known as UCH-L1)
PI3K	phosphatidylinositol 3-kinase
PTK	protein tyrosine kinase
SV40	simian virus 40
UCH-L1	ubiquitin carboxy terminal hydrolase-L1 (e.g. PGP9.5)

Chapter 1 : INTRODUCTION

1.1 Cancer cell signalling

In multi-cellular organisms, cells need to interact with each other in order to control their growth, differentiation, survival and perform diverse physiological functions. Cells emit and receive signals dictating their state of growth either by direct cell-to-cell interaction or via secreted molecules. These signalling molecules activate receptors embedded in the cell membrane, leading to the activation of transduction cascades which can ultimately lead to gene activation or repression. According to the specificity, strength and duration of the molecular signals received, the cell will proliferate, differentiate, enter into growth arrest, or undergo apoptosis. Signalling pathways are of great complexity and although major routes have been deciphered, the spatio-temporal organisation and components of intracellular machinery are poorly understood (reviewed in Hunter 2000). A multitude of proteins are involved in specific signalling events, however, the picture can be simplified by the classification of their components according to their protein structure or functional domain content. The major events governing signal transduction involve cascades of protein-protein or protein-lipid interactions regulated by post-translational modifications such as phosphorylation. Many protein interaction domains have been defined such as the SH2, PTB, SH3, LIM, and PDZ domains which are implicated in diverse signalling pathways (Pawson 1995).

This complex machinery is highly regulated although alteration of the normal intracellular signals can lead to the development of diseases such as cancer. It is now acknowledged that a succession of several genetic changes is required for the progressive conversion of normal human cells into cancerous cells and signalling molecules are often implicated in this process. Hanahan and Weinberg have proposed a model of tumourigenesis, whereby several physiological conditions are attained by cells as they become tumourigenic (Hanahan and Weinberg 2000). These "hallmarks" of cancer are summarised as self-sufficiency in growth signals, insensitivity of cells to growth-inhibitory signals, capability to evade programmed cell death, limitless replicative

potential, increased angiogenesis and ability to invade tissues and metastasise. Many components of the signalling pathways involved in these processes have been shown to be deregulated in cancers. In proliferative signalling pathways alone, numerous proto-oncogenes or tumour suppressors have been identified which cause amplification or ablation of signalling, over-proliferation and eventual tumour formation. However, although several markers of the disease are already well characterised, there is still a need to develop targeted therapies. A single protein, although of high importance does not reflect the state of a cell as it relies on the interaction with many other cellular components. Attempting to target one effector of a pathway needs to be done with great care as it could disrupt other vital processes in a cell. In order to better develop such treatments, it would thus be necessary to identify all the players involved in tumour progression and to fully understand the function of these target proteins.

1.2 The ErbB receptor tyrosine kinase family

One of the major mechanisms used to mediate transmembrane signalling involves tyrosine phosphorylation of receptors to generate binding motifs. A large family of receptor tyrosine kinases have been identified and characterised (Schlessinger 2000). These receptors can be classified into 14 different sub-groups based on their structural organisation. The main feature that they share, with the exception of the insulin receptor, is their single transmembrane domain which consists of a hydrophobic helix that connects an extracellular ligand binding domain and an intracellular kinase domain. Upon specific ligand binding, activation of their intracellular tyrosine kinase domain triggers tyrosine phosphorylation which leads to a complex succession of downstream signalling events, ultimately resulting in diverse cellular processes such as cellular proliferation, differentiation, cell motility and transcriptional activation. These receptors also play a fundamental role in development, as shown by gene knock-out studies (Lee *et al.* 1995; Chan *et al.* 2002). Many of these receptors have been previously implicated in the development of major diseases and particularly, a significant number of receptor tyrosine kinases (RTKs) have been identified as oncogenes and are implicated in various human cancers (Schlessinger 2000; Zwick *et al.* 2002).

One of the best-studied sub-families of RTKs is the epidermal growth factor receptor (EGFR) family. There are four members in this family, EGFR (also known as ErbB-1), ErbB-2 (HER2), ErbB-3 and ErbB-4. These receptors have two cysteine-rich extracellular ligand binding domains, a single transmembrane domain and an intracellular domain consisting of the tyrosine kinase domain and a carboxy-terminal tail which possesses functional tyrosine phosphorylation sites. The first event leading to receptor activation is the binding of a specific ligand to the extra-cellular domain of the receptor. This event subsequently triggers the formation of dimers and the trans-activation of intracellular tyrosine kinase activity, leading to the autophosphorylation of specific tyrosine residues in the C-terminal tail (reviewed by Weiss and Schlessinger 1998). These phosphorylated residues then act as recruitment sites for specific adaptor proteins and signalling molecules. Despite their similarity, each member of the ErbB family can trigger the activation of distinct, but overlapping patterns of downstream signalling through the recruitment of different subsets of signalling molecules (Hackel *et al.* 1999). In addition, each receptor can bind a specific subset of a large family of ligands (reviewed by Peles and Yarden 1993; Riese and Stern 1998). These are the epidermal growth factor (EGF)-related peptides, which fall into three major groups according to receptor specificity. EGF, amphiregulin (AR) and transforming growth factor-α, all bind to EGFR (ErbB-1); betacellulin, epiregulin and heparin binding EGF-like growth factor, have the potential to bind to both EGFR and ErbB-4 receptors, and the Neu differentiation factors or heregulins (Hrg) which are specific ligands for ErbB-3 and ErbB-4 receptors. The ErbB-2 receptor is the only receptor for which no specific ligand is known.

Each member has the potential to be activated by homodimerisation upon ligand binding, but heterodimers with other members of the ErbB family are also formed. It has been shown that the dimerisation partner of each family member can dictate the specificity of the signalling activated downstream (Olayioye *et al.* 1998). Importantly, although ErbB-2 receptor does not itself have a specific ligand, it is thought to be the preferred heterodimerisation partner of the other ErbB family members (Tzahar *et al.* 1996; Graus-Porta *et al.* 1997). Despite this, ErbB-2 has also been reported to form

homodimers in ErbB-2 overexpressing cell systems (Lindberg *et al.* 2002). In addition, ErbB-2 heterodimerisation enables the ErbB-3 receptor to be activated, since ErbB-3 homodimers cannot signal due to a lack of intrinsic tyrosine kinase activity (Carraway *et al.* 1994; Guy *et al.* 1994; Kim *et al.* 1998). Thus, for one of the ErbB family members to be activated, a combination of events such as the presence of specific ligands and the availability of a co-receptor for homo- or heterodimer formation is necessary.

As mentioned, each receptor possesses distinct multiple sites of tyrosine phosphorylation mainly localised in the C-terminal tail, which act as consensus motifs for specific binding of effector proteins and initiation of signalling (Figure 1.1). Several studies have been used to map the consensus tyrosine phosphorylation sites of the intracellular cytoplasmic tail of each ErbB family member and to identify the substrates which are recruited at specific phosphotyrosine residues (reviewed by Weiss and Schlessinger 1998; Olayioye *et al.* 2000).

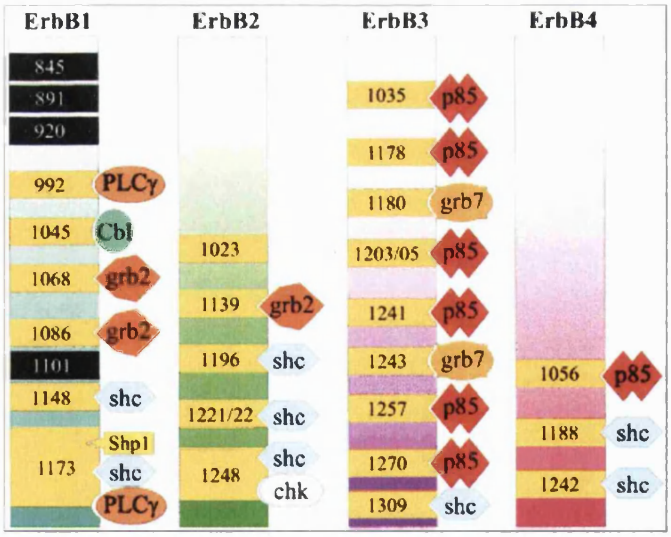


Figure 1.1: Summary of known ErbB family consensus phophotyrosine residues and recruitment of signalling molecules reported (review Olayioye *et al.* 2000). Autophosphorylation sites are shown in yellow and Src kinase phosphorylation sites are shown in black.

Three main classes of substrates are recruited by activated ErbB receptors: adaptor proteins, including Shc, Grb2, Grb7 and Gab1; kinases, such as phosphatidylinositol 3-kinase (PI3K) through its p85 subunit, Src and Chk and finally the phosphatases, Shp1 and Shp2. All ErbB receptors possess at least one docking site for the recruitment of Shc, which binds via its protein tyrosine binding (PTB) domain or its src homology-2 (SH2) domain. However, other effector proteins are more specific and for instance ErbB-1 (EGFR) is the only ErbB member able to recruit c-Cbl or Shp1. In addition, ErbB-3 possesses six known phosphorylation sites able to recruit the regulatory subunit of PI3K, p85, and is therefore thought to be the best activator of downstream PI3K signalling (Cantley *et al.* 1991; Songyang *et al.* 1993).

This diversity in recruitment together with the preferential heterodimerisation reveals the complexity and diversity in signalling that is generated through the activation of this family of receptor tyrosine kinases. As shown in Figure 1.2, the recruitment and activation of effector proteins is an important step in the initiation of various downstream signalling cascades. These include the mitogen-activated protein kinase (MAPK) and the PI3K pathways according to the ErbB dimerisation profile (Neve *et al.* 2002). The MAPK pathway is activated downstream of the ErbB family members via the recruitment of Shc, Grb2 and SOS. This pathway consists of a ras, raf, Mek, MAPK and Rsk signalling cascade, which thus extends from the membrane to the nucleus and leads to the activation of downstream transcription factors namely Elk-1/p62TCF, c-Myc, C/EBPβ (Johnson *et al.* 1996; Olayioye *et al.* 2001). The PI3K pathway is also activated downstream of ErbB-3/ErbB-2 heterodimer formation thus leading to further downstream activation of the Akt/PKB and S6 kinase cascades (Vanhaesebroeck *et al.* 1997). Activation of these signalling cascades leads to transcriptional activation of specific genes via transcription factors such as Myc, Jun and Fos or the STATs. Such gene expression can be required for proliferation, apoptosis, differentiation and cellular migration or adhesion. Furthermore, each signalling cascade is not isolated, but can also interact with other pathways (such as the cytokine signalling pathway) and the amplitude and duration of the signals transmitted will dictate the quality of the output signal.

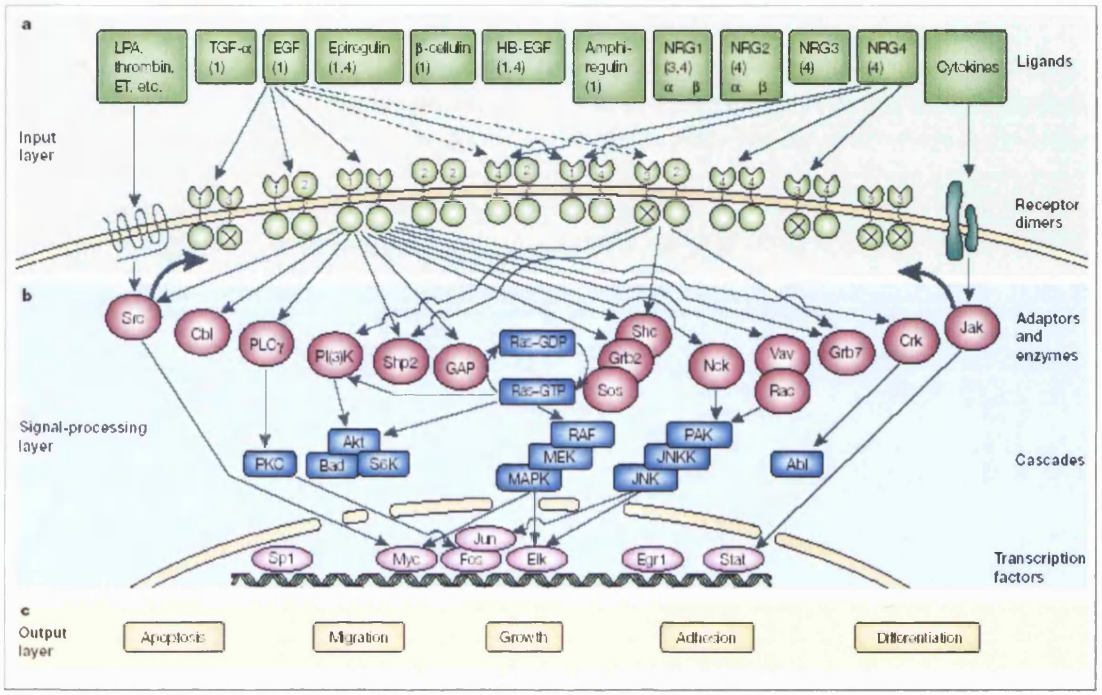


Figure 1.2: The complexity of the ErbB signalling network (from Yarden and Sliwkowski 2001). This figure summarises the complexity of the ErbB signalling network. It can be described as several layers of activation. From an array of specific ligands, ErbB receptors are activated in combinations of dimers subsequently leading to the recruitment of various specific and overlapping effector proteins. Kinase activation enables further cascade activation of signalling molecules. Amongst those, the best studied pathways activated downstream of the ErbB family members are the mitogen-activated protein kinase (MAPK) pathway via the recruitment of Shc, Grb2 and Sos and the phosphatidylinositol 3-kinase (PI3K) pathway leading to the cascade activation of Akt and the S6Kinase. The next layer of this cascade activation results in the activation of specific transcription factors such as Myc, activated protein-1 (AP1) (Jun, Fos) or Elk leading to gene expression.

In addition to this array of signalling events, ErbB receptors are rapidly down-regulated, which is a mechanism employed by the cell to switch off the signal. Firstly, this involves the recruitment of protein phosphatases, such as PTP-1B (Ostman and Bohmer 2001; Haj *et al.* 2003). These play an important role in the maintenance of ligand-independent receptor phosphorylation, but also in the dephosphorylation of ligand-

induced receptor activation. This was revealed by inhibition or knock-out of PTP-1B (PTP-1B$^{-/-}$) leading to the sustained activation of RTKs (Haj *et al.* 2003). In addition, receptor down-regulation occurs by rapid endocytosis via coated-pits and subsequent intracellular degradation (Sorkin and Waters 1993; Sorkina *et al.* 2002). This process appears to be strictly controlled and varies between each ErbB receptor. EGFR was shown to be endocytosed and degraded via the ubiquitin-proteasome pathway, due to its ability to recruit c-Cbl, a ubiquitin E3 ligase (Levkowitz *et al.* 1998; Levkowitz *et al.* 1999). However, this mechanism is not shared by other ErbB members which are thought to have impaired ligand-mediated endocytosis (Sorkin *et al.* 1993; Baulida *et al.* 1996; Levkowitz *et al.* 1999).

Due to their implication in different forms of cancer, both EGFR and ErbB-2 have been the focus of intense research. EGFR was the first ErbB receptor to be cloned (Ullrich *et al.* 1984) and it has been found overexpressed in several different types of cancers such as non-small cell lung carcinoma, bladder, cervical and ovarian cancers and in squamous cell carcinomas of the head and neck (Nicholson *et al.* 2001). Increased EGFR signalling due to overexpression is thought to play a causative role in tumourigenesis. Several tyrosine kinase inhibitors have been developed in order to specifically target this receptor for prospective therapies, such as the compound ZD1839 (Iressa) (Ciardiello 2000), or the monoclonal antibody directed against EGFR, cetuximab (C225) (Baselga 2001). These EGFR targeted therapies were designed to prevent activation of receptor and both are in Phase III clinical trials (Overholser *et al.* 2000).

Due to its pivotal role in the ErbB signalling network, ErbB-2 was shown to be a potent activator of downstream signalling events in response to a variety of EGF family ligands. Moreover, it is known to play a major role in the progression of some human cancers. Indeed, ErbB-2 has been reported to be an important marker in invasive breast cancer (Bange *et al.* 2001) and gene amplification and protein overexpression is observed in 10 to 30% of invasive breast cancers and in 40 to 60% of intraductal breast carcinomas (Slamon *et al.* 1987). ErbB-2 overexpression has been correlated with the reduced survival of patients with breast cancer and poor patient prognosis (Slamon *et al.* 1987). ErbB-2 protein expression status can dictate responsiveness to breast cancer therapies and

is therefore a useful diagnostic marker. The main therapeutic approach used in breast cancer involves the removal of the tumour tissue in combination with chemotherapeutic treatments and anti-metabolic drugs. However, such treatments can present adverse effects for the patient due to their high toxicity, and their efficacy and specificity are poor. Improving the selectivity of cancer therapy by sparing non-cancerous cells and specifically treating transformed cells therefore represents an important aim. New targeted therapies have been developed in order to reduce the adverse effects of drug treatment on the patients and to improve the outcome of the treatment. However, the molecular profile of all breast cancers is diverse and efficient therapies require a good understanding of the oncogenic characteristics of each cancer. For instance, ErbB-2 overexpressing breast cancer patients have shown poor responsiveness to conventional therapies such as endocrine therapy and chemotherapeutic agents (Bange *et al.* 2001). A neutralising humanised antibody targeted against ErbB-2, Herceptin (trastuzumab), has been developed and has shown good outcome in the reduction of the growth of ErbB-2 overexpressing breast tumours (Carter *et al.* 1992). Trastuzumab has been in Phase I-III clinical trials and the Federal Drug Agency (FDA) approved its clinical use in 1998. However, the defined mechanisms by which ErbB-2 down-regulation occurs following treatment with trastuzumab and leading to tumour growth inhibition are not well understood. Improved targeted therapies need to be developed and thus further understanding of the mechanism involved in the mediation of breast cancer by ErbB-2 overexpression needs to be assessed. For this purpose, the focus of the work presented in this thesis was to investigate the mechanisms involved in ErbB-2-mediated tumour formation in breast cancer.

1.3 Models in cancer research

In the search for new-targeted therapies against cancers, the use of cancer models can be advantageous as it enables the accurate and reproducible study of the various properties of human cancer cells. Although clinical cancer biopsies obtained from patients are more representative of the real state of the illness, such samples are usually very heterogeneous and also difficult to obtain in large quantities. The prerequisite of a

good model of cancer is that it has to mimic to the best extent the biological behaviour of an *in situ* cancer. This is not always evident as model cancer cell systems are also prone to genetic instability during their development, and often they lack the cellular or tissue environment which they are supposed to mimic. For a single disease, multiple models can be used and these can be either animal models or *in vitro* cellular models, and although no model is ideal, the most appropriate will match best the experimental question asked. Several models of breast cancer have been developed which are used to understand the events leading to the development of cancer. Many breast cancer cell lines such as SKBr3 BT474 and SUM cell lines, have been derived from pleural effusions of patients with metastatic disease and therefore are more representative of the final stages of the tumour. However, little is known about the earlier stages of tumour formation. Cell models have been generated from primary tumour materials, although the creation of these pure cell lines often leads to biased clonal selection of cells which does not reflect the original genetic profile of the primary tissue from which it arose.

Cell lines can also be derived by the manipulation of normal tissues and such models enable the understanding of the early events in cancer development. It is now well established that for a normal cell to be transformed into a cancerous cell, several genetic events need to occur (Hanahan and Weinberg 2000). Therefore, to efficiently convert normal primary cells into tumourigenic cells, the alteration of several oncogenes needs to be carried out (Land *et al.* 1983). For instance, the expression of the simian virus 40 (SV40) large T antigen (LT) can lead to the tumourigenic conversion of primary human epithelial and fibroblast cells in conjunction with ectopic activation of the telomerase catalytic subunit (*hTERT*) and an oncogenic allele of H-*ras* (Hahn *et al.* 1999). In the present study, a cellular model of breast cancer was chosen to investigate the role of ErbB-2 overexpression in cellular transformation. Stamps et al. developed an immortalised luminal epithelial cell line, HB4a, derived from reduction mammoplasty tissue of a normal subject (Stamps *et al.* 1994). Most human breast cancers arise from the luminal epithelium of human mammary gland (Bartek *et al.* 1991) and this cell type was therefore ideal for the study of the early stages of breast cancer. Indeed, the cells display several markers of the luminal epithelial cell type and have a characteristic luminal

epithelial cellular morphology in culture. These cells were immortalised with a temperature-sensitive mutant of SV40 large T antigen (ts A58-U19) (Jat and Sharp 1989). It is important to bear in mind that this process of immortalisation involves the sequestration and inactivation of the tumour suppressor proteins pRb and p53 by the SV40 large T antigen. These proteins normally down-regulate the cell cycle, but also act as sensors of cell damage to induce cell cycle arrest and apoptosis, given the appropriate stimuli.

In order to understand the role of ErbB-2 overexpression in transformation, HB4a cell lines were stably transfected with a human ErbB-2 cDNA construct. Clones were selected by fluorescence activated cell sorting (FACS) based on the levels of surface ErbB-2 expression they showed. A panel of cell lines were generated which overexpressed various levels of ErbB-2 (Harris *et al.* 1999). In this thesis, the parental cell line, HB4a and one of the ErbB-2 overexpressing clones, C3.6, were used in order to establish the differential protein expression profile between the cell lines related to ErbB-2 overexpression. In the original work, the C3.6 clone was shown to overexpress levels of surface ErbB-2 similar to the levels observed in human breast cancers (Harris *et al.* 1999). In addition, another clone, C5.2 was also characterised and was shown to express 10-fold higher levels of ErbB-2 than the C3.6 clone. These cell lines were further characterised and their transformed phenotype was investigated in relation to ErbB-2 overexpression (Harris *et al.* 1999). Compared to the HB4a, the C3.6 and C5.2 cells were shown to have an altered morphology indicating that they had lost cell-to-cell contact inhibition and acquired cytoskeletal changes characteristic of transformed cells. In addition, the ErbB-2 overexpressing cell lines were shown to display an increased proliferative capacity and showed enhanced growth responsiveness to EGF and Hrgβ1, growth factors that are specific for EGFR and ErbB-3 respectively. The C3.6 and C5.2 cells were also found to display an increase in anchorage-independent growth, as measured by their ability to form colonies on soft agar. However, these cells did not induce the formation of tumours after inoculation into nude or SCID mice, showing that they do not display the phenotype of a fully transformed cell line.

This model cellular system was first used for the study of the effect of ErbB-2

overexpression on mitogenic signalling and cell cycle progression (Timms *et al.* 2002). This work showed that the C3.6 and HB4a cells were able to activate downstream signalling via the MAPK and the PI3K signalling pathways in response to EGF and Hrgβ1 treatment. However, the ErbB-2 overexpressing cells had enhanced MAPK signalling, but not PI3K signalling. In addition, both EGF and Hrgβ1 could induce myc, cyclin D1, cdk6 and cyclin E, known activators of the cell cycle, and this was more prominent in the ErbB-2 overexpressing cells. The cell cycle inhibitor $p27^{kip}$ was also suppressed in the C3.6 cells, which has been previously shown in other models of ErbB-2 overexpression (Lane *et al.* 2000; Lenferink *et al.* 2001). These combined changes induce an increase in cdk2 activity that promotes more rapid progression through the cell cycle.

Together, this prior work showed that this cellular model of breast cancer is a valid model for the study of breast cancer and was therefore used in this study to examine the effects of ErbB-2 overexpression on the overall protein expression profile.

1.4 Cancer research and proteomic analysis

As previously mentioned, the development of cancer occurs through multiple steps and hence involves the deregulation of many cellular processes. Conventional analysis of signal transduction pathways usually involves the study of a targeted pathway, and the understanding of an altered phenotype. However, a new era of study is emerging, which involves the large-scale analysis of the global protein profile, or proteome, of a cell type or tissue. Such an approach, under ideal conditions, is unbiased and potentially provides a comprehensive view on the protein expression profile of a cell, thus opening the field of cancer research to a new level of understanding.

1.4.1 Overview on proteomics

The understanding of the proteome has evolved with the ever-increasing knowledge of the genome, and more specifically the human genome. In addition, more accurate high-throughput technologies have made possible the development of approaches with high potential for global protein analysis, enabling the study of complex events in a single step as opposed to conventional targeted approaches. Despite the fact

that the main techniques used in proteomic analysis were developed in the early 1970s (Kenrick and Margolis 1970), this new terminology, proteomics, was first mentioned in 1994 at a "2-D Electrophoresis" Symposium held in Siena, Italy. The study of the proteome represents the "large-scale" analysis of proteins expressed in a genome. More specifically, differential proteome analysis is the comparison of protein expression profiles, for example of seperated proteins between two or more samples on a global scale. It was mainly in the 1990s, with the emergence of high-sensitivity analytical chemistry for protein identification using mass spectrometry, that the field of proteomics became useful in biological studies. Major improvements in mass spectrometry have replaced the former method of choice for protein identification by Edman degradation, an approach which although very powerful, was time consuming and required high amounts of sample. The field of proteomics has rapidly evolved and the study of the proteome now includes an array of methods to evaluate not only changes in protein expression in a cell in response to stimuli or between biological samples, but also the examination of protein-protein interactions, activity status, and protein sub-cellular localisation. Although the main platform used for protein identification relies on the resolution of proteins by gel-based approaches and their further identification using mass spectrometry (MS) (Figure 1.3), new approaches and workflows are now emerging and make possible large-scale protein separation and identification using automated multi-dimensional liquid chromatography-MS, and high-throughput data analysis. These approaches will be described below.

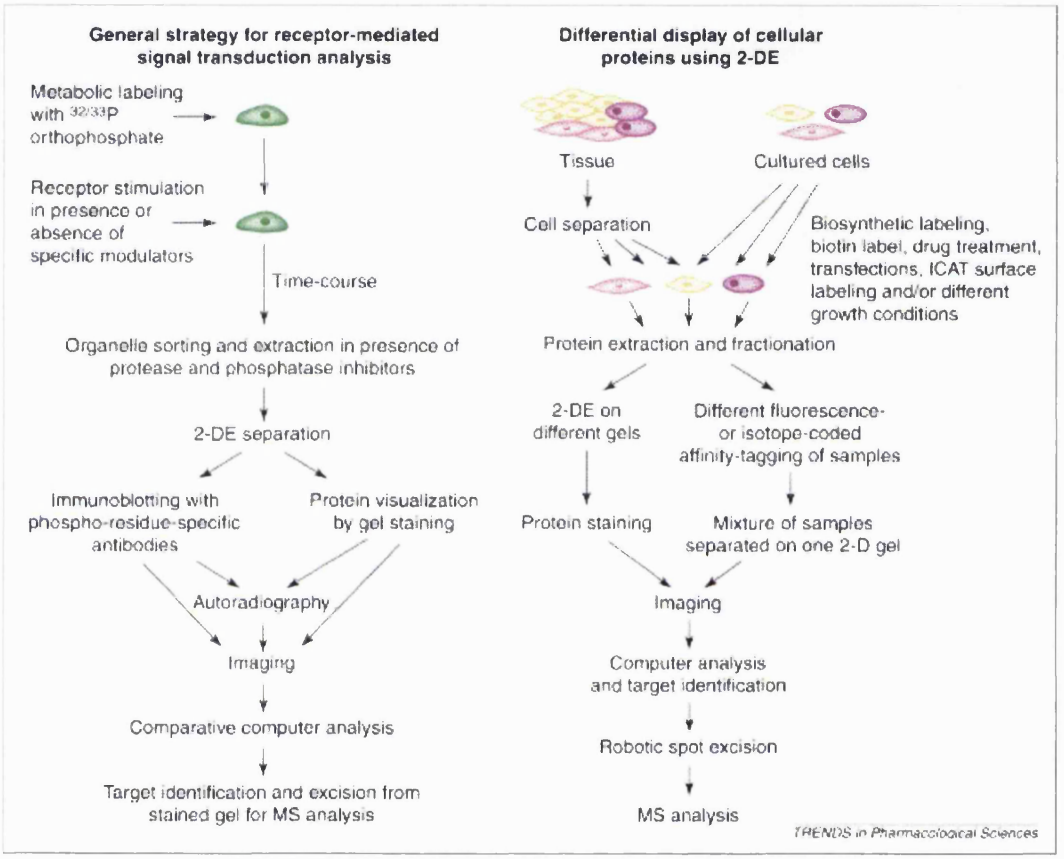


Figure 1.3: Overview on 2DE-based proteomic strategies to study signal transduction pathways (from Naaby-Hansen *et al.* 2001). Several proteomic platforms can be developed to characterise complex signalling pathways, comparing several cell/tissue types under various growth conditions. Multiple gels can subsequently be compared using appropriate computer analysis softwares and putative biomarkers of interest can then be further analysed by mass spectrometry (MS).

1.4.1.1 Principles of two-dimensional gel electrophoresis

The main approach used in the study of the proteome is a combination of two-dimensional gel electrophoresis for protein separation with the detection of separated proteins and the use of mass spectrometry for protein identification. Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) was first developed in the early 1970s (Kenrick and Margolis 1970) and was optimised using denaturing isoelectrofocusing (Klose and Spielmann 1975; O'Farrell 1975; Scheele 1975). However, the use of 2D-PAGE was greatly improved with the development of immobilised pH gradient (IPG) gels, which gave increased protein resolution in the first dimension (Bjellqvist *et al.* 1982). Subsequently, pre-cast IPG strips were made commercially available, making 2D-PAGE a more reproducible technique which could be easily applied by many users. However, it was with the emergence of mass spectrometry that 2D-PAGE became more widely used as a proteomic tool. The main challenge of 2D-PAGE is to achieve high protein resolution with increased sensitivity in order to provide pure and abundant samples for further mass spectrometric analysis. In addition, for differential protein expression analysis, it is important to attain high reproducibility in protein separation from experiment to experiment.

In principle, two-dimensional gel electrophoresis (2DE) relies on the separation of proteins in the first dimension according to their charge and then in the second dimension according to their relative molecular weight. On a first dimension, proteins are separated by isoelectrofocusing (IEF) in a pH gradient and become focused when they acquire a zero net charge at their isoelectric point (pI). The proteins are then separated on a second dimension according to their mass by conventional SDS-PAGE. The electrophoretic separation of a protein is highly specific to its chemical properties and therefore it becomes possible to accurately compare proteins analysed on distinct 2DE gels. Although the use of IPG strips has considerably improved the consistency of data produced in different laboratories, the datasets generated often contain a lot of redundancy and it is now accepted that only a fraction of the proteome can be detected on a 2DE gel. It has become clear that in many biological comparisons, identification of components expressed at low levels cannot be achieved by conventional gel analysis methods, hence

the need to optimise the protein separation efficiency of 2DE gels and increase the yield in protein detection. Full descriptions of current methodologies and optimised techniques have been reported (Rabilloud *et al.* 1994; Gorg *et al.* 1999; Gorg *et al.* 2000). In order to reduce the complexity in protein samples and the problem of multiple proteins migrating at the same coordinates, workers have developed narrow range IPG strips (Gorg *et al.* 1999; Westbrook *et al.* 2001) or used larger gels (Klose 1999). These approaches also enabled the improved separation of larger amounts of protein sample on a gel. Another significant improvement has been in the solubilisation of proteins using detergents that are compatible with isoelectrofocusing and gel electrophoresis, providing a means of increasing sample loading (Rabilloud 1998). Sample application can also be optimised in order to improve the entry of proteins into the first dimension. Several methods have been reported, such as cup-loading which have made possible the analysis of up to 1 mg of proteins loaded on an IPG-strip (Hanash *et al.* 1991). Another drawback of protein separation by 2DE is the poor resolution of basic proteins during the isoelectrofocusing. These proteins represent a large fraction of the proteome and include membrane and nuclear proteins. This limitation could be due to inadequate buffering by ampholytes in this pH range and loss of reductive capacity resulting in protein precipitation following spurious disulfide bond reformation. Several optimisations have been made in order to improve the solubility of membrane proteins on a 2DE gel (Rabilloud 1994; Herbert 1999; Rabilloud *et al.* 1999). This can be achieved using stronger detergents or high pH range IPG strips to improve the solubility and resolution of the highly basic proteins. However, this means that several 2DE gels have to be run to separate the acidic and basic fractions of the proteome and all proteins cannot be separated on the same 2DE gel. Another very important area of improvement in 2DE gel-based proteomics has been the development of more sensitive and reproducible visualisation methods for gel-separated proteins as discussed below.

1.4.1.2 An array of methods for the detection of gel separated proteins

Following on from their electrophoretic separation, it is necessary to detect and quantitate proteins in order to select the proteins of interest for further analyses by mass

spectrometry. The challenge is to achieve the detection of the complete proteome, hence to visualise and accurately quantitate high as well as very low abundance proteins. In addition, the protein detection method has to match the requisite imposed by the separation method, as it should not affect protein solubility during the electrophoretic process. Importantly, the method of protein detection employed also needs to be compatible with the method of choice for downstream protein identification. Several reviews have reported the methods of choice for detection of gel-separated proteins (Patton 2000; Rabilloud 2000; Patton 2002).

The most commonly used approaches rely on in-gel staining of proteins subsequent to electrophoretic separation. Post-electrophoretic staining of gel-separated proteins can be achieved by noncovalent staining with organic dyes, which rely on the affinity between dye and proteins. For instance, protein staining with a coomassie brilliant blue (CBB R-250) solution in methanol and acetic acid is commonly used. Although this method is rapid and provides a uniform staining of proteins across gels, it is relatively insensitive with a detection limit of 100-200 ng. In addition, this staining method requires the extensive destaining of gels, which can create variations in background staining from gel-to-gel. This method was further improved with the use of colloidal CBB (CBB G-250), a di-methylated form of CBB used in an acidic solution containing ammonium sulphate (Neuhoff *et al.* 1988). The limit of detection of colloidal CBB staining is around 10-20 ng. Another advantage of colloidal coomassie staining is that it does not require de-staining of the gels compared to conventional coomassie staining, therefore increased reproducibility in protein detection from gel-to-gel is achieved. Despite their relative insensitivity, both CBB R-250 and G-250 are commonly used for the analysis of abundant proteins or those that have been overexpressed for post-translational modification analysis, since these dyes are both compatible with further mass spectrometric analysis or Edman sequencing (Scheler *et al.* 1998).

The most commonly used high sensitivity detection methods rely on the precipitation of silver with proteins in acidic or alkaline conditions, a method that can show up to 50-fold higher sensitivity when compared to conventional coomassie staining with CBB R-250. With such approaches, gels are usually fixed and washed in order to

remove excess SDS or salts which could react with silver and produce a high background. Numerous silver staining protocols have been described, although the most sensitive approach uses aldehyde-based fixatives which interfere with subsequent protein identification by mass spectrometry or Edman sequencing. This is due to the fact that proteins are covalently cross-linked to the gel and cannot be efficiently eluted out of gel pieces for further analysis. Other silver staining methods have been developed to overcome these problems and do not use such harsh fixatives (Shevchenko *et al.* 1996; Yan *et al.* 2000). However, these methods provide less sensitivity in protein detection and only surface staining of gels is required if samples are to be compatible with downstream mass spectrometric identification. In addition, staining from gel-to-gel is somewhat variable, as this is not an end-point staining method (the extent of staining can be variable from gel to gel). Furthermore, this staining method appears to be heterogeneous as some proteins show a higher affinity for silver staining than others, making it a poor quantitative approach (Rabilloud *et al.* 1994; Lopez *et al.* 2000). Thus, although silver staining is a highly sensitive method, it is not really suitable for differential analysis and cross-comparison of samples between gels.

So-called reverse staining has also been applied for protein visualisation, using potassium chloride, zinc chloride and copper chloride. These methods rely on the precipitation of salts, which occurs faster with the free salts present in the gel than with the proteins. Therefore this method stains the background of the gel, leaving proteins as transparent spots which can be viewed with appropriate back illumination. The most sensitive staining uses zinc chloride-imidazole (Ortiz *et al.* 1992) where 10 to 20 ng of protein can be detected (Castellanos-Serra 1999). Although this method is rapid, the extent of reverse staining is not fixed and the reverse image tends to fade quickly. Some reports have shown the compatibility of this staining method with further mass spectrometric analysis (Castellanos-Serra 1999).

Other post-electrophoretic protein visualisation methods have used fluorescent dyes, which can be detected under a standard 300 nm UV light or using an imaging device equipped with the appropriate lasers and wavelength filters. The main advantage in the use of such fluorescent stains is that fluorescence gives a very broad linear

dynamic range of detection. This greatly facilitates the process of protein quantitation in biological samples where protein abundance can range over seven or eight orders of magnitude. Hydrophobic dyes such as 1,8-ANS (1- anilinonaphthalene 8-sulfonate) and bis-ANS (bis (8-p-toluidino-1-naphthalenesulfonate)) have been previously used and shown to display a sensitivity level similar to that observed with conventional CBB (Hartman and Udenfriend 1969; Pina *et al.* 1985), and (Horowitz and Bowman 1987). Nile red staining was more recently described as a hydrophobic probe interacting with SDS-protein complexes (Daban *et al.* 1991; Daban 2001). Binding of Nile Red fluorophores relies on the formation of protein-SDS complexes with which the fluorescent dyes can interact. Consequently, this method also requires the use of low concentrations of SDS for the gel electrophoresis to avoid increased fluorescent background. This method shows good sensitivity and detection of as little as 5 ng of protein has been reported (Daban 2001). However, the staining is not stable due to the insolubility of the Nile red dyes in water (Daban 2001).

More recently, Molecular Probes have developed a panel of fluorescent dyes, the SYPRO®dyes (Steinberg *et al.* 1996; Steinberg *et al.* 2000; Berggren *et al.* 2002). The SYPRO Orange, Red and Tangerine dyes also react with SDS-protein complexes, similarly to the Nile red fluorescent dye and the sensitivity reported was similar to that observed with Nile red, although these dyes are more water-soluble. However, for optimum staining the presence of detergent, although necessary, can induce the formation of an apolar environment in the gel that affects protein-staining efficiency. Two transition metal chelate stains have also been developed, SYPRO Ruby and SYPRO Rose stains. The most promising is SYPRO Ruby, a ruthenium-based metal chelate stain showing high sensitivity down to 1 ng of protein (Lopez *et al.* 2000; Berggren *et al.* 2002). These fluorophores react directly with proteins via lysine, arginine and histidine residues, therefore enabling the sensitive and specific detection of all gel-separated proteins (Patton 2000). Importantly, this method was also shown to be compatible with further mass spectrometric analysis (Steinberg *et al.* 2000). SYPRO Ruby staining is rapid as it consists of a single step followed by minimal washes to remove excess dye and can serve as an end point staining method allowing consistent staining from gel-to-gel. However,

this method is a non-covalent staining approach and therefore, consistency in staining and washing times needs to be maintained since the staining is reversible. Consequently, this non-covalent staining method represents a great improvement in the detection of gel-separated proteins for subsequent proteomic analysis, although gel-to-gel variation still remains an obstacle for both statistical and accurate differential analysis (see below).

Protein detection can also be achieved by covalent labelling pre- or post-electrophoretic separation. As previously mentioned, protein modification prior to their separation can be restrictive as covalent modification can affect protein electrophoretic mobility. A commonly used covalent labelling method prior to electrophoresis is radioactive labelling. This method is highly sensitive and provides a very wide linear dynamic range of protein detection. Radiolabelling is mainly accomplished by incorporation of radioactive metabolites during the growth of cells. For instance, ^{3}H, ^{35}S, ^{32}P or ^{33}P can be incorporated into proteins by addition of growth media containing isotopically labelled nutrients. *In vivo* labelling with ^{35}S-methionine and cysteine in particular enables the analysis of newly synthesised proteins and their quantification (Westbrook *et al.* 2001). Proteins can also be specifically labelled, for example specific complexes or lysates can be labelled by *in vitro* kinase assay. Radiolabelled proteins are then detected in-gel by exposure to film (autoradiography) or using a phosphorimager. The latter imaging method provides a wider dynamic range than autoradiography and enables better quantitative analysis. However, concerns remain as to the effects of addition of radioactive compounds to cells. Indeed, DNA damage can occur, as well as elevation of p53 levels leading to cell cycle arrest and apoptosis. Another important issue and drawback in the use of radioactive compounds is their hazardous nature and their cost. Limitations in the use of radioactive compounds now represents a high priority for many laboratories and finding alternative methods is favoured.

Fluorescent tags have also been used to covalently label proteins. Most covalent labelling has been performed on gel-separated proteins, and several fluorophores, such as fluorescamine, fluorescein isocyanate, N-(7-di-methylamino-4-methylcoumarinyl) maleimide and dansyl chloride, have been used for this purpose. Dyes such as 5-(4,6-dichlorotriazin) aminofluorescein, monobromobimane, 2-methoxy-2, 4-diphenyl-3 (2H)

furanone (MDPF) and dansyl chloride have also been used for covalent labelling prior to gel electrophoresis, although low sensitivity and poor resolution on 2DE gels was reported (as reviewed in Patton 2000; Rabilloud 2000). Indeed, the major problem faced was that most of the dyes possess a charge which affects the electrophoretic properties of proteins. In addition to this, covalent protein modification can also dramatically affect the ability to subsequently identify the proteins by mass spectrometry.

Unlu *et al.* recently described a new covalent labelling approach for differential protein detection, known as difference gel electrophoresis (DIGE) (Unlu *et al.* 1997). A combination of two fluorescent protein tags, the NHS-cyanine dyes Cy3 and Cy5, was developed, which enabled differential labelling of two samples prior to their electrophoretic separation on the same 2DE gel. Labelling conditions were established for which protein covalent labelling with these fluorescent dyes prior to gel electrophoresis could be achieved without adversely affecting protein mobility and keeping the sensitivity of detection high and linear over a wide dynamic range. This approach was then further optimised and commercialised by Amersham Biosciences for the analysis of *Escherichia coli*. Although in its early development at the beginning of this thesis, 2D-DIGE was chosen for further optimisation and was then used to investigate the differential protein expression profile of a model cell system of ErbB-2-dependent breast cancer. A complete description of this approach is given in Chapter 3 of this thesis.

Amongst all the detection methods described, fluorescence labelling appears to be the best method of choice due to its sensitivity and wide linear dynamic range. Historically, the main method for visualisation of fluorescently labelled proteins required the use of UV trans-illumination and direct visual detection. However, many fluorphores are excited optimally outside the UV range and detection systems have been developed which use specific lasers for excitation. A single laser source cannot provide the appropriate excitation over the entire wavelength range of different fluorophores and so some devices use a broadband visible light source in combination with narrow band filters enabling the selection of a specific wavelength adequate for each fluorphore of interest. Similarly, the detection device can be equipped with a selection of emission filters, allowing the detection of fluorescence emission for a range of different

fluorophores. Several types of detectors can be used, such as a cooled charged-coupled device (CCD) or photomultiplier tube (PMT) and in this study, a 2920 2D-MasterImager was used. This device, commercialised by Amersham Biosciences, enables optimal excitation and detection of the fluorescent dyes (cyanine dyes and SYPRO Ruby dyes) (Figure 1.4).

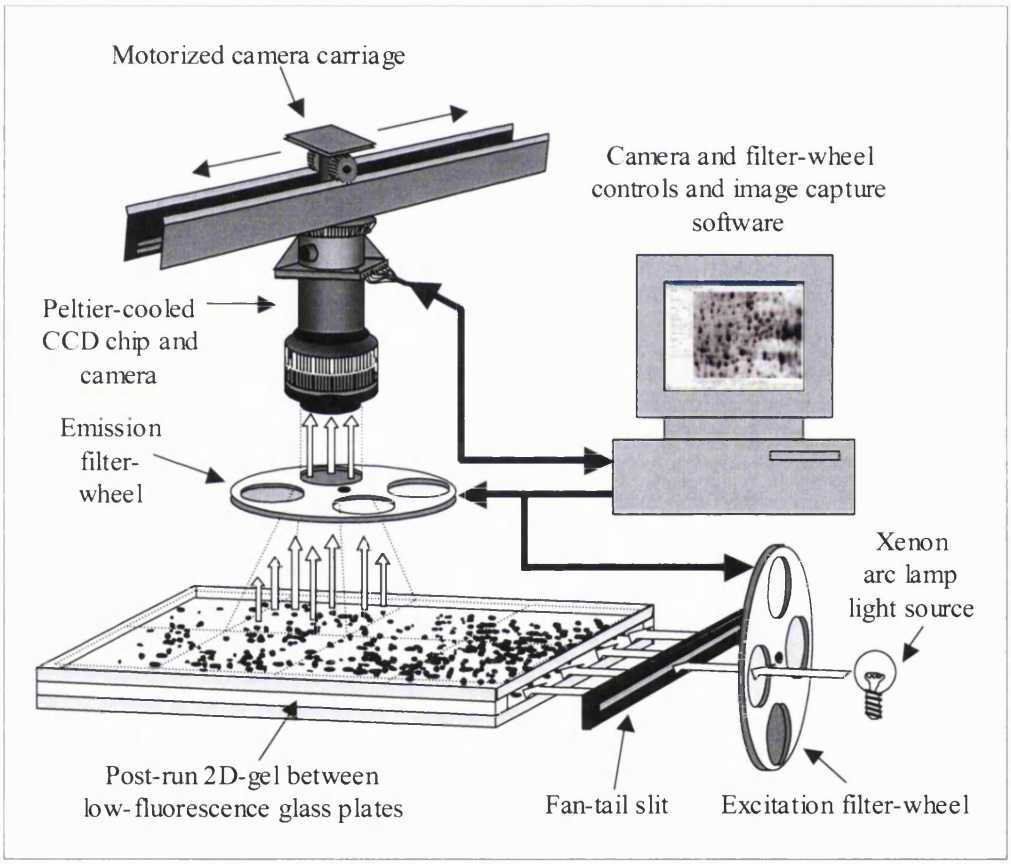


Figure 1.4: Schematic representation of the Amersham Biosciences 2920 2D-MasterImager (Timms unpublished). The gel is directly imaged between low-fluorescence glass plates after 2D-gel electrophoresis. Each separate fluorescent image can be collected individually using filter wheels for the appropriate excitation and emission wavelengths and images are then recorded. This device was used for the image acquisition of cyanine dye-labelled and SYPRO Ruby-stained gels.

This instrument was derived from the previously developed Arthur 1442 Multi-Wavelength Fluorimager (PerkinElmer, Inc.). This fluoroimager has a high intensity xenon arc lamp source equipped with filter wheels enabling the selection of narrow range wavelengths and is combined with a cooled CCD camera providing sensitive detection of fluorescent signals. Using this device, the gel is scanned directly between glass plates using an edge illumination (Figure 1.4). Another imager, the Typhoon 9410 (Amersham Biosciences, UK), relies on a multiple laser source and photomultiplier detection device providing improved fluorescent detection and high image resolution. This imaging device has also been optimised for detection of the cyanine dyes.

1.4.1.3 2D gel-based proteomics in cancer research

Over the past 5-10 years, 2DE-based proteomic analyses have been employed for the identification of new biomarkers of diseases (Hanash 2000). In cancer research alone, many groups have used this approach and have reported extended lists of differentially regulated proteins in cancer cells compared to normal tissues.

In a study of the protein profiling of bladder squamous cell carcinomas, Ostergaard *et al.* identified several highly expressed protein markers used to define the state of differentiation of each cell type tested (Ostergaard *et al.* 1997; Celis *et al.* 1999). Here, 150 fresh bladder tumour samples were analysed by 2D-PAGE combined with protein identification by mass spectrometry. This work revealed the reproducibility and power of 2D-PAGE analysis in the search for potential biomarkers of diseases, enabling the classification of cancers and tissues. Another more recent study reported the 2DE-profiling of lung adenocarinoma (Chen *et al.* 2002). A comparison of 93 lung adenocarcinomas and 10 normal lung samples was carried out and 9 potential biomarkers of disease were identified by mass spectrometry and further validated by western blotting and immunohistochemistry analysis.

A good example of the power of 2DE-based proteomic approaches for cancer research and cell signalling was shown in the study of the MAPK signalling pathway (Lewis *et al.* 2000). In this study, cell lines were used and treated with activators or inhibitors of the MAPK signalling pathway and differentially regulated proteins in

response to specific stimuli were further analysed by mass spectrometry. This enabled the identification of 25 targets of the MAPK signalling pathway with several of those previously reported as important effectors of this pathway. This represents a large-scale experiment as for each of the five time points tested, 2DE gels were run in triplicate from separate experiments. Proteins were then detected by MS-compatible silver staining and protein patterns were compared by gel image analysis using an appropriate curation software (Melanie 2.1).

Advances in proteomics analysis in the study of breast cancer for the identification of biomarkers have enabled the generation of protein profiles characteristic of diseased tissues (as reviewed by Hondermarck *et al.* 2001). Furthermore, several protein expression maps have been characterised and stored in databases. A study showed the clustering analysis of normal adult human luminal epithelial breast cells purified from mammoplasties. Protein expression maps (PEMs) of 20 samples run in duplicate were analysed by 2DE and matched (Page *et al.* 1999), thus providing important cartographies of protein expression profiles of a panel of human breast tissues. Harris *et al.* also carried out a clustering analysis of a variety of human breast cancer cell lines (Harris *et al.* 2002). In this work, 26 breast cancer cell lines and three control cell lines derived from normal breast or benign disease tissue were analysed by 2DE and each sample was run in triplicate. In general, such analyses involve many gels to be run in replicate in order to avoid gel-to-gel variations greatly increasing the workload involved in gel analysis, which is very laborious. In addition, extensive work is needed for the subsequent statistical image analysis. However, this latter work showed the power of 2DE gel for the analysis of protein expression patterns, since reproducible comparison between various normal and breast cancer cell lines could be carried out. Interestingly, the analysis of PEMs could enable the prediction of *in vivo* tumour forming ability based on patient prognosis. Such work represents an interesting basis for the proteomic analysis of ErbB-2 mediated transformation in luminal mammary epithelial cells envisaged in this thesis.

1.4.1.4 The limitations of 2DE based methods

As mentioned previously, 2DE-based proteomic studies are somewhat limited since only a fraction of the proteome can be visualised on a 2DE gel (as reported by Corthals *et al.* 2000). It is estimated that of the tens of thousands of genes expressed in a human cell, only a small fraction (two to three thousand proteins) can be routinely resolved and visualised on a 2DE gel. In addition, this subset of proteins analysed by 2DE methods represents the most highly abundant and soluble proteins. Moreover, each protein expressed may have different splice variants, proteolysis products or be post-translationally modified, therefore increasing the number of protein features potentially detectable from the proteome. The dynamic range of expression of proteins is very high in biological samples, extending over seven or eight orders of magnitude, with a small proportion of the most abundant proteins making up a high percentage of the total mass, and potentially masking the low abundance proteins (Wilkins *et al.* 1996). According to theoretical studies, to analyse a protein expressed at 1000 copies/cell, it would be necessary to load 10 mg of total lysate, assuming that proteins can be visualised at the low nanogram level on a gel. Since high loading precludes high-resolution separation, the low copy number proteins will be masked or not detected on an average 2DE gel stained with the most sensitive stain. This represents a major caveat in this type of analysis since these low abundance proteins could be the most relevant in cancer cell signalling.

It is also important to note the limitations in protein separation in terms of the physico-chemical properties of the protein. Very basic or hydrophobic proteins, such as membrane proteins, are not well separated in the first dimension or are poorly transferred to the second dimension. It may also be difficult to detect low molecular weight proteins as they are often more difficult to stain. Thus overall, only a fraction of the proteome can be separated and visualised on a 2D gel. Gygi *et al.* showed in previous work that the combination of 2DE and mass spectrometry only detected the most abundant proteins in the yeast *Saccharomyces cerevisiae* (Gygi *et al.* 2000). In a small genome such as the yeast genome, it is possible to estimate the level of expression of a protein according to its propensity to use only one of several codons to incorporate a specific amino acid into the polypeptide chain. According to this estimation, proteins which are largely expressed

appear to have a higher codon bias (>0.2). The authors evaluated the protein abundance based on its codon bias and using these estimations, showed that amongst the proteins visualised on a 2D gel and identified by MALDI-MS, the majority were proteins expressed at high copy number per cell. This study thus inferred the need for the development of larger-scale techniques which could make possible the detection of low abundance proteins, for instance using enrichment techniques such as sub-cellular fractionation or liquid chromatography purifications. Indeed, some studies have shown that improved coverage is attained by sub-cellular fractionation, allowing enrichment of specific cellular compartments (Lopez *et al.* 2000; Bell *et al.* 2001). Other new methods aim to improve sample recovery and increase proteome coverage, using immunoaffinity purification of tagged proteins combined with multidimensional liquid chromatography and high-accuracy sequencing by mass spectrometric approaches (see below).

A further caveat to conventional two-dimensional based studies for the analysis of differential protein expression is its semi-quantitative nature. Protein expression profiling requires the comparison of multiple replicate gels in order to generate a statistically valid analysis. However, high sensitive protein staining methods are often non-end point methods and therefore protein detection is somewhat variable from experiment to experiment (as seen with silver staining approaches) obligating the need for comparison of many replicate gels for statistical validity in the results. Most importantly, although the methods of protein separation have been improved and reproducible protein expression patterns are detected, gel-to-gel variations are still observed which limit the accuracy of comparison between samples. Indeed, even minimal variations between experiments such as the temperature or the buffer components, the electrophoretic rigs or even the batch of IPG strips used, can cause variations between gels. Although various proteomic studies have shown successful results in the comparison of multiple samples and enabled the identification of various protein targets (Lewis *et al.* 2000; Benvenuti *et al.* 2002; Chen *et al.* 2002), these kind of studies are labour intensive as many replicate gels need to be run and complex and time consuming image analysis is required.

The ability to run different samples on the same gel, so that proteins are subjected to the same physical microenvironment during gel electrophoresis, would therefore be a

huge technical advance. Covalent labelling with the pair of fluorescent tags described by Unlu *et al.* therefore represents an attractive tool for improved reproducibility for the comparison of multiple samples (Unlu *et al.* 1997). Moreover, fluorescence detection using this method offers a high sensitivity and broader linear dynamic range for better quantitation.

1.4.2 Mass spectrometry in proteomics

Mass spectrometry is an analytical tool allowing the determination of the mass of a molecule with high accuracy. Mass spectrometers are composed of an ion source where the sample is applied, a mass analyser which measures the mass-to-charge ratio (m/z) of ionised particles and a detector. Two major types of approach can be used in the study of proteins and peptides; matrix assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOF-MS) (Karas and Hillenkamp 1988; Tanaka *et al.* 1988), and electrospray ionisation mass spectrometry (ESI-MS) (Whitehouse *et al.* 1985; Fenn *et al.* 1989), which rely on distinct methods of ionisation (Figure 1.4A). The first one, MALDI-MS relies on ions generated from a solid phase using laser pulses (Karas and Hillenkamp 1988). The sample is usually applied in a matrix solution that facilitates the ion formation by absorption of photon energy from the laser beam (Figure 1.5A). The most commonly used matrices for protein/peptide analysis are 2,5-dihydroxybenzoic acid (DHB) and α-cyano-4-hydroxycinnamic acid (CHCA). The second one, ESI-MS generates ions from a liquid phase. The sample, in a solvent mixture (usually 1:1 water/organic solvent-formic acid), is directly sprayed into the mass spectrometer where an electrostatic field is formed between the capillary and the walls of the mass spectrometer and as droplets travel, they evaporate and the charged particles enter into a gas phase (Figure 1.5A).

Both ionisation methods can be performed in the positive as well as the negative ion mode but typically, for peptide mass analysis, ionisation scans are performed in the positive ion mode. Each ion is separated in the mass spectrometer according to their m/z ratio and several mass analysers have been developed for improved ion separation, resolution and mass accuracy.

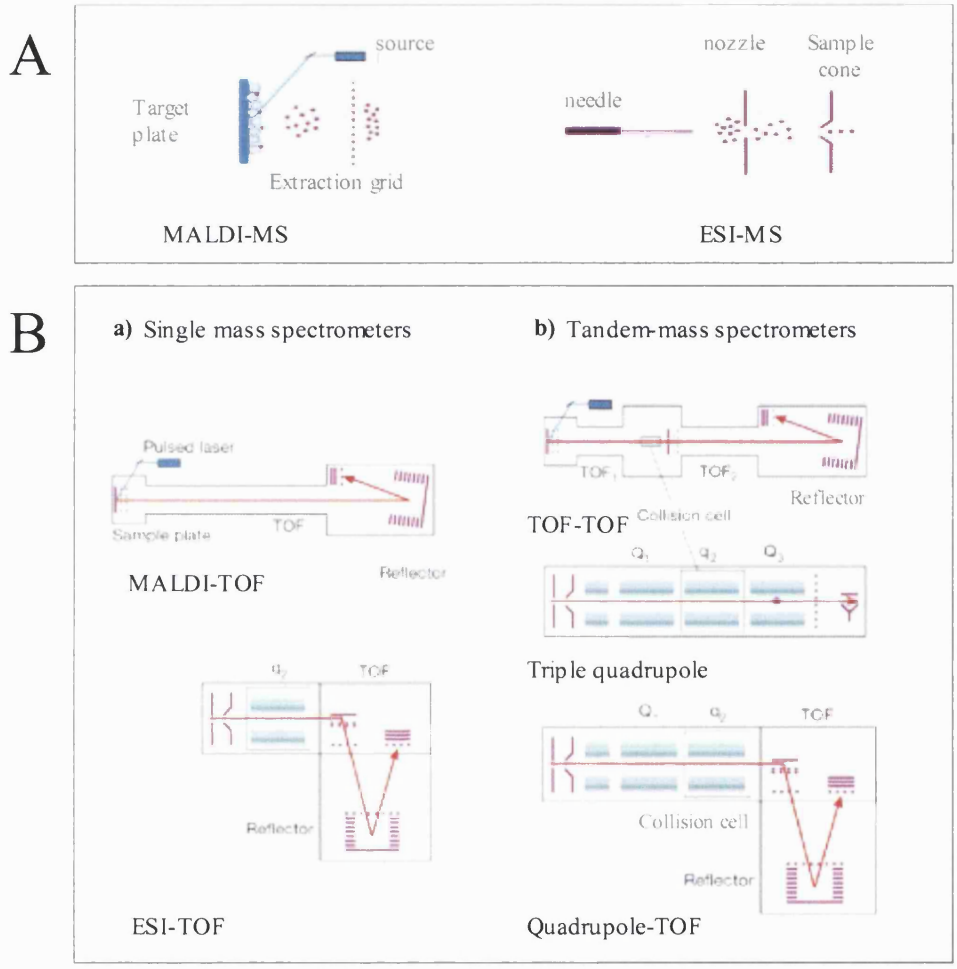


Figure 1.5: Principles of mass spectrometry and instrumentation (adapted from Aebersold and Mann 2003). A) Two main ionisation methods can be applied for protein identification and characterisation. These are matrix assisted laser desorption/ionisation mass spectrometry (MALDI-MS) and electrospray ionisation mass spectrometry (ESI-MS). B) Several mass analysers have been developed and they can be combined for improved mass resolution and sensitivity.

The main types used for biological samples are the quadrupole, quadrupole ion trap, time-of-flight (TOF) and the ion cyclotron resonance (ICR) analysers (Figure 1.5B), and they each differ in the way ions are separated.

Protein analysis by mass spectrometry can be carried out by peptide mass mapping analysis or by peptide fragmentation approaches to generate peptide sequences by tandem mass spectrometry (MS/MS). Peptide mass mapping, also called peptide mass fingerprinting (PMF), is the method of choice for analytical proteomic analysis (Henzel *et al.* 1993; James *et al.* 1993; Mann *et al.* 1993; Yates *et al.* 1993). Briefly, a protein is first proteolysed using specific enzymes to generate a peptide mixture. Subsequently, the masses of the peptide mixture can be accurately measured in the mass spectrometer. The list of peptide masses is then matched against theoretical masses for proteins stored in a database. Several databases can be used, which contain theoretical protein cleavages according to the enzyme of choice and possible post-translational modifications and known chemical modifications. A list of potential matches is given and ranked according to a scoring system depending on best probability of a match being real. Confidence in the identification then depends on the species studied, the mass and pI of the protein observed on a gel, the number of peptides matching the theoretical masses, the mass accuracy of detection, the protein sequence coverage reached, the number of missed cleavages during the proteolysis process and also the type of modifications observed, which reflects the process of sample handling. Protein identification by peptide mass fingerprinting was carried out in this study and is further described in Chapters 2 and 4.

The main approach routinely used for peptide mass mapping relies on MALDI-MS instrumentation, as this ionisation method generates predominantly singly charged ions $[M+H]^+$ and mass spectra are therefore easier to interpret in terms of the peptide mass profile. In addition, this mode of ionisation is more tolerant to high salt concentrations or sample contaminations, as sample crystallisation enables a degree of on-target sample purification.

The other method of choice for protein/peptide analysis is peptide fragmentation by tandem mass spectrometry. In addition to accurate information on the sequence of a peptide and thus the identification of a protein, tandem mass spectrometry is also an

important tool for the analysis of post-translational modifications, such as glycosylation, phosphorylation or acetylation. In tandem mass spectrometry, selected parent ions are fragmented in the mass spectrometer to generate smaller ions with the loss of amino acids (Biemann and Scoble 1987; Hunt *et al.* 1987; Wilm *et al.* 1996). Two main approaches have been used, post source decay (PSD), in a MALDI-MS instrument, and collision induced dissociation (CID) in an ESI-MS instrument. Basically, the common principle of these methods is that ions fragment along the backbone molecule using a high energy, inducing the formation of a series of smaller daughter ions. According to the peptide bond fragmented, the ions are given an annotation as shown in Figure 1.6. Generally, the ions most likely to be generated are the "b" and "y" ions, which represent fragmentations at the amino bond (Figure 1.6).

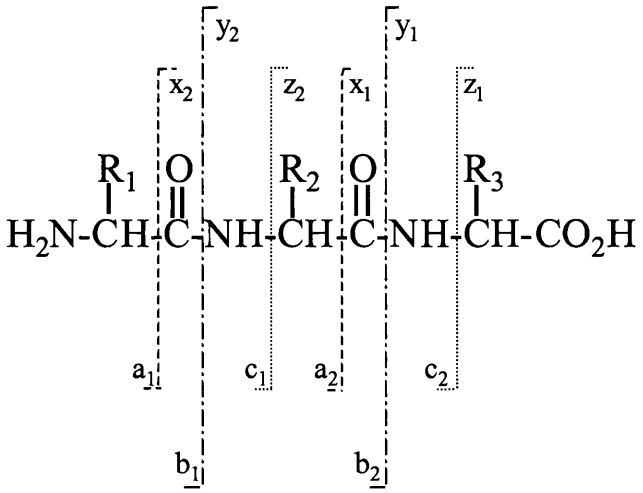


Figure 1.6: Representation of the fragment ions generated from a peptide (from Biemann 1990). A series of daughter ions is generated upon fragmentation at high collision energy. According to the ionisation method applied, the fragment ion pattern will vary. Commonly, in a collision-induced dissociation, the bond most likely to fragment will be the amino bond, therefore generating mainly b and y ions. On the other hand, in post-source decay (PSD), other ions are observed such as a and x ions.

The sequence can thus be determined from the difference in mass between the ions of a series. Analysis of peptide sequence by PSD can be done post peptide mass mapping. A parent ion is selected and further analysed by induced laser energy forcing

the peptide to fragment. The advantage of this method is that several parent ions can be selected from the same aliquot of sample applied onto the target. However, the mass spectra are often harder to interpret due to the complex ion mixtures generated (such as "a" and "c" ions) (Figure 1.6). However, the preferred method for peptide sequence analysis uses CID fragmentation. This can be carried out in a tandem mass spectrometry instrument (for instance by ESI-MS/MS with a Q-TOF or also by MALDI-TOF-MS instruments equipped with a collision cell, such as a MALDI-TOF-TOF-MS). Briefly, peptide mixtures are first analysed in the mass analyser and selected parent ions of interest are focused and fragmented in a collision cell, where increased energy is applied through collision with inert gas molecules to induce fragmentation, and subsequently analysed in a second mass analyser. Then, the mass of each peptide fragment is measured, which all differ by the mass of amino acid residues. In an ideal situation, a sequence of "b" and "y" ions is generated with loss of a single amino acid per peptide backbone allowing the partial sequence to be deduced. Again, the list of fragments is matched against a theoretical database, although manual interpretation of data is often needed for improved confidence in the analysis.

Tremendous progress has been made in recent years in the development of optimised instrumentation for increased mass accuracy, sensitivity, yield of detection and automation. The size of instruments has also decreased and several mass analysers have been combined, providing elaborate and compact analytical instruments for high sensitivity and high-resolution biological mass spectrometry. In addition, high-throughput automation is constantly pushing the detection and identification limits further, allowing a promising future for proteomic analyses.

1.4.2.1 Chromatography based proteomic approaches and mass spectrometry

Other alternatives to conventional gel electrophoresis have also been used for large-scale protein separation and proteomic analysis. In conventional electrospray mass spectrometry, the peptide mixtures are often pre-separated by reverse phase high performance liquid chromatography (RP-HPLC), and directly eluted into the mass spectrometer. This enables the resolution of peptides according to their retention time on

the column during chromatography with simultaneous mass analysis. In addition, HPLC is also used to remove excess salts or other solvents which can affect the ionisation process or complicate the mass spectra and hence diminish the total ion signal for a given peptide. Due to several caveats of the 2DE-based protein separation methods, researchers have used other chromatography approaches to separate complex protein mixtures. These have been combined, greatly reducing sample complexity for accurate and efficient protein identification by mass spectrometry. For instance, two-dimensional and three-dimensional chromatographic separation methods have been used, combining strong cation exchange, reversed phase and affinity columns. Yates and co-workers developed a chromatographic platform for optimum protein purification and separation using a combination of separation methods, the so-called multi-dimensional protein identification technology (MudPIT) (Washburn *et al.* 2001; Washburn *et al.* 2002). In this approach, a complex protein mixture is proteolysed and directly analysed following a two-dimensional liquid chromatography (2D-LC) by strong cation exchange (SCX) coupled to a reverse phase (RP). The chromatographic device is directly connected to a tandem mass spectrometer and resolved peptides directly eluted into the ESI-MS/MS. Therefore all peptides are systematically sequenced and a complex protein mixture can be accurately identified. In contrast to 2DE gel based methods, liquid chromatography can resolve a wide range of proteins encompassing high and low molecular weights and extreme isoelectric points (Koller *et al.* 2002). In addition, larger sample volumes can be applied to columns, which overcomes the problems of overloading faced by the 2DE gel approaches. This multidimensional chromatographic approach has been applied for the characterisation of the yeast genome and showed the power of this technology for the resolution and analysis of low-abundance as well as hydrophobic proteins (Washburn *et al.* 2001). In addition, a complete study of the life cycle of *Plasmodium falciparum* was carried out (Florens *et al.* 2002). This work represented a high-throughput proteomic analysis with the identification of 2,415 parasite proteins, representing 46% of the gene products.

Although attractive, these methods rely on the exhaustive analysis of complex datasets as all peptides are systematically sequenced independently of their expression

levels. In addition, this method is not a quantitative tool and therefore could not be used for comprehensive differential protein expression analysis. In order to circumvent the lack of quantitation in the chromatographic approaches, several groups have developed differential tagging strategies. In one of these, heavy (deuterated) and light (hydrogen) isotope coded affinity tags (ICAT) are used to differentially label tryptic peptides on cysteine residues prior to sample mixing and chromatographic separation (Gygi *et al.* 1999). Such a strategy therefore enables the quantitative comparison of expression ratios between two samples. Enrichment of only the tagged peptides is achieved, since the tags also possess a biotin group allowing avidin affinity purification. However there are several drawbacks to this method such as low yield in sample labelling and recovery of tagged peptides. Stable isotope tagging can also be achieved by enzymatic incorporation. For instance, ^{18}O is transferred from $H_2{}^{18}O$ to the proteins during proteolysis (Mirgorodskaya *et al.* 2000; Yao *et al.* 2001), or by metabolic labelling with stable isotope labelled amino acids (SILAC) such as ^{13}C-arginine or ^{13}C- leucine (Ong *et al.* 2002).

Despite the power of these methods in terms of coverage and sensitivity they are still somewhat limited as only a small number of samples can be compared at once and they necessitate the use of complex integrated softwares in order to deconvolute and quantitate the data generated. In addition, they are not appropriate for the differential protein expression profiling or the comparison of multiple samples in the search for biomarkers of diseases, such studies would be too complex to perform. Therefore it is important to combine this tool with other purification methods. Indeed, the multidimensional chromatography has also been combined with other affinity methods such as immobilised metal affinity chromatography (IMAC) (Ficarro *et al.* 2003), which can be optimised for the specific enrichment of phosphopeptides (Neville *et al.* 1997).

1.4.3 Functional proteomics

Functional proteomic analysis is designed to understand the changes in protein expression within the context of a biological process and is crucial in the analysis of disease states. It is evident that a protein acts within the context of multiple events and in

conjunction with many other components. The idea of functional proteomics is therefore to obtain a more complete understanding of cellular processes at the protein level, not only in terms of expression level but also in terms of understanding protein interaction networks and protein post-translational modifications and their role in cell physiology.

Protein interaction studies are commonly carried out by direct interaction analysis using specific purification methods such as direct immunoprecipitation or pull-downs of epitope tagged overexpressed constructs. The co-immunoprecipitated complex can be resolved by liquid chromatography or gel electrophoresis and the components of the complex analysed by mass spectrometry or else by specific detection by immunoblotting. This method has been extensively used, although it is often difficult to be confident about the nature of the interaction and according to the specificity of the protein-protein interaction, the complex can be more or less easily extracted. However, problems associated with lack of specificity of antibody-antigen interactions have to be assessed prior to any functional analysis. Importantly, the large excess of immunoglobulin used to affinity purify the complex of interest often represents a source of contamination affecting the ability to identify *bona fide* co-immunoprecipitating proteins.

A widely used approach in the study of protein-protein interaction is the yeast-two hybrid approach (Ma and Ptashne 1988; Fields and Song 1989). This method enables the identification of putative interacting partners of a selected protein (method described by Van Criekinge 1999). Briefly, the protein of interest is fused to a DNA-binding domain (DB) to form the "bait" and screened against a library of cDNA clones fused to an activation domain (AD) also called the "prey". The formation of the complex DB-AD enables transcriptional activation and expression of a reporter gene. The ability of a cDNA clone to induce the expression of the reporter gene is representative of the specific protein-protein interaction between the protein of interest and an unknown potential interacting partner. Screening of a whole genome can be done and systematic analysis of protein interactions can be achieved independently of any knowledgeable information on the plausibility of such protein interactions. However, this approach presents certain limitations, since it is carried out in the yeast (*S. cerevisiae*), a system which is not representative of other biological systems of interest. For instance some proteins cannot

fold correctly or display the appropriate post-translational modifications needed for their active protein-protein interaction. In addition, some proteins are not suited for yeast-two hybrid analysis, such as the ones with putative roles as transcription factors. The chimera can also alter the conformation of proteins and inhibit protein-protein interactions. However, it represents an important starting point for the identification of potential interacting proteins.

Large-scale analysis of multi-protein complexes has also been performed by immunoprecipitation-based approaches. These methods have been further improved with the advances in mass spectrometry as an analytical tool, enabling the systematic identification of interacting partners. A FLAG epitope-tagging method was used for characterisation of the nucleolin-binding ribonucleoprotein complex (Yanagida *et al.* 2001). Here, proteins of choice are fused to an epitope tag, such as Myc (EQKLISEEDL), HA (YPYDVPDYA) and VSV-G (YTDIEMNRLGK), and antibodies against these tags are used to immunoprecipitate the protein of interest and its co-immunoprecipitating partners. The protein-complex can then be analysed by conventional 1D-SDS-PAGE combined with MALDI-MS analysis for protein identification. Such approaches enable the precipitation of proteins with increased affinity due to the use of the tag. Another affinity approach called tandem affinity prurification (TAP) was also used in the yeast system (Rigaut *et al.* 1999). This purification is carried out in two affinity steps, to improve the purity and specificity of the multi-protein complex without removing low affinity interactions. The target protein is fused with a calmodulin binding protein (CBP) tag and a protein A tag, separated by a specific cleavage site for the tobacco etch virus (TEV) protease. Multi-protein complexes are first purified through an IgG column (for protein A binding) and eluted by enzymatic cleavage with the TEV protease. Contaminants are then removed on a second low affinity column with calmodulin beads (for calmodulin binding proteins CBP) and pure complexes recovered by gentle elution. Using this approach, native proteins complexes can be further purified and identified by conventional gel-based or chromatography-based protein resolution and mass spectrometry. This technique was applied for the large-scale analysis of the yeast proteome and identified 232 multiprotein complexes from 1,739 tagged genes (Gavin *et*

al. 2002). Although originally developed in the yeast system, some reports have shown its potential application in mammalian cell systems (Cox *et al.* 2002) and this method was also recently used in combination with double-stranded interference RNA technology, which enabled the accurate specificity of protein interaction by knocking down the endogenous gene (Forler *et al.* 2003).

Large-scale protein interaction analysis has also been developed using protein array technology and has been applied to the yeast system enabling the large-scale analysis of protein interactions (Zhu *et al.* 2001). Multiple applications of protein arrays have been developed (reviewed by Jenkins and Pennington 2001). The principle of such approach is that an array of molecules are immobilised on a target and allowed to interact with a complex protein mixture. Specific protein-protein complexes can then be detected using various means of detection, for instance by fluorescence tagging or by direct mass spectrometric analysis such as the surface enhanced laser desorption/ionisation-MS (SELDI-MS) approach (Issaq *et al.* 2002). Specific interaction can then be screened across the target, giving rise to a list of potential interacting partners.

This scale of study represents an extensive work and relies on the knowledge of the genome of the species analysed. Here, the studies reported have been carried out in the yeast, which shows less complexity compared to human systems. However, all these newly developed methods are very promising for the future of proteomic analysis and some of them have been developed for the diagnosis of diseases, such as the protein chip microarray and SELDI approaches.

1.5 Aims of this study

The field of proteomic analysis has evolved considerably over the last 10 years and major improvements in the techniques as well as instrumentation have been developed thus providing powerful tools for the analysis of complex cellular events. The aim of the study presented in this thesis was to attempt to understand the role of ErbB-2 overexpression in cellular transformation and subsequent development of cancers, mainly breast cancer. In the search for new molecular markers of ErbB-2 positive human breast cancers, a proteomic strategy was envisaged. ErbB-2 is thought to play a key role in the

control of complex signalling events including interactions with other ErbB family members and also cross-talk with other signalling pathways. ErbB-2 overexpression is therefore estimated to affect several known key cellular processes as well as new uncharacterised cellular functions, which could be important to assess in the attempt to identify new potential targets for breast cancer therapy. In order to obtain a general view on the complex machinery involved in ErbB-2 overexpression-mediated cellular transformation, it was apparent that a large-scale protein analysis approach would be advantageous. 2D-DIGE, a new method that putatively overcomes problems of gel-to-gel variation and increase the linear dynamic range of detection was evaluated. Subsequently, a 2D-DIGE optimised protocol was applied for improved reproducible and quantitative proteomic analysis. This approach was first described by Unlu *et al.* and permits the differential labelling of proteins prior to their electrophoretic separation on a 2DE gel. Since this technique was early in its development, it had to be first optimised and validated for protein expression profiling. This work is described in Chapter 3 of this thesis.

The goal of this study was to identify new proteins, which are differentially regulated in response to ErbB-2 overexpression and following treatments with specific EGF-family growth factors. For this purpose, a 2D-DIGE-protein detection platform was developed and a workflow established for rapid, accurate and sensitive quantitation of protein differences which could be subsequently identified by mass spectrometry. The study of ErbB-2 overexpression was carried out in a model breast cancer cell system, comprising of a paired parental cell line, HB4a and a derived clone, C3.6, which overexpresses ErbB-2. The differential protein expression analysis of this system, using 2D-DIGE is described in the Chapters 4 and 5. Furthermore, in order to understand the impact on the differential regulation of selected potential markers of ErbB-2 overexpression, it was then crucial to validate the protein changes in the context of cancer cell biological study. Following on from the proteomic DIGE analysis, further functional characterisation of potential targets of ErbB-2 overexpression was addressed using other biochemical and cell biological approaches. The consequences of the protein expression pattern imposed by ErbB-2 overexpression were further investigated in the context of

other signalling pathways and cellular processes, with particular emphasis on the regulation of the other ErbB family members (Chapters 5 and 6).

Chapter 2 : MATERIALS AND METHODS

All chemicals and reagents were purchased from Sigma-Aldrich Company Ltd., UK, unless otherwise stated.

2.1 Tissue culture

2.1.1 Cell growth

HB4a, C3.6 and other clones derived from the HB4a cell line, were obtained from Dr. M. O'Hare as described (Harris *et al.* 1999). They were maintained in 1640 RPMI media containing 10% (v/v) foetal calf serum (FCS), supplemented with L-glutamine (2 mM), streptomycin (100 μg/ml), penicillin (100 IU/ml) (all purchased from Gibco-Invitrogen Corp., UK), insulin (5 μg/ml) and hydrocortisone (5 μg/ml) in tissue culture dishes at 37°C in a 10% CO_2 humidified incubator. Cells were passaged when they reached 70 to 80% confluence by trypsinisation according to standard procedures.

A normal breast luminal epithelial cell line, Lum878, immortalised with SV40 large T antigen and *hTERT*, was maintained following the same growth conditions as those previously described for HB4a and C3.6 cells. The breast tumour cell lines, BT474, SKBr3, MB361, MB435, MB453 and MCF7 were maintained as reported (http://www.atcc.org/SearchCatalogs/CellBiology.cfm). Briefly, they were grown in DMEM media (Gibco-Invitrogen Corp., UK) containing 10% (v/v) FCS, supplemented with streptomycin (100 μg/ml) and Penicillin (100 IU), in tissue culture dishes at 37°C in a 10% CO_2 humidified incubator.

Stocks of luminal epithelial cells were stored in liquid nitrogen and new batches of low passage cells were used after 25 passages in order to avoid potential genetic drift, which can occur during prolonged passaging of cells. For freezing, trypsinised cells were pelleted and resuspended in 10% (v/v) DMSO in 90% (v/v) FCS. Cells were slowly taken to −70°C in a propanol container and then transferred to liquid nitrogen for long-term storage. Rapid thawing was carried out in a 37°C water bath and excess DMSO was removed by centrifugation prior to plating. Cells were allowed to recover for three passages prior to sampling.

2.1.2 Cell treatment with growth factors and inhibitors

Cells were treated for various periods of time with specific growth factors. In this work, two ligands specific for ErbB family members were used, epidermal growth factor (EGF) (R&D systems) and Heregulinβ1 (Hrgβ1) (R&D systems). For treatment, cells were first serum-starved in 0.1% (v/v) FCS in 1650 RPMI media supplemented with L-glutamine (2 mM), hydrocortisone (5 μg/ml) and antibiotics for 48 hours. They were then treated for various times with either 1 nM (6 ng/ml) of EGF or 1 nM (8 ng/ml) of Hrgβ1. Specific inhibitors were also used to study downstream signal pathways and were typically added 30 min prior to treatment with growth factors. A specific tyrosine kinase inhibitor of EGFR and ErbB-2, AG1478 (Biosource, California, US) was used at 5 μM, a concentration which can inhibit tyrosine phosphorylation of both EGFR and ErbB-2 receptors and abrogate downstream activation of the MAP Kinase and PI3 Kinase pathways (Chan and Timms, unpublished data). Inhibition of the proteasome was carried out using 1 μM PS341 (Millenium Pharmaceuticals, Inc.) (Adams *et al.* 1999). Inhibition of protein synthesis was achieved using 10 μg/ml cycloheximide (CHX).

2.2 Sample preparation

2.2.1 Cell lysis-Protein extraction

Cells were always lysed at sub-confluency (80% confluency). They were washed twice in 0.5x ice-cold phosphate-buffered saline (PBS) and lysed in the appropriate lysis buffer. The conditions varied according to the subsequent protein separation method used or for fractionation analysis and immunoprecipitation work.

For one dimensional sodium dodecyl sulphate polyacrylamide gel electrophoresis (1D-SDS PAGE) and subsequent western blotting, cells were lysed in an NP40 lysis buffer (50 mM HEPES; 150 mM NaCl; 1% NP40; 1mM EDTA) supplemented with protease inhibitors and phosphatase inhibitors (100 μg/ml AEBSF, 17 μg/ml aprotinin, 5 μM BpVphen, 5 μM fenvalerate, 1 μg/ml leupeptin, 1 μM okadaic acid, 1μg/ml pepstatin, 2 mM sodium orthovanadate). Routinely, cells from a 15 cm plate were lysed in 700 μl of

lysis buffer and scraped on ice. Lysates were left 20 min on ice and cell debris was cleared by centrifugation at 14,000 rpm for 10 min at 4°C.

Protein concentrations were determined using a Bradford microtitre plate assay (Spectra max plus, Molecular Devices). Routinely, 2 µl of total cell lysate was used and mixed with 200 µl of Coomassie protein assay reagent (Pierce-Perbio, IL, USA). Absorbance readings were taken at 595 nm on a microtitre plate spectrophotometer. Standard curves were determined using dilutions of bovine serum albumin (BSA) and samples were assayed in triplicate.

Samples were reduced and denatured in Laemmli sample buffer (50 mM Tris pH 6.8, 10% (v/v) glycerol, 2% SDS (w/v), 0.1% (w/v) bromophenol blue, 2% β-mercaptoethanol) with boiling for 5 min at 100°C.

For conventional two-dimensional polyacrylamide gel electrophoresis (2D-PAGE), cells were lysed in conventional 2DE lysis buffer (4% CHAPS, 8 M urea, 2 M thiourea, 65 mM dithiothreitol (DTT)) containing protease and phosphatase inhibitors (see above). Samples were passed five times through a 25G needle to homogenise and shear DNA and insoluble cell debris was cleared by centrifugation at 14,000 rpm for 10 min at 10°C. Following protein quantitation by Bradford assay, 200 µg of total cell lysate was taken for 2DE and the final volume was adjusted to 350 µl with 2DE lysis buffer, and containing 2% (v/v) of a 50% (v/v) ampholine/Pharmalyte pH 3 to 10 mixture (Amersham Biosciences, Inc., UK) and bromophenol blue.

For 2D-difference gel electrophoresis (2D-DIGE), cells were lysed and scraped in a 2D-DIGE lysis buffer (4% CHAPS, 8 M urea, 2 M thiourea, 10 mM Tris, pH 8.3). DTT was omitted from the buffer since high concentration would interfere with dye labeling efficiency. In addition, the pH was adjusted to 8.3 with 10 mM Tris since dye labeling is most efficient between pH 8-9. Protease and phosphatase inhibitors were also used (see above) but AEBSF was omitted, as it was not compatible with dye labelling. Following protein quantitation by Bradford assay, aliquots of 100 µg of total cell lysate were used for further DIGE labelling with NHS- Cyanine dyes as described in section 2.2.2.

2.2.2 *Protein labelling with NHS-cyanine dyes (DIGE labelling)*

The NHS-Cyanine dyes (Cy2, Cy3, Cy5: Figure 3.2A Chapter 3) were purchased from Amersham Biosciences and subsequently synthesised "in-house" by Dr. P. Gaffney (LICR). Cyanine dye stock solutions of 1 mM in dimethyl formamide (DMF) could be kept at -20°C for up to 4 months without loss of labelling efficiency. Cyanine dye dilutions for labelling were made fresh in DMF and kept in the dark and on ice prior to use. Routinely, protein extracts prepared in 2D-DIGE lysis buffer (see section 2.2.1) were "minimally" labelled with dye to protein ratios of 4 pmol of dye/μg of protein. Labelling was performed on ice in the dark for 30 min and reactions were quenched with a 50-fold molar excess of free lysine to dye for 10 min on ice in the dark.

For the sensitivity test, a pure solution of 10 mg/ml of transferrin in 2D-DIGE lysis buffer was prepared and 50 μg of transferrin was labelled with 200 pmol of cyanine dye. The reaction volume was kept to a minimum and reaction was carried out on ice in the dark for 30 min. 10 nmol of free lysine was then added for 10 min to stop the reaction for and 5x Laemmli sample buffer added to fully reduce and denature the sample. Serial dilutions of labelled transferrin were then loaded and run on 8% 1D-SDS-PAGE gel.

For the proteomic analysis of HMLEC cell lines, 100 μg of each sample was "minimally" labelled with 400 pmol of either Cy3 or Cy5 in a minimal reaction volume. The reaction was performed in the dark on ice for 30 min and then quenched with 20 nmoles of free lysine for 10 min. Labelled samples were pooled appropriately and proteins were reduced by addition of DTT to 65 mM for 15min. The volume was adjusted to 350 μl with 2D-DIGE-lysis buffer, with concentrations of DTT and ampholyne/pharmalyte (pH 3 to 10) adjusted to 65 mM and 2% (v/v) respectively, and bromophenol blue added.

For the analysis of growth factor treatment on the proteome of the HB4a and C3.6 cell lines, cell lysates were prepared as described above from appropriately serum-starved or stimulated cells. Four time points were tested and each sample was analysed in triplicate. Cell lines from the same time point were compared in pairs with an internal standard run on the same 2DE gel. These were run in triplicate, representing a total of 12 gels (Figure 4.8, Chapter 4). Routinely, 100 μg of HB4a lysate was labelled with 200

pmol of Cy3 and 100 μg of C3.6 lysate was labelled with 200 pmol of Cy5. A pool of all samples was made and 100 μg of pool was labelled with 200 pmol of Cy2 and used as a standard to be run on all gels. Labelled samples were mixed prior to protein separation by gel electrophoresis. All gels were cast and run at the same time. For automated protein picking for subsequent mass spectrometric (MS) identification, gels were bonded to the inner glass plates. The inner glass plates were treated with 1 ml of Bind Silane solution (8 ml ethanol, 200 μl acetic acid, 1.8 ml ddH_20, 8 μl Bind Silane (Plus One, UK)) and left to dry at room temperature for 1 hour prior to gel casting. Reference spots were stuck to the inner glass plate and were used to assign spot coordinates for the automated picking. These were fluorescent spots which could be detected during the gel-imaging process.

2.3 Protein separation by polyacrylamide gel electrophoresis

For 1D-SDS PAGE, proteins were separated on a Hoeffer gel rig system (Amersham Biosciences, UK) using standard procedures. Adjusted percentage SDS-PAGE gels were run according to the molecular weight of the protein of interest. For western blotting, 30 μg of total cell extract was loaded onto the gel (see section 1.7.1).

For 2D-PAGE, samples were applied to immobilised pH-gradient (IPG) strips 18 cm, non-linear, pH 3-10 (Amersham Biosciences, UK) by over-night gel rehydratation at room temperature. For cyanine dye-labelled samples, IPG strips were rehydrated in the dark to avoid photo-bleaching of the dyes. Isoelectrofocusing was performed on a Multiphor II apparatus (Amersham Biosciences, UK) at 20°C with the following steps applied, 300 V for 30 min increasing to 3500 V over 3 hours, then 3500 V was maintained for a total of 80 kVhrs. IPG strips were then equilibrated in equilibration buffer (EB) (50 mM Tri-HCl pH 6.8, 6 M Urea, 30% (v/v) glycerol, 1% SDS) with gentle rocking. The equilibration was carried out for 15 min in EB containing 65 mM dithiothreitol (DTT) to reduce disulfide bonds and for a further 15 min in EB containing 240 mM iodoacetamide (IA) to alkylate reduced thiol groups. Equilibrated IPG-strips were then transferred onto 18 cm x 20 cm x 1.5 mm thick 12% acrylamide/0.8% bis-acrylamide gels. IPG strips were overlaid with 0.5 % (w/v) of low-melting point agarose in running buffer (25 mM Tris base, 192 mM glycine, 0.1% (w/v) SDS) containing

bromophenol blue. Routinely, electrophoresis was carried out in cooled tanks (Protean II-Biorad Laboratories, Inc., USA) at 10°C, 30 mA per gel, until the bromophenol blue dye-front had run-off. It was important to ensure that the dye front had completely run-off, otherwise the bromophenol blue would interfere with fluorescence imaging.

2.4 Protein detection-Gel staining and imaging

Several gel staining methods were used and compared for post-staining ability and sensitivity.

2.4.1 Coomassie brilliant blue staining (CBB-R250)

Gels were fixed for a minimum of three hours in fixing solution (40% (v/v) methanol, 7.5% (v/v) acetic acid). They were then incubated in 0.1% CBB-R250 in 45% (v/v) methanol, 10% (v/v) acetic acid for a minimum of 1h and then destained in 50% (v/v) methanol, 10% (v/v) acetic acid until the background staining was clear.

2.4.2 Silver staining

A mass spectrometry (MS)-compatible silver staining protocol was employed, according to Shevchenko *et al.* (Shevchenko *et al.* 1996). Briefly, gels were fixed for a minimum of three hours in fixing solution (40% (v/v) methanol, 7.5% (v/v) acetic acid) and then washed twice for 15 min in ddH_2O in order to remove excess acetic acid and SDS. Gels were sensitised in 0.02% (w/v) sodium thiosulfate for 1 min and washed twice in ddH_2O for 1 min each. Gels were then incubated for 20 min in 0.15% (w/v) silver nitrate at 4°C and then washed twice in ddH_2O for 1 min each before development in 2% (w/v) sodium carbonate with 0.02% (v/v) formaldehyde. The reaction was stopped with 5% (v/v) acetic acid. Gels were stored in 1% (v/v) acetic acid at 4°C for eventual analysis by mass spectrometry.

Another silver staining protocol was also used for the comparison of protein detection methods, however, this staining protocol is not mass spectrometry compatible since it includes cross-linking of proteins with glutaraldehyde (Hochstrasser and Merril

1988). Briefly, gels were fixed for a minimum of 1 hr in fixing solution (40% (v/v) methanol, 10% (v/v) acetic acid) and washed in ddH_2O for 5 min. Gels were then soaked in 25 % (v/v) glutaraldehyde solution for 30 min in a fume hood and were then extensively washed in cold ddH_2O 4 times for 30 min. Staining was performed in a staining solution (2% silver nitrate, 0.133% (v/v) ammonium hydroxide, 20 mM sodium hydroxide) for 10 min with gentle rocking. Gels were then washed 3 times 5 min in ddH_2O and the staining was developed in 0.01% (w/v) citric acid, 0.04% formaldehyde and stopped in 5% acetic acid.

2.4.3 SYPRO®Ruby fluorescent staining

Routinely, bonded gels with separated cyanine dye-labelled proteins were post-stained with SYPRO®Ruby fluorescent protein stain (Molecular Probes, USA). Following electrophoresis, gels were fixed for at least 3 hours in fixing solution (30% (v/v) methanol, 7.5% (v/v) acetic acid, in water). Lower concentrations of methanol and acetic acid were used in order to avoid gel shrinkage and cracking of bonded gels. Gels were washed twice in water and then incubated in SYPRO Ruby staining solution for a minimum of 3 hours in the dark. Gels were then washed briefly in water to remove excess dye, which would otherwise cause high background staining.

2.4.4 Gel imaging

For gels stained using visual staining methods (coomassie blue staining and silver staining), gel images were acquired using a BioRad GS800 scanning densitometer. Images were exported as 8-bit tiff images from QuantityOne software (BioRad, Laboraories Inc., USA) into Photoshop or processed as described in section 2.4.5.

For 2D-DIGE analysis, gels were imaged directly between glass plates directly after electrophoresis run using the 2920 2D-Master Imager (Amersham Biosciences, UK). This charge coupled device (CCD)-based instrument possesses two six-position filter wheels (excitation and emission) enabling simultaneous scanning under different wavelengths specific for each cyanine dye used and for SYPRO Ruby. Excitation and

emission wavelengths used are shown in Table 2.1. Images with optimal pixel intensity, but below pixel saturation (<64,000 pixels) were obtained by adjustment of the exposure times for each dye labelled sample. The images generated were exported as 16-bit tiff images for further protein expression profile analysis and quantitation as described in section 2.4.5. Gels post-stained with SYPRO Ruby fluorescent stain were also imaged using the 2920 2D-Master Imager. Again, sub-saturation images were obtained by adjustment of exposure times.

	Excitation (nm)	Emission (nm)
Cy2	**480**	**530**
Cy3	**540**	**590**
Cy5	**620**	**680**
SYPRO Ruby	**400**	**632**

Table 2.1: Excitation and emission wavelengths used for the detection of each cyanine dye and SYPRO Ruby fluorescent stain using the 2920 2D-Master Imager.

2.4.5 Image analysis

Images were curated and compared prior to selection of proteins of interest for picking and subsequent identification by MS. More obvious differences could be observed visually by pseudo-colouring and direct overlay of images from the same gel using Adobe Photoshop. However, for improved image analysis and statistical validation, curation softwares were employed. Several softwares are commercially available for gel image processing and two softwares were employed in this study, Melanie III (Swiss Institute of Bioinformatics, Geneva, Switzerland) and DeCyder (Amersham Biosciences, UK). The advantage of the Melanie III software was that spot detection could be performed with adjusted parameters allowing the comparison of distinct gel images produced using different staining methods, gel-imaging systems or gel imaging devices.

However, more manual intervention was required, which was time consuming and user-based, making it more likely to generate variations and hence diminish the level of reproducibility of the experiment. The DeCyder software was used for automated 2D-DIGE-gel analysis and was designed for the comparison of differentially labelled samples run on the same gel gel. The software consists of two modules, the differential in-gel analysis (DIA) module and the biological variation analysis (BVA) module. The DIA module detects overlapping features of paired gel images generated from the same 2DE and automatically performs spot detection, filtering, background subtraction, normalisation and quantitation of spot differences. Subsequently, detected spots from these paired images are matched using the BVA module which also performs a user-guided statistical analysis to find consistent differences between all matched spots from multiple samples. For such analysis, an internal standard is employed and run on each gel against the analytical sample(s). This facilitates matching across gels and allows a spot-by-spot internal normalisation for improved reproducibility and accurate quantification of protein abundance. Spot filtering can then be performed and only spots of statistical relevance and sufficient volume are selected for the analysis. Routinely, thresholds of +/- 1.23 average fold-difference were set with p values <0.05 from triplicate samples and a volume cut-off >100,000 pixels (undetectable by most sensitive staining methods) were selected for peptide mass mapping by MALDI-MS.

2.5 Protein identification by mass spectrometry

2.5.1 In-gel digestion

All samples to be analysed by mass spectrometry were prepared in a category 6 clean room in order to avoid keratin and other contaminations. Gel pieces were excised either manually for silver/coomassie stained gels or automatically using a Syprot spot picking robot (Amersham Biosciences, UK). For automatic spot picking, two markers were added to the glass plates during gel casting and were used as reference for the determination of spot coordinates. A picking list was generated with Melanie III or DeCyder and exported to the spot picker as a text file.

A previously described in-gel digestion protocol was employed with modifications (Shevchenko *et al.* 1996). Briefly, silver stained gel pieces were destained with 1% (v/v) hydrogen peroxide (H_2O_2) for 10 min where applicable. Next, gel pieces were washed 3 x in 25 mM ammonium bicarbonate (NH_4HCO_3)/50% (v/v) acetonitrile. Proteins were reduced in 10 mM DTT in 25 mM NH_4HCO_3 for 50 min at 50°C and subsequently alkylated in 50 mM iodoacetamide (IA) for 60 min in the dark at room temperature. Removal of excess alkylating agent was performed by two washes in 25 mM NH_4HCO_3/50% acetonitrile and the gel pieces were dried in a speed vac centrifuge (Speed Vac, Savant, NY, USA) for 10 min or until dryness. 5 µl of a 10 ng/µl solution of sequencing grade modified trypsin (Promega, Southampton, UK) in 25 mM NH_4HCO_3 was added to dry gel pieces, which were allowed to re-swell for a minimum of 5 min. Gel pieces were then overlaid with 25 mM NH_4HCO_3 solution and left for a minimum of 18 hrs at 37°C. Tryptic peptides were extracted with 50% acetonitrile/5% trifluoroacetic acid and vacuum centrifuged to complete dryness. They were then re-suspended in 5 µl of ddH_2O and kept at -20°C for mass spectrometric analysis.

2.5.2 MALDI-MS and peptide mass fingerprinting

Matrix-assisted laser desorption mass spectrometry (MALDI-MS) with peptide mass mapping was used for protein identification. Routinely, 0.5 µl of tryptic digest was mixed with 1 µl of matrix solution (saturated aqueous 2,5-dihydroxybenzoic acid (DHB)), onto a sample plate and dried using a hair dryer to allow the formation of crystals. MALDI mass spectra were acquired using a Reflex III time of flight (TOF) mass spectrometer (Bruker Daltonik, Bremen, Germany) in the reflector mode. The mass spectrometer was calibrated with external standards using a standard mixture of peptides (calibration mixture 2 from Sequazyme kit, Applied Biosystems, USA). Internal calibration of the mass spectra was carried out using reference trypsin autolysis peaks: m/z 842.51 and m/z 2211.1046. Prominent peaks in the mass range between m/z 700 to 4000 were used to generate a peptide mass fingerprint.

2.5.3 Database searching

Each peptide mass fingerprint was matched against the updated NCBI database using Protein Prospector version 3.4.1, MS-Fit (University of California, San Francisco, USA), which calculates theoretical accurate tryptic masses for all proteins in the database. A scoring system was used for validation of protein identities. Typically, a positive identification was accepted when a minimum of 6 peptide masses matched the theoretical masses for a particular protein with a mass error of +/- 50 ppm and up to 2 missed cleavages were allowed. Moreover, the matched peptides had to represent a minimum of 30% of the sequence of the protein identified and the protein hit had to display the approximate molecular weight and pI as the spot detected on the gel. When a protein hit fulfilled the specified thresholds of validation for protein identification, unmatched peptides were systematically re-submitted to the database in a search for possible multiple proteins per gel plug and potential sites of post-translational modification were checked (e.g. phosphorylation).

2.6 Validation and functional proteomics

2.6.1 Immunoblotting/western blotting

Cell extracts were separated by 1D or 2D-gel electrophoresis and electro-blotted onto polyvinylidene fluoride (PVDF) membrane (Immobilon P, Millipore) using a wet transfer tank in transfer buffer (195 mM glycine, 25 mM Tris, 20% (v/v) methanol). Membranes were blocked for 1h with 5% (w/v) low fat milk in Tris buffered saline (50 mM Tris pH 8, 150 mM NaCl) with 0.05% Tween-20 (TBS-T) or in 5% BSA in TBS-T for anti-phosphotyrosine immunoblotting. Membranes were then incubated for a minimum of 1h in a primary antibody solution in TBS-T. The primary antibodies used, their dilutions and sources are shown in Table 2.2. Membranes were washed in TBS-T (three times 15 min) and then probed with the appropriate horseradish peroxidase (HRP)-coupled secondary antibody (Amersham Biosciences, UK). After further washes in TBS-T, immunoprobed proteins were visualised using the enhanced chemiluminescence (ECL) method (NEN-Life Science Products, Inc., USA).

2.6.2 *Immunoprecipitation*

Proteins were immunoprecipitated from 300-600 μg of total cellular extract in NP40 lysis buffer. The cleared supernatant was incubated with primary antibody (1-2 μg) and protein A or protein G (according to the type of antibody used) coupled to sepharose beads, for 2 hrs at 4°C with rotation. Immunecomplexes were washed 5 times in NP40 lysis buffer and then boiled for 5 min in Laemmli sample buffer for complete reduction and denaturation prior to separation by 1D-SDS-PAGE. Immunoprecipitates were then analysed by western blotting with appropriate antibodies.

For large-scale immunoprecipitations, 4 mg of total cell lysate were used with 20-40 μg of antibody. 1D-SDS-PAGE separated immunecomplexes were analysed by silver staining to visualise co-immunoprecipitating proteins. Non-specific binding was controlled with a parallel immunoprecipitate using a protein A purified non-immune IgG, from rabbit serum, following an identical procedure. Specific bands were then cut and processed by mass spectrometry as described in section 2.5.

2.6.3 *Immunostaining*

Sub-cellular localisation of different proteins was investigated by immunofluorescence. Cells ($2x10^5$) were seeded into 12 well plates containing sterile cover slips and allowed to grow to 80% confluency. Due to the sensitive growth of the HB4a cell line on cover slips, cells were serum starved for only 24 hrs as opposed to 48 hrs in 0.1% FCS/1640 RPMI starvation media. Cells were then treated for various time points with 1 nM (6 ng/ml) EGF and then washed in ice cold PBS and fixed in 4% paraformaldehyde for 20 min at room temperature. The cells were then permeabilised in 0.2% (v/v) Triton X100 in TBS for 5min and washed twice in TBS and blocked in 3% BSA for 30 min at room temperature. Primary antibodies were used at a dilution of 1:200 or 1:100 and incubation was carried out for 30 min at 37°C (Table 2.2).

Antibody	Clone	Supplied/Donated by	Monoclona/Polyclonal	Western blotting	Immuno staining
14-3-3 beta	(A-15): sc-17288	Santa cruz biotechnology, Inc., UK	goat pAb	1:1000	
Cbl	(C-15): sc-170	Santa cruz biotechnology, Inc., UK	rabbit pAb	1:1000	
Copine III	#6064	E.Grimm, MD Anderson cancer centre, USA	rabbit pAb	1:2000	1:200
CPS 1		L.Graves, UNC, USA	rabbit pAb	1:1000	
cyclin D2	DCS 3.1	G.Peters, Cancer Research UK, London, UK	mouse mAb	1:500	
cytokeratin 17	(C46): sc-8421	Santa cruz biotechnology, Inc., UK	mouse mAb	1:1000	
ErbB-1 (EGFR)	1005: sc-03	Santa cruz biotechnology, Inc., UK	rabbit pAb	1:2000	
ErbB-1 (EGFR)	Ab-1 (528)	NeoMarkers, Lab Vision Corp., USA	mouse mAb	1:1000	1:200
ErbB-2-Neu	(9G6): sc-08	Santa cruz biotechnology, Inc., UK	mouse mAb IgG1	1:1000	1:200
ErbB-3	(C17): sc-285	Santa cruz biotechnology, Inc., UK	rabbit pAb	1:1000	
ErbB-4	Ab-4 (H3.90.6)	NeoMarkers, Lab Vision Corp., USA	mouse mAb	1:1000	
ezrin	(C-19): sc-6407	Santa cruz biotechnology, Inc., UK	goat pAb	1:1000	
Ku 86	(M20): sc-1485	Santa cruz biotechnology, Inc., UK	goat pAb	1:1000	
lamin B	(C-20): sc-6216	Santa cruz biotechnology, Inc., UK	goat pAb	1:1000	
Lamp-1	(H-228): sc-5570	Santa cruz biotechnology, Inc., UK	rabbit pAb		1:100
large T antigen		P. Jat, LICR, UK	mouse mAb	1:500	
Mn-SOD	# 06-984	Upstate biotechnology, NY, USA	rabbit pAb	1:1000	
MxA		K. Melen, NPHI, Helsinki, Finland	rabbit pAb	1:500	
MxA		O. Haller, Freiburg, Germany	mouse mAb	1:500	
p27		Transduction Laboratories, BD biosciences, Canada	mouse mAb	1:1000	
PGP9.5	Z02361	Affinity research products Ltd., Exeter, UK	rabbit pAb	1:1000	1:100
P-Tyr	(PY99): sc-7020	Santa cruz biotechnology, Inc., UK	mouse mAb IgG1	1:1000	

Table 2.2: List of antibodies used in this work for the validation of differential protein expression analysis.

When double staining was performed, a second incubation with a primary antibody from a distinct species was carried out under identical conditions. Goat anti-mouse antibodies conjugated to tetramethylrhodamine isothiocyanate (TRITC) and goat anti-rabbit antibodies conjugated to fluorescein isothiocyanate (FITC) (Jackson Immunoresearch Laboratories Inc., USA) were used for the detection of primary antibodies. Cell preparations were then mounted on a glass microscope slide with DAKO fluorescent mounting medium (DAKO Corporation, CA, USA). Images were mainly collected with a coolview 12 bit integrating cooled CCD camera (Photonic Science, UK) mounted over an Axiophot microscope fitted with a X63 NA or a X100 1.4 oil immersion objective (Zeiss, UK). All immunocytochemistry collected with the CCD camera were taken under the same magnification. Some images were also acquired with a LSM 510 confocal laser scanning microscope (Zeiss, UK) fitted with a X10 eyepiece, using either X40 NA 1.3 or X63 NA 1.4 oil immersion objectives. Image files were collected as a matrix of 1024x1024 pixels describing the average of 16 frames scanned at 0.062 Hz. FITC and TRITC were excited at 488 nm and 543 nm respectively, and visualized through 540±25 and 608±32 nm bandpass filters, respectively.

2.6.4 Kinase assay

An *in vitro* kinase assay was performed in order to examine the activity associated with Copine III. For this purpose, 400 μg of total cellular extract was incubated for 2 hrs at 4°C with rotation with 5 μl of rabbit polyclonal Copine III antibody and 40 μl of a slurry of protein A coupled to sepharose beads (Amersham Biosciences, UK). Immunoprecipitates were washed in NP40 lysis buffer and then split into two equal fractions. One fraction was kept for analysis by western blotting; the other fraction was used to perform an *in vitro* kinase assay. Assays were carried out in a kinase buffer (25 mM HEPES pH 7.5, 100 mM NaCl, 10 mM $MnCl_2$, 5 mM $MgCl_2$), containing 10 μCi of [γ-^{32}P] ATP and 2 μg of myelin basic protein (MBP) as substrate for 30 min at 37°C. Reactions were terminated with 2x Laemmli sample buffer, boiled and then run on a 1D-SDS-PAGE and transferred onto a PVDF membrane. Membranes were allowed to dry and then exposed to a phosphorimager screen for 24 hrs prior to scanning on an Fx-

phosphorimager (BioRad Laboratories Inc., USA).

2.6.5 Sub-cellular fractionation

Three 15 cm culture dishes of confluent cells were harvested in 2 ml of hypotonic buffer (10 mM HEPES pH 7.4, 5 mM KCl, 5 mM $MgCl_2$) and cells allowed to swell on ice for 30 min prior to homogenisation with a Dounce homogeniser using at least 20 strokes for complete homogenisation. The efficiency of the homegenisation was checked by light microscopy. A crude nuclear fractionation was prepared by centrifugation for 5 min at 600xg using a bench top centrifuge (Biofuge, Heraueus-Kendro Ltd., USA) at 2500 rpm. The pellet, P1, was washed in hypotonic buffer and resuspended in 1x NP40 lysis buffer to a final volume of 2.5 ml. The supernatant was centrifuged at 100,000xg using an Optima ultracentrifuge (Beckman Coulter, Inc., USA) for 30 min enabling the separation of the cytosolic supernatant fraction (S100) and the membrane fraction (P100). P100 was also washed in hypotonic buffer and the final pellet was resuspended in 2.5 ml of 1x NP40 lysis buffer. The S100 fraction was collected and 5x NP40 lysis buffer was added to give a final volume of 2.5 ml. Equal volumes of each fraction were run on 1D-SDS PAGE gels and analysed by western blotting with appropriate antibodies to assess the localisation of specific proteins.

2.6.6 Biotinylation of membrane proteins

To assess the ratio of expression of EGFR at the cell surface compared to the total levels, surface proteins were labelled with EZ-Link Sulfo-NHS-Biotin (Pierce-Perbio, IL, USA). Briefly, cells were serum starved in 0.1% FCS/1640 RPMI for 48 hours and then treated with 1 nM (6 ng/ml) EGF for various times. Prior to lysis, cells were washed twice in 0.5x ice cold PBS and labelled with 500 ng/ml of Sulfo-NHS-Biotin for 15 min on ice. Excess Biotin was removed by several washes and quenched by further incubation with DMEM 10% FCS for 15 min on ice. Cells were then lysed in NP40 lysis buffer supplemented with protease, phosphatase and kinase inhibitors, and lysates were precipitated with streptavidin coupled to agarose beads (Pierce-Perbio, IL, USA).

Incubation was carried out at 4°C overnight and beads were washed four times in NP40 lysis buffer. Protein complexes were eluted in Laemmli sample buffer and analysed by western blotting.

Chapter 3 : TWO-DIMENSIONAL DIFFERENCE GEL ELECTROPHORESIS (2D-DIGE): METHOD DEVELOPMENT AND OPTIMISATION

3.1 Introduction to 2D-DIGE

Global differential protein expression analysis is a powerful tool in defining the processes involved in cancer and in identifying diagnostic and prognostic markers of cancers. Traditionally, 2D gel electrophoresis has been used for large-scale protein separation and comparison of protein expression across samples. In conventional 2DE, gel-separated proteins are visualised mainly using post-electrophoretic staining methods such as silver staining, coomassie staining or non-covalent fluorescent staining (see Chapter 1). Although gel electrophoresis methods have been improved and the use of pre-cast immobilised pH gradient strips allows increased reproducibility, gel-to-gel variation is still a problem. Thus, any differential expression analysis requires the comparison of many replicate gel images, before significant and accurate changes in expression between multiple samples can be assigned. The problem is further compounded by the fact that the most sensitive methods of protein visualisation have a poor linear dynamic range of detection (silver, coomassie), give variable staining (silver) or are expensive and hazardous to use (metabolic radio labelling). To overcome these problems, Minden and co-workers described a novel strategy for differential protein expression analysis using covalent modification of proteins with fluorescent tags (United states patent Minden *et al.*; Unlu *et al.* 1997). In its original form, this new approach allowed the direct comparison of two samples through differential labelling with spectrally distinct fluorescent dyes. Labelled samples are mixed prior to protein separation on the same 2D gel, imaged at different wavelengths, allowing a more accurate determination of protein differences between samples. Another advantage is that these fluorphores have a linear dynamic range of detection that is superior to other non-linear protein detection methods

such as silver staining or coomassie staining (Patton 2000). Therefore, the use of these dyes should increase the sensitivity of detection for low-abundance proteins, as well as allowing accurate quantification of more abundant proteins. In this chapter, a description of the method and its development and optimisation is presented. A workflow is established for further application in the examination of proteomic changes in a model system of ErbB-2-dependent breast cancer.

3.2 Covalent fluorescent protein tagging

In its current form, two dimensional difference gel electrophoresis (2D-DIGE) enables the direct comparison of two or three protein samples on the same 2DE gel. Proteins are covalently modified on lysine residues with spectrally distinct fluorescent cyanine dye molecules prior to their separation by gel electrophoresis (Figure 3.1). Although the use of covalent modification for protein detection is not novel in 2DE gel analysis, in 2D-DIGE, a pair of charge and mass matched tags is used making it possible to label and run different samples on the same 2DE gel, greatly alleviating the problem of gel-to-gel variation. It is easy to appreciate the power of this method as different samples are subjected to the same electrophoretic conditions providing increased reproducibility and accuracy in comparing protein expression levels.

Cyanine dyes were first described by Mujumdar *et al.* (Mujumdar *et al.* 1989; Mujumdar *et al.* 1993). These fluorophors have a general structure which can be modified to create a panel of reactive fluorescent tags. The dyes developed by Unlu *et al.* are N-hydroxy-succinimidyl (NHS)-ester derivatives of the fluorescent cyanines 3 and 5 (NHS-propyl-Cy3, NHS-methyl-Cy5) (Figure 3.2A) (Unlu *et al.* 1997).

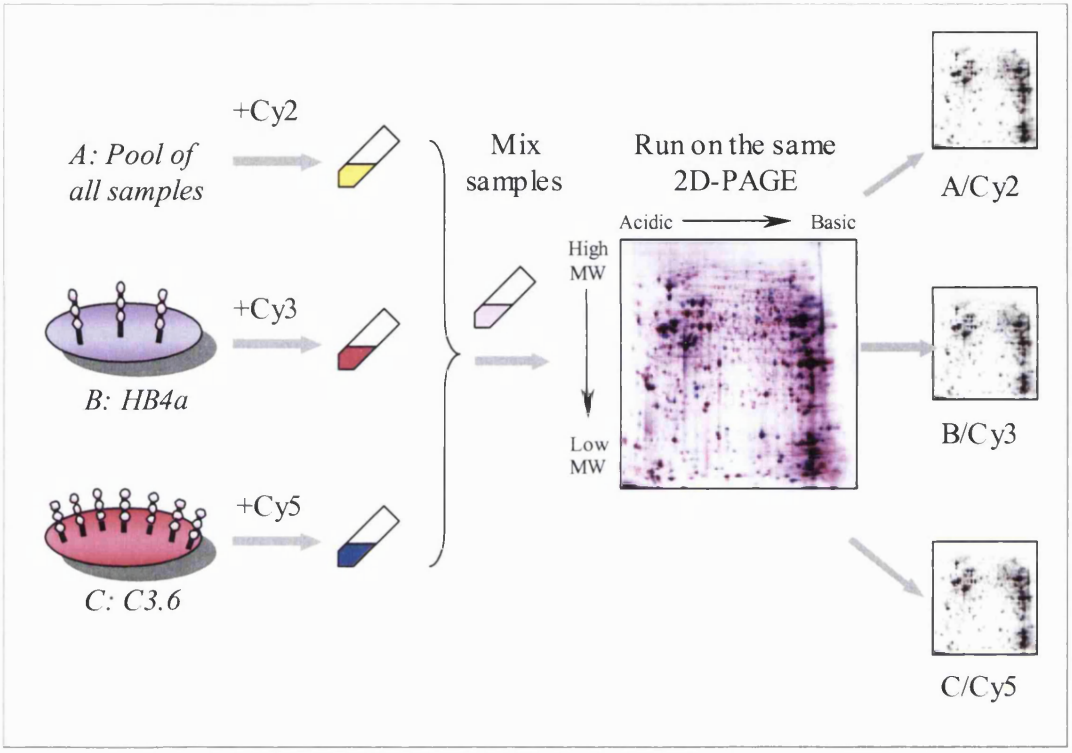


Figure 3.1: The principle of 2D-difference gel electrophoresis (2D-DIGE). Up to three samples can be covalently labelled with a distinct fluorescent cyanine dye (Cy2, Cy3 or Cy5) and mixed prior to protein separation by 2DE. Proteins are visualised by fluorescence imaging using a multi-wavelength fluorescence detection device generating three single, but directly super-imposable images of the three samples analysed.

Three cyanine dyes are now commercially available that have distinct and non-overlapping excitation and emission spectra which are not altered by the derivatisation process (Figure 3.2C). This enables their specific detection with a multi-wavelength detection device. Using these dyes, proteins are covalently labelled on lysine ε–amino groups (ε-NH_2-lys) by nucleophilic attack resulting in the loss of the NHS group (Figure 3.2B). The advantage of using lysine labelling is that almost all proteins contain at least one lysine residue and therefore most proteins will become labelled. The labelling reaction appears to be specific for ε-NH_2-lys lysine groups, although the NHS-ester could also react with the N-terminal α-amino groups of proteins (α-NH_2-Nterm).

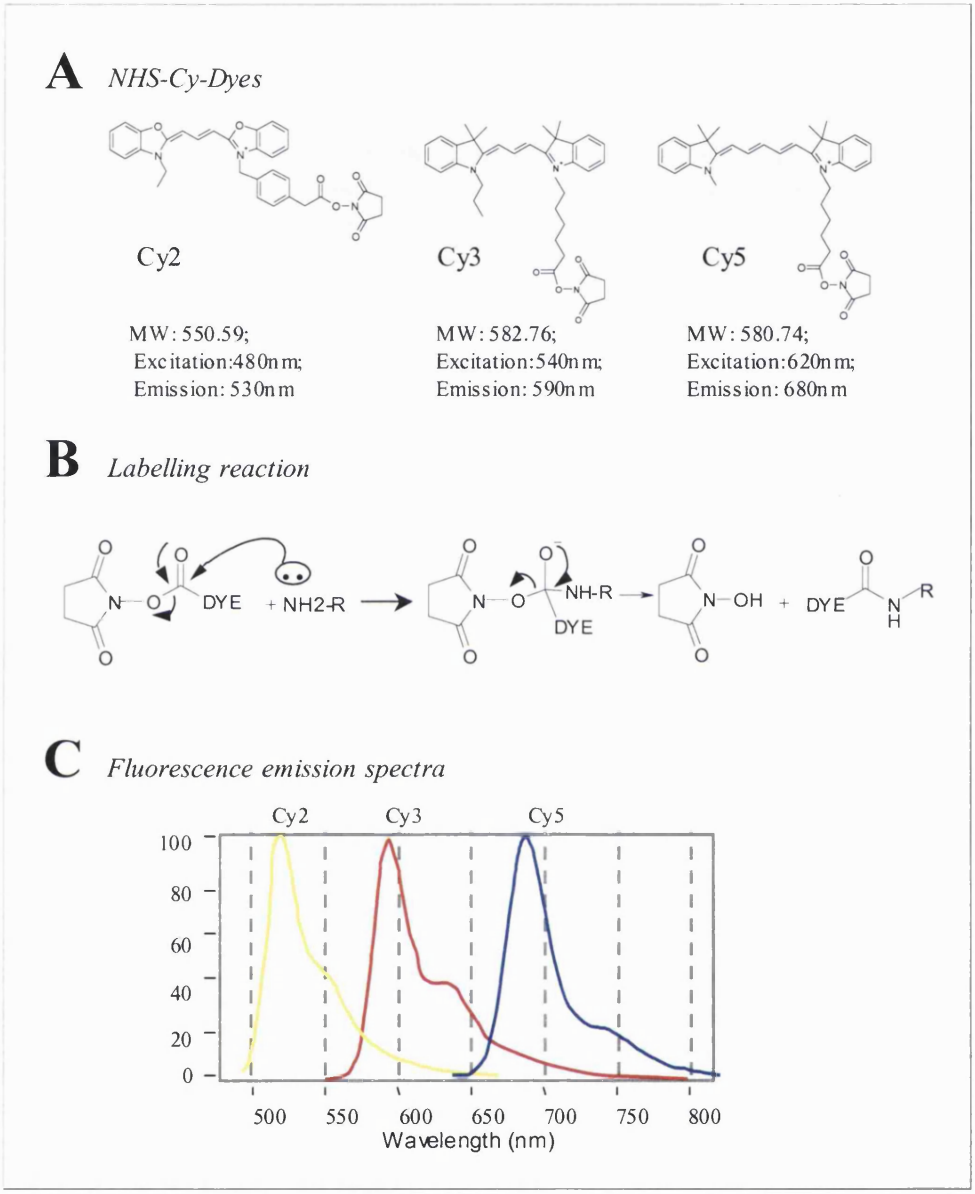


Figure 3.2: Characteristics of the NHS-Cy-dyes. A) Structure of the NHS-Cyanine dyes. Cy2, 3-(4-carboxymethyl)phenylmethyl)-3'-ethyloxacarbocyanine halide *N*-hydroxy-succinimidyl ester, Cy3, 1-(5-carboxypentyl)-1'-propylindocarbocyanine halide *N*-hydroxysuccinimidyl ester, Cy5, 1-(5-carboxypentyl)-1'-methylindodicarbocyanine halide *N*-hydroxysuccinimidyl. Each dye has a similar molecular weight and single positive charge matching the charge of the modified residue. B) The dyes have an *N*-hydroxysuccinimidyl ester reactive group triggering covalent interaction with amine groups of lysine residues. C) Each dye displays a distinct emission spectra, enabling the individual detection of differentially labelled proteins at the appropriate wavelength without overlap of signals.

Conceptually, the dye modification should not perturb the electrophoretic mobility of labelled proteins, and so the dyes should be mass and charge matched. For this reason, the size of the aliphatic chain (Figure 3.2A) was originally modulated to maintain a similar molecular weight between each dye. Additionally, the dyes possess a positive charge which matches the positively charged amino groups they modify. Thus, dye modification should not impair the isoelectrofocusing of the derivatised proteins, an issue that was shown to be problematic in previous work (Urwin and Jackson 1993). Since the original development of the NHS-Cy3 and Cy5 dyes (Unlu *et al.* 1997), and during the work carried out here, a third dye, NHS-ethyl-Cy2 (Cy2) (Figure 3.2A), was introduced, and all three cyanine dyes are now commercially available (Amersham Biosciences). However, in the present study, the NHS-Cy3 and NHS-Cy5 dyes were synthesised in-house by Dr. Piers Gaffney, which enabled a drastic cut in the cost of experiments. Although the original synthesis was briefly reported by Unlu *et al* (Unlu *et al.* 1997), a modified protocol had to be set up to efficiently synthesise these reagents (Gaffney P., unpublished data). The purity of the dyes was assessed (98%) and they appeared to be stable over long periods of storage (data not shown).

3.3 Optimisation of labelling conditions

Conditions had to be established and optimised for labelling of biological samples for 2DE separation and subsequent mass spectrometric identification. Although of relatively low molecular weight, the dyes could affect the electrophoretic mobility of proteins or cause protein precipitation when used in large excess. Indeed, it had been reported that protein precipitation was correlated with the level of labelling reached (Unlu *et al.* 1997). However, the level of labelling has to be sufficient to provide a good sensitivity of detection. For this purpose, a so-called "minimal" labelling strategy was developed, as described by Unlu *et al.* (Unlu *et al.* 1997). Initially, a dye-to-protein ratio of 16 μmol/μg was estimated to be sufficient for minimal labelling of 1-2% of all lysine residues in a whole cell lysate with labelling of an average of 20% of the molecules of a particular protein on one lysine. This should allow for uniform labelling of all proteins without drastically altering their electrophoretic mobilities. Subsequently, further

modifications were introduced to improve labelling efficiency and protein solubility (Tonge *et al.* 2001). These included the removal of DTT in the labelling reaction, which can react with free dye molecules at high concentrations and the optimisation of the pH of the reaction to a more basic environment (pH 8.3), facilitating the nucleophilic attack of the dye on amino groups. These improvements also included a reduction in the amount of dye used to 4 pmol/μg of protein, thereby minimising problems with insolubility and preventing large-scale and heterogeneous mobility shifts. Moreover, improved protein detection was attained using a xenon lamp/CCD type analytical scanning instrument, the 2920 2D-Master Imager, which offered a higher range of sensitivity of detection than the system used by Unlu *et al.* (Unlu *et al.* 1997).

Preliminary work was carried out by Amersham Biosciences to establish the optimal conditions for protein labelling with *E. coli* samples. Hence, there was a necessity to test, optimise and validate 2D-DIGE methodology in the human luminal epithelial cell system to be used in this study. Routinely, HMLEC cells, which are adherent cells, were grown in tissue culture dishes and washed in phosphate buffered saline (PBS) before lysis. It was found that the high concentrations of salt involved had adverse effects on both the protein labelling and the protein separation on a 2D gel. However, low-salt buffers could not be used for washing because cellular osmotic shock could occur. Dilution of the PBS solution (0.5x PBS) improved the labelling and protein separation without affecting the cell lysis. Moreover, the lysis conditions mainly used for 2D gel electrophoresis analysis of these cell lines required the use of 65 mM DTT for sulfhydryl reduction. Free amines as well as DTT or ampholytes are potential quenching agents for the dyes, diminishing the labelling efficiency. Therefore, reduction of proteins was performed after labelling and alternative reagents were used, which would not interfere with the dyes. Thiourea was used as it was shown to provide better solubilisation of proteins for 2DE (Rabilloud 1998), and did not adversely affect dye labelling, so long as the pH was adjusted to pH 8.3, which is optimal for the nucleophilic attack to occur on the amino groups of lysines with facilitated proton exchange (data not shown). Additionally, cells are often lysed in the presence of protease, phosphatase and kinase inhibitors, which could also interfere with the labelling process. Based on the

chemical nature of these inhibitors, a limited inhibitor-cocktail was tested for optimum protein extract and labelling. Routinely, 1mM EDTA, 1 μg/ml pepstatin, 1 μg/ml leupeptin, 17 μg/ml aprotinin, 2 mM sodium orthovanadate, 1 μM okadaic acid were shown to be compatible without producing any detectable effects on labelling efficiency.

Next, several parameters were investigated such as the level of sensitivity reached and the equivalence of protein labelling between each dye. Cyanine dye labelling of proteins was also compared to other commonly used methods of detection for conventional proteomic analysis. Additionally, it was important to check that there was no cross-talk between the channels of excitation and emission of each dye. The ability to isolate proteins of interest for further protein identification was assessed and a phenomenon of apparent shift in molecular weight between the labelled and un-labelled protein fractions analysed was observed. The need to post-stain 2D-DIGE gels was revealed and a compatible post-staining method was established, which would match with the workflow. Subsequently, the compatibility of cyanine dye covalent tagging with further protein identification by mass spectrometry was assessed and an efficient proteomic workflow was further established.

3.4 2D-DIGE is a sensitive method

For 2D-DIGE analysis, it was important to establish the practical level of sensitivity achievable and to compare it to other commonly used methods of detection. Firstly, a standard protein, transferrin, was used to examine minimal labelling levels following previously described conditions and using a dye-to-protein ratio of 4 pmol/μg. This protein was a good control as it contains 58 lysine residues and is of relatively high molecular weight (70 kDa). 50 μg of transferrin were labelled with 200 pmol of Cy3 dye and different dilutions of variable amounts of labelled protein were then analysed by 1D-SDS-PAGE. Protein detection was carried out using the 2920 2D-Master Imager (Amersham Biosciences). Since this imaging device provides edge illumination, the gel could be imaged directly between glass plates following 1DE. Gels were routinely exposed for 20 s for optimum detection of labelled transferrin down to 1-2 ng of loaded protein (Figure 3.3A).

The cross-talk detected during gel imaging between channels was also assessed. Similar to the previous test, 50 μg of transferrin was labelled with Cy2, Cy3 and Cy5 using the labelling ratio given above. Different amounts of labelled transferrin were then loaded onto the same gel and imaged using the 2920 2D-Master Imager at the appropriate excitation and emission wavelengths (Table 2.1, Chapter 2). As shown in Figure 3.3B, no cross-talk between channels could be observed even when gels were overexposed to give a saturated signal. However, this experiment also showed that Cy2, gave a stronger background signal. In addition, a tiling effect occurred when the gels had to be exposed for long periods of time for improved detection.

Protein detection levels were compared between the cyanine dye labelling method and other established sensitive methods of detection which are based on metal ion reduction with silver. For this purpose, the same amount of Cy3-labelled transferrin was compared to un-labelled transferrin stained with either silver nitrate (according to Shevchenko *et al.* 1996), a mass spectrometry-compatible superficial staining, or by diamine silver staining (according to Hochstrasser and Merril 1988). This experiment showed that both silver staining methods and DIGE labelling were equally sensitive and that 2-4 ng (e.g. 25-50 fmol) of loaded protein could be detected (Figure 3.3C). Although the silver staining method described by Hochstrasser *et al.* showed slightly better sensitivity of detection than DIGE labelling, the use of a cross linking agent (glutaraldehyde) makes this method incompatible with mass spectrometry and therefore it cannot be used for proteomic analysis where proteins are to be picked from gels for MS identification.

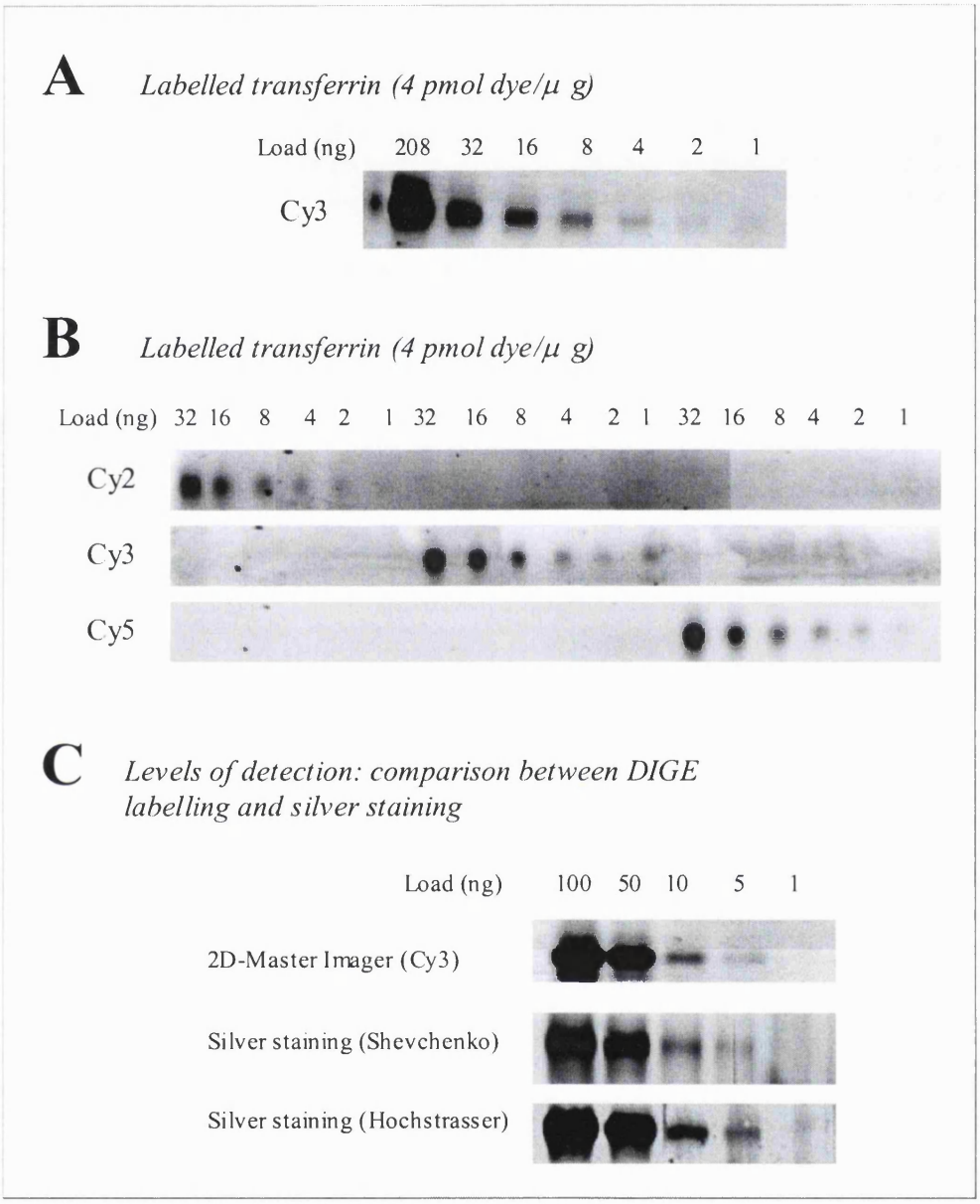
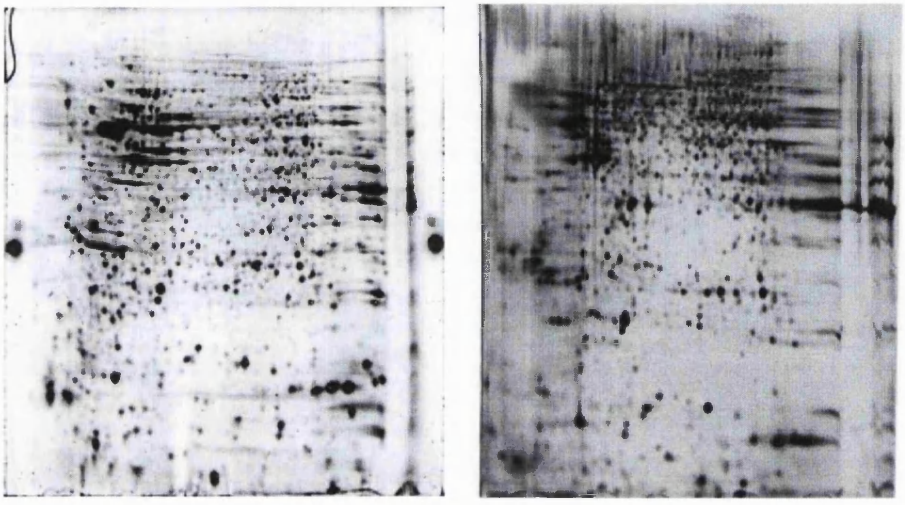


Figure 3.3: Sensitivity test of cyanine dye-labelling. A) 50 μg of a standard protein, transferrin, was labelled with 200 pmol of Cy3, using the labelling conditions described in Chapter 2. Serial dilutions of labelled transferrin were loaded on a 12 % SDS-PAGE gel and imaging was carried out using the 2920 2D-Master Imager. B) Transferrin labelled with Cy2, Cy3 or Cy5 was only detected at specific emission and excitation wavelengths. C) Levels of sensitivity were compared to two other protein detection methods: superficial silver staining (Shevchenko) and diamine silver staining (Hochstrasser). These methods are described in Chapter 2.

This data shows that the method of DIGE labelling has sufficient sensitivity for the detection of low amounts of protein and is comparable to silver staining. However, as seen in Figure 3.3C, superficial silver staining (Shevchenko staining method) is a non-linear detection method and suffers from the problem that high amounts of protein are labelled to saturation, unlike the DIGE-labelled proteins. Additionally, in comparison to silver staining methods, DIGE labelling and detection was a more rapid technique, which did not give rise to high background and other staining artefacts often encountered with silver staining methods (Figure 3.3C).

Following on from this study, the global protein detection efficiency of DIGE was tested using mammalian cell lines. Protein lysates of HMLECs were prepared in a lysis buffer compatible with cyanine dye labelling (CHAPS/urea/thiourea/Tris pH8.3) (Chapter 2). Briefly, 100 μg of the same cell lysate were labelled with Cy3 and Cy5 under "minimal" labelling conditions (4 pmol/μg) and the samples were mixed prior to protein separation on the same 2DE gel. Protein detection patterns were then compared between cyanine dye labelled, silver stained and SYPRO Ruby fluorescent protein stained samples by running three different gels (Figure 3.4). Qualitatively speaking, the Cy-dye and SYPRO Ruby images were most similar and features were sharp unlike silver staining which produced somewhat streaky images, and smeared gel features, particularly in the high molecular weight range. Some proteins were preferentially detected with one staining method over another, as shown by the differential staining pattern (Figure 3.4). This was most likely due to the different chemistries involved.

Additionally, as expected, spot intensities for the same proteins varied between the three detection methods as shown by comparison of grey scale intensities (data not shown). Proteins were also equally well detected from the low to high molecular weight range of the 2DE gel, therefore suggesting that a diverse array of proteins can be cyanine dye labelled with no preferential labelling conferred by individual protein sequences.



Sypro Ruby staining (200 µg)

Shevchenko silver staining (200 µg)

2D-DIGE labelling
(Cy3 image from 100 µg)

Detection method	No. of features
Cy3 labelling	1278
SyproRuby staining	1357
Silver staining	1338

Figure 3.4: Comparison of three protein detection methods for 2DE. Identical samples were run on three distinct 2D gels and either covalently labelled prior to gel electrophoresis with Cy3 or stained after gel electrophoresis by silver staining or fluorescent SYPRO Ruby staining. The inset table shows the number of features detected in each gel using Melanie III software.

For accurate comparison of protein detection patterns, each gel image was curated using Melanie III software and spot numbers were compared (Figure 3.4). The protein features were manually curated and all features below a normalised percentage volume of 0.01% were omitted, as were dust particles and background staining. This manual gel image analysis had to be undertaken as gel images from distinct detection methods were being compared. Results revealed that a similar number of protein features were detectable by silver staining, SYPRO Ruby staining and DIGE labelling with approximate values of 1338, 1357 and 1278 protein features being detected respectively. Overall, this work confirmed the efficiency of cyanine dye labelling for proteomics analysis and showed that the method was faster than the other two techniques tested. Furthermore, it provided sensitive detection of proteins separated by 2DE which was comparable with silver staining and SYPRO Ruby staining.

3.5 2D-DIGE is a highly reproducible and accurate proteomic technique

The next step was to verify the reproducibility in labelling with one dye versus another which was crucial in the validation of the 2D-DIGE analysis system. To test this, two aliquots of the same sample were labelled with Cy3 and Cy5 and both labelled samples were run on the same 2D gel. Gel images were initially pseudo-coloured and merged using the Adobe Photoshop software (Figure 3.5A). Each image had an identical protein pattern and no major changes could be detected by eye. A more detailed gel image analysis was then applied using DeCyder software, which is optimised for 2D-DIGE-image analysis (described in Chapter 2 and in more detail in Chapter 4). This analysis revealed that an identical sample labelled reciprocally with Cy3 and Cy5 and run on the same 2DE gel had an almost identical protein detection pattern (Figure 3.5). Out of 3054 protein features detected, 3012 displayed a similar volume ratio between a +/- 1.23 fold threshold, which represents a standard deviation (SD) of 1. Thus, the difference in fold ratio between the dyes was below 1.23 fold and hence this value could be considered as a cut off threshold in further experiments.

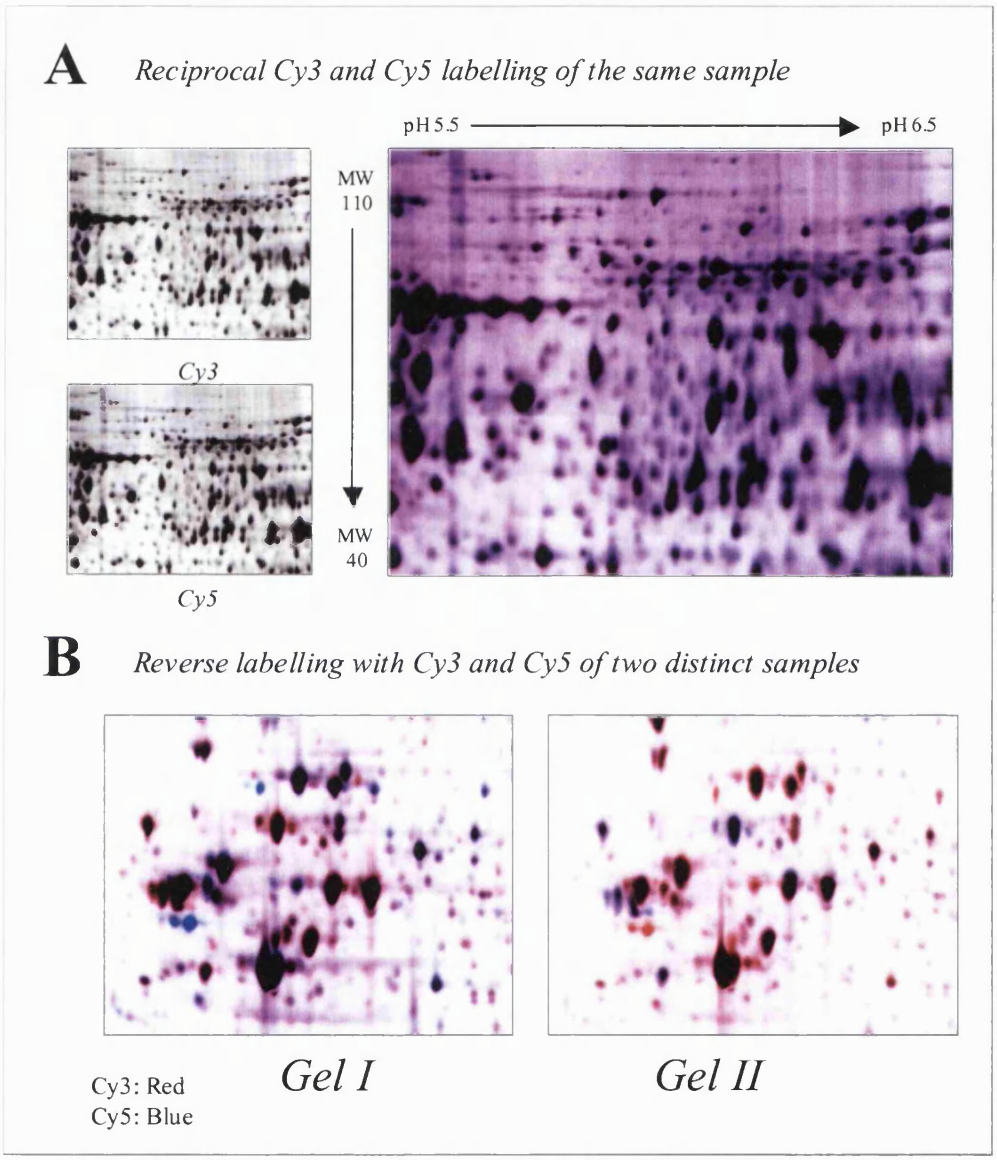


Figure 3.5: Equivalence of 2D-DIGE labelling and detection of differences in the protein expression profiles of HB4a and C3.6 cell lines. A) Identical cellular lysates from HB4a cells were labelled with Cy3 and Cy5 following the conditions described in Chapter 2. Mixed samples were then run on the same 2DE gel. Images were pseudo-coloured and merged using the Adobe Photoshop software. B) Two cell lysates, HB4a and C3.6 were labelled with Cy3 and Cy5 respectively, mixed and run on the same 2DE gel (Gel I) and also labelled reversibly (Gel II). Gels were imaged on the 2920 2D-Master Imager and exported images superimposed using Adobe Photoshop software. Proteins displaying identical intensities appear in black, proteins expressed in one sample compared to the other appear in red (Cy3) or blue (Cy5).

This showed that Cy3 and Cy5 can label proteins in a reproducible manner as 98% of total detected proteins had protein expression volumes falling within the SD. The detailed comparison between HB4a reciprocally labelled with Cy3 and Cy5 could be represented on a scatter plot (Figure 3.6).

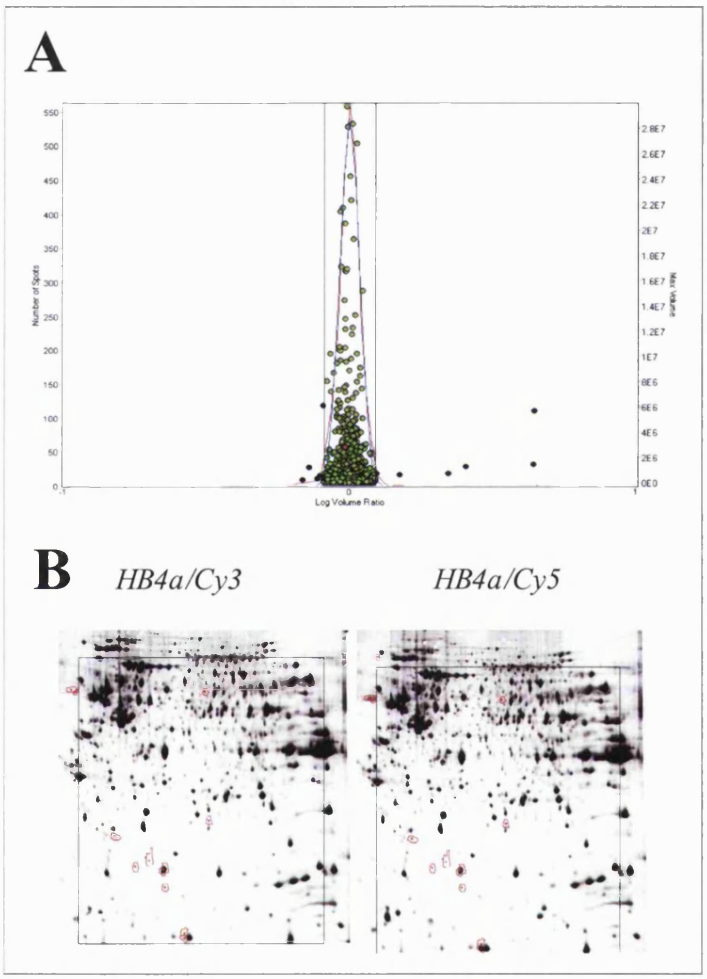


Figure 3.6: Cy3 and Cy5 give reproducible protein detection patterns. Two aliquots of HB4a cellular lysate were labelled with Cy3 or Cy5, mixed and run on the same 2DE gel. Gel images were then curated using the DIA module of the DeCyder software. A) A scatter plot of spot volume and number versus volume ratio is shown with a threshold ratio set at 1.23 (+/- 1 SD) (vertical lines). B) Cy3 and Cy5 separated gel images. Protein features outside the threshold ratio of 1.23 are circled in red.

Protein features, which did not fall within the standard deviation, were considered as significantly differentially expressed between the samples. As a result, this showed that most protein features were considered similar and the proteins with a statistically different protein expression profiles were mostly very low abundance proteins. 2D-DIGE was further tested for sensitivity and reproducibility for the detection of differences in protein expression between the parental HB4a and the derivative ErbB-2 overexpressing HMLEC cell line, C3.6. A reciprocal labelling strategy was employed: two identical samples per cell line were reversibly labelled with Cy3 and Cy5 and mixed with their complementary dye partner. Two 2D gels were run and protein expression patterns and volume ratios of features between the cell lines were compared.

By visual comparison, very similar patterns of protein features were evident, although a number of differences could be detected in both images, indicating significant and reproducible differences in the protein profiles of the two cell lines (Figure 3.5B). This also indicates that there was no preferential labelling with one dye over the other. This comparison of Cy3 and Cy5 labelling was carried out at a time before Cy2 became commercially available. However, later experiments showed that Cy2 behaved in much the same way as the other dyes for labelling although a higher background was observed for Cy2 labelling (data not shown).

Recently published data from other laboratories has reflected the results obtained herein and confirmed the importance of 2D-DIGE labelling for reproducible and accurate statistical analysis. One detailed study was carried out by Tonge *et al.* to validate the use of 2D-DIGE for statistical protein expression analysis (Tonge *et al.* 2001). The authors attempted to evaluate the quantitative variability in 2D-DIGE analysis. The work was focused on the reproducibility in cyanine-dye labelling between each of the three available cyanine dyes. Variation in spot quantitation between gels and between cyanine dye-labelled samples was assessed. An identical sample was labelled with each one of the available dyes and the three-dye-labelled sample was run on the same 2D gel. An evaluation of spot intensities was done, revealing a high correlation between gel images. Ideally, two images corresponding to the same sample should present a normalised spot volume ratio of 1.00. However, this work showed ratios between identical protein spots

varying from 1.13 to 4.27, with smaller volume spots giving rise to larger ratios, hence a poorer correlation. However, the same variability was observed for all three dyes.

Reproducibility in labelling from separate runs was also assessed in this current study during an initial characterisation of the HB4a and C3.6 cell lines and showed that differential protein expression was reproducible from gel-to-gel but also from experiment-to-experiment (as presented in the next Chapter). Furthermore, for large-scale protein expression analysis, multiple gels are cross-compared and for accurate statistical analysis, it is crucial to adopt a technique which enables reproducible 2DE gel analysis. Although 2D-DIGE increases the accuracy in the comparison of paired samples, gel-to-gel variations still represent an important caveat in the comparison of biological conditions across time points or samples. However, 2D-DIGE can also be used with an internal standard run together with the analytical samples and used in order to circumvent problems associated with gel-to-gel variation. This internal standard or reference sample is run on all gels, enabling the normalisation of each matched spot through a common reference spot for better quantitation of intensities and statistical analysis. This could be achieved by labelling a pool of all samples with one dye, for instance Cy3 and each distinct sample with Cy5, and running paired samples on the same 2D gel. This type of work was described by Yan *et al.* where fluorescent 2D-DIGE was validated by applying it to the study of the known effects of benzoic acid treatment on the proteome of *Escherichia coli* (Yan *et al.* 2002). With the introduction of the third dye, Cy2, this allowed the standard to be run on each gel with two distinct samples labelled with Cy3 and Cy5. Thus, differences between Cy3 and Cy5 labelled samples can be linked through the internal standard labelled with Cy2 as ratios Cy3/Cy2 compared to Cy5/Cy2 across all gels. Because Cy2 gave a stronger background than Cy3 and Cy5, it was preferentially used to label the standard pool of samples. Next, a large-scale 2D-DIGE study was carried out to analyse the effect of growth factor treatment on protein expression patterns in the HB4a cell model system and this work will be further described in Chapter 4.

3.6 Post-staining of 2D-DIGE gels

The rationale for using this technique of 2D gel electrophoresis is to increase the separation of a complex protein mixture in order to detect changes in their level of expression and to allow the identification of the proteins of interest using an analytical approach such as mass spectrometry. Although mass spectrometry is in itself a very sensitive approach, it is important to collect a pure and MS-compatible sample for optimum characterisation. Since cyanine dye labelling involves minimal labelling of proteins, one can expect that the majority of proteins present on a 2D gel are un-labelled. However, due to the added mass of the dye molecule (approximately 550-580 Da), a shift in molecular weight may be apparent between the modified and the unmodified molecules. The effect of the protein shift was analysed by comparison of gel images from a cyanine dye labelled gel and its fluorescent SYPRO Ruby post-stained image. As shown in Figure 3.7, this shift was more prominent in the lower mass range of the gel, as expected, due to the ratio of the added mass of a dye molecule compared to the mass of the protein. The shift was also non-linear across the whole molecular weight range of the gel, potentially complicating any statistical estimation of the shift for "blind" automated spot picking. For further mass spectrometric characterisation of proteins of interest, it was therefore important to find a way to accurately pick spots representing the majority of un-labelled protein. Thus, this work shows that post-electrophoretic staining of the gels cannot be avoided.

Each visualisation method tested enabled protein detection and no reverse staining was observed, suggesting that the covalent modification of proteins on their lysines did not alter their potential to be stained with any of the methods tested. As referred to previously, the best-matched protein detection pattern was encountered using SYPRO Ruby as a post stain. As shown in Figure 3.8, despite the differences in protein labelling and the shift in molecular weight observed between DIGE labelling and SYPRO Ruby staining (Figure 3.7), most protein features detected by DIGE could be seen on the SYPRO Ruby post-stained gel image.

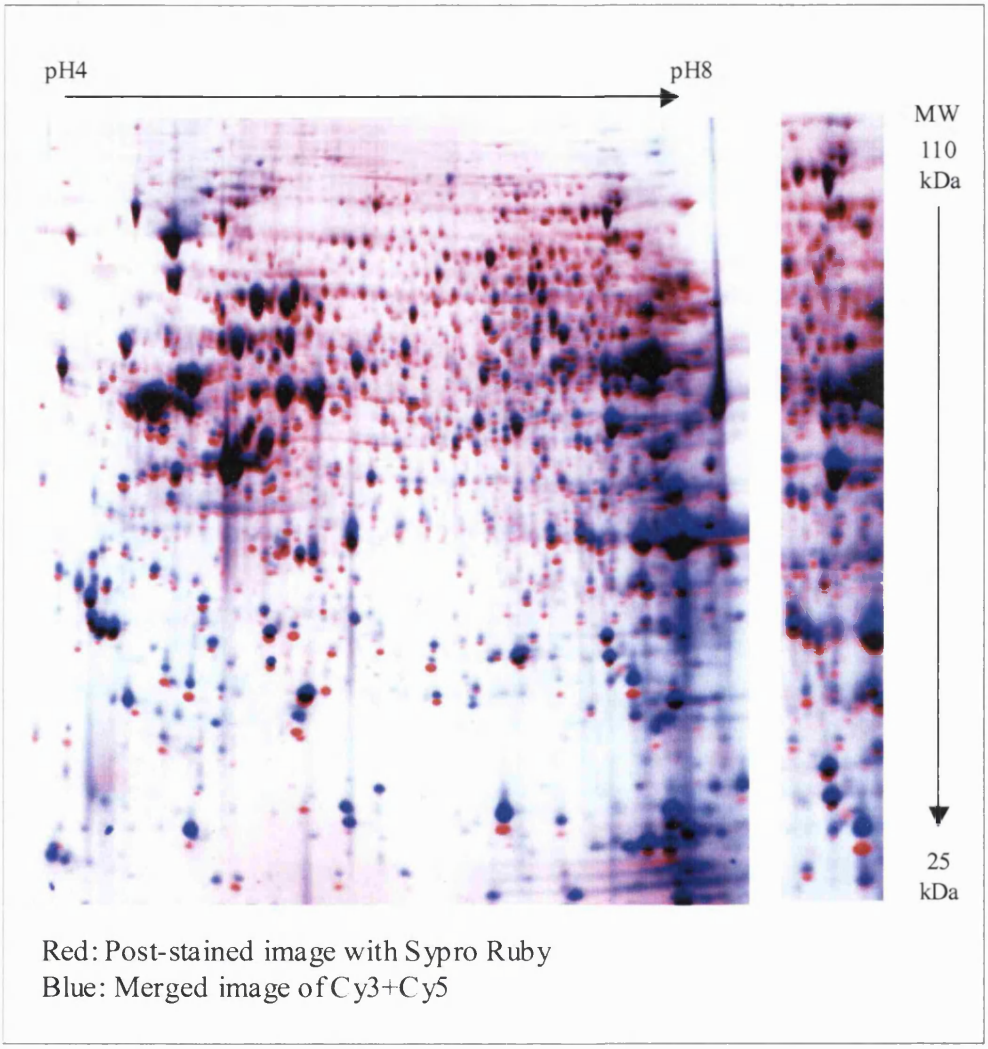


Figure 3.7: Evaluation of the protein shift in the second dimension caused by protein labelling with cyanine dyes. HB4a cell lysate was labelled with Cy3 and Cy5 dyes following minimal labelling. The DIGE image was merged (blue). The same 2DE gel was then post-stained with SYPRO Ruby stain (red). The protein pattern was compared by overlaying images. A shift in molecular weight could be observed between the cyanine dye labelled proteins and the un-labelled proteins, and this was more prominent in the lower molecular weight region of the gel. The right-hand panel represents an enlarged portion of the gel image.

An accurate comparison of protein detection capability between fluorescent SYPRO Ruby staining and DIGE labelling was carried out using MelanieIII software, revealing that 1.4 +/- 0.1 times more protein features could be detected using SYPRO Ruby staining than cyanine dye labelling. Thus, SYPRO Ruby labelling would appear suitable for 2D-DIGE post-staining.

To facilitate gel picking accuracy and speed (and for further high-throughput analysis), an automated gel picker was employed. Optimally, gels were bonded to one of the glass plates using a Bind Silane solution (Plus One, UK) prior to casting. This strategy prevents movement of the gels during picking. However, it was discovered that superficial MS-compatible silver staining did not efficiently stain the bonded gels and the sensitivity reached was greatly diminished since only one surface of the gel is stained. This made silver staining a poor DIGE post-staining method if automated picking was required. On the other hand, fluorescent SYPRO Ruby staining of bonded gels provided a similar sensitivity of detection to that seen with non-bonded gels. Thus, fluorescent SYPRO Ruby stain was consequently used for all 2D-DIGE post-electrophoretic staining in the course of this study.

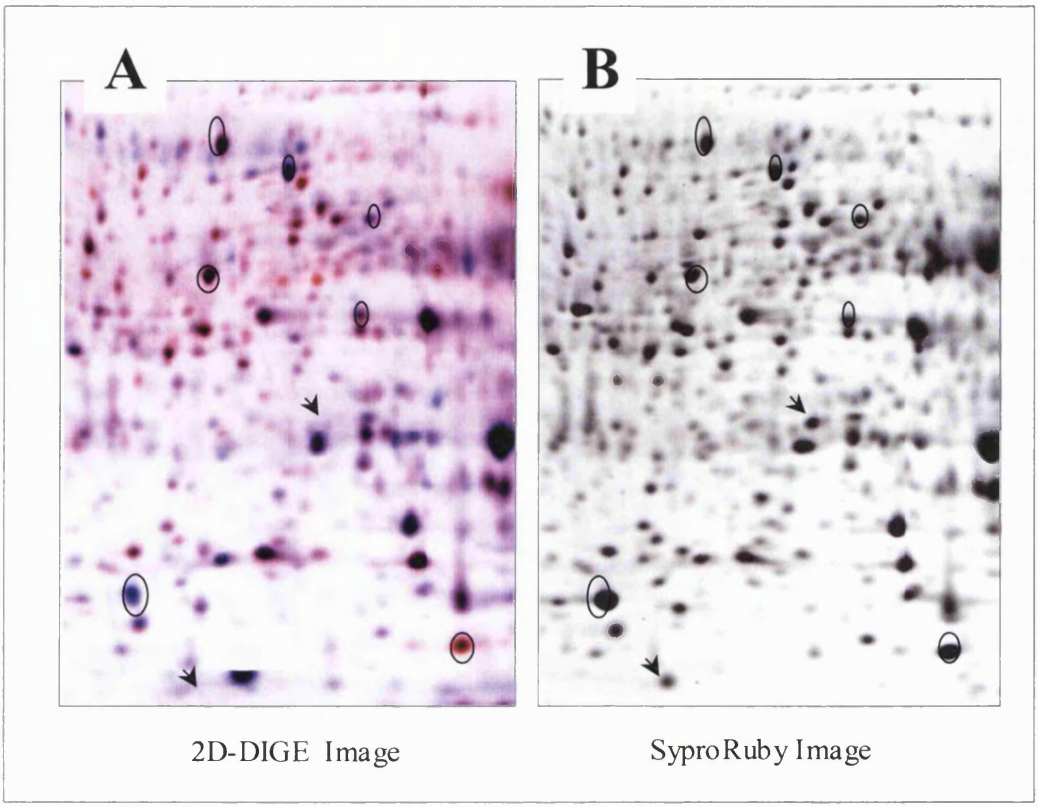


Figure 3.8: Assessment of SYPRO Ruby staining as a DIGE post-staining method. A) HB4a and C3.6 cells were minimally labelled with the Cy3 and Cy5 dyes respectively and gel images were merged using Adobe Photoshop software. B) The DIGE gel was fixed and post-stained with SYPRO Ruby and imaged using the 2920 2D-Master imager. Both protein detection patterns were compared and most proteins could be matched across gels images. Landmarked spots are shown with the circles. Several proteins were preferentially detected with SYPRO Ruby over cyanine-dye labelling (shown with the arrows).

3.7 Compatibility with downstream mass spectrometric identification

For differential expression analysis, it was not only important to find an appropriate analytical protein detection method, which is highly reproducible, sensitive and quantitative, but also to find an approach compatible with the subsequent identification of differentially regulated and expressed protein spots. The main approach used for the identification of proteins from gels is mass spectrometry (MS). Routinely,

protein identification is carried out by proteolytic cleavage of proteins into peptides, which are then analysed by MALDI-MS.

However, this procedure requires that any pre-treatment of samples is compatible with further MS analysis. Thus the compatibility of DIGE labelling and SYPRO Ruby post-staining with MS analysis was further assessed. Although it was shown that some staining methods are incompatible with MS, SYPRO Ruby staining had been previously reported to be a mass spectrometric compatible detection method (Lopez *et al.* 2000). Because cyanine dye labelling involves covalent modification of proteins on their lysine residues, this could interfere with protein proteolysis. Indeed, trypsin, the favoured proteolytic enzyme for protein digestion, cleaves at the C-terminal side of lysine and arginine residues and consequently the dye molecule may inhibit trypsin cleavage by conferring a steric hindrance. The labelling conditions used here involve minimal labelling of proteins, and therefore a low percentage of lysine modifications are expected. As discussed in more detail in the next chapter, protein identification did not seem to be affected by covalent modification of proteins, thus showing the compatibility of 2D-DIGE with mass spectrometric analysis. This was also reported in a recent work carried out by Yan *et al.* (Yan *et al.* 2002). Thus, 2D-DIGE seems to be a good detection and quantitation method that is compatible with MS-based proteomic analysis.

3.8 Conclusions and Discussion

In this Chapter, a detailed evaluation of the 2D-DIGE technique for comprehensive proteomics analysis was performed. It was shown that a combination of the three mass-and-charge-matched fluorescent dyes, Cy2, Cy3 and Cy5 could be used for sensitive, non-biased, accurate and reproducible analysis of differential protein expression analysis. The levels of sensitivity were established for a standard protein, transferrin, and up to 1-2 ng of loaded sample could be detected on a 1DE gel. Comparison with other detection methods currently used for 2D gel analyses revealed that DIGE labelling competes well with these methods enabling sensitive protein detection. Additionally, the specificity of excitation and emission of the dyes was shown to be distinct and there was no cross-talk between channels during detection. Importantly,

reproducibility was validated and identical protein labelling was achieved with each dye suggesting that there was no preferential labelling by different dyes.

The main improvement provided by this approach was that the comparison of two to three samples could be achieved on the same two-dimensional gel, therefore avoiding the problems associated with gel-to-gel variation. In addition, the use of DIGE analysis enabled the cross comparison between gels providing the use of a third fluorescent dye as an internal standard (pool of all samples) allowing accurate statistical analysis of multiple samples, as applied in a biological experiment and described in Chapter 4. Consequently, the DIGE approach was an important improvement for the comparison of multiple samples for 2DE-based differential protein expression analysis. In addition, this method provided a more rapid detection, since images were obtained directly after gel electrophoresis and did not require prolonged fixing and/or long staining incubations. Importantly, the same extent of labelling was reached in replicate labelling experiments and an identical detection pattern was obtained. Conversely, such labelling was not achieved with the use of silver staining and coomassie staining.

However, it was still necessary to use a post-staining approach for accurate spot picking and subsequent protein identification due to the apparent shift in molecular weight which was attributed to cyanine dye addition. The SYPRO Ruby protein stain appeared to be the most adequate post-DIGE staining method as the protein detection pattern was most similar between DIGE labelling and SYPRO Ruby staining. Post-staining required cyanine-dye and SYPRO Ruby stained images to be matched, representing an additional step in the workflow and a potential caveat in the technique. However, images could be easily overlaid and differentially expressed proteins identified from SYPRO Ruby post-stained images for spot picking. It would ultimately be advantageous to bypass this additional step of secondary staining, which is time consuming as well as limiting, since not all proteins can be matched and therefore picked for further mass spectrometric analysis. One objective would be to perform saturation labelling of proteins to achieve a direct correspondence of protein mass and labels and to make this possible alternative covalent tags could be used. Conditions are currently being optimised in our laboratory for optimum saturation labelling of cysteine-containing

proteins. A pair of dyes have already been synthesised by Dr. Piers Gaffney, ICy3 and ICy5, and the compatibility of such saturation labelling with 2DE and subsequent protein identification by mass spectrometry is under investigation.

In summary, these results demonstrate the power of 2D-DIGE for proteomics analysis. Application to the analysis of HMLECs cell lines was assessed, optimum labelling was determined, and reproducible differences in protein expression profile between HB4a and C3.6 cells could be detected. A detailed 2D-DIGE analysis in the study of ErbB-2 overexpression on the protein expression profile of HMLECs cell lines is discussed in the next Chapter.

Chapter 4 : 2D-DIGE PROTEOMIC ANALYSIS OF ErbB-2-DEPENDENT EXPRESSION IN A MODEL BREAST CANCER CELL SYSTEM

4.1 Introduction

In the previous Chapter, the potential of 2D-difference gel electrophoresis (2D-DIGE) analysis for sensitive, reproducible and accurate quantitative analysis of differential protein expression was shown (Chapter 3). However, its usefulness as a proteomic technique needed to be tested in a biological context. In this Chapter, the utility of 2D-DIGE for the differential expression analysis of cancer cells in cultures is reported and this technique was further evaluated in the study of the proteomic effects of ErbB-2 overexpression in a model breast cancer cell system. As previously described, ErbB-2 represents an important marker of breast cancer cells as it is found overexpressed in 25 to 30% of human breast cancers. However, the impact of ErbB-2 overexpression on the mechanisms involved in the development of cancer remains unclear (see Chapter 1). The aim of this work was to investigate the effect of ErbB-2 overexpression on global protein expression and to assess the efficacy of 2D-DIGE analysis for such biological study. This study was carried out using a relevant model cell system of ErbB-2 overexpression by direct 2D-DIGE comparison of a parental breast luminal epithelial cell line HB4a, and an ErbB-2 overexpressing derivative, C3.6 (detailed in Chapter 1).

The identification of differentially expressed proteins was then performed by MALDI-MS analysis and the protein identities are discussed as potential downstream targets of ErbB-2 overexpression in breast cancer. Furthermore, the effects of growth factor treatment (Hrgβ1 and EGF ligands) on the protein expression profile were also addressed in the context of ErbB-2 overexpression and cellular signalling. Subsequently, the comparison between the protein profiling data and a parallel transcription profiling study of the HB4a and C3.6 cell system was carried out.

4.2 ErbB-receptor expression and cell morphology of HB4a and C3.6 cells

The study of the effects of ErbB-2 overexpression on the proteome was carried out by differential analysis of protein expression in human mammary luminal epithelial cells (HMLECs), the HB4a and C3.6 cell lines. Previous studies from our laboratory have shown that the ErbB-2 overexpressing clone (C3.6) has increased responsiveness to mitogenic signalling and a certain level of growth factor independence compared to the HB4a line (Harris *et al.* 1999; Timms *et al.* 2002). In this study, ErbB-2 overexpression was shown to result in an increased and prolonged activation of the downstream signalling pathways of the ErbB receptors, most notably through the activation of the ERK1/2 MAPK pathway. Growth analysis revealed a higher proliferation rate in the C3.6 cell line compared to the parental line, indicating an effect of ErbB-2 on the cell cycle of these cells (Timms *et al.* 2002). The levels of ErbB-2 expressed at the cell surface were previously measured by FACS analysis and compared to other well-defined breast cancer cell lines overexpressing ErbB-2 (BT474 and SKBr3) (Harris *et al.* 1999). From these results, the level of overexpression of ErbB-2 in the C3.6 cells was estimated to be approximately 10-fold more than that of the parental cell line, HB4a, whilst another clone, C5.2, expresses approximately 100-fold more ErbB-2 than the parental cell line, and this was used to further validate potential ErbB-2-dependent targets. The C3.6 cell line was chosen for this study as it expressed levels of ErbB-2 similar to that seen in human breast tumours. Although the C5.2 cells would have been a good control for very high expression of ErbB-2, this clone was not chosen for this proteomic study as it seemed to have two sub-cellular populations, with a luminal phenotype and a fibroblastic phenotype.

The total cellular levels of ErbB-2 were estimated by western blotting in the three cell lines, HB4a, C3.6 and C5.2 (Figure 4.1A). This data showed that the levels of ErbB-2 were similar to the levels estimated previously by fluorescence activated cell sorting FACS analysis (Harris *et al.* 1999). The levels of other ErbB family members (EGFR and ErbB-3) were also examined (Figure 4.1A).

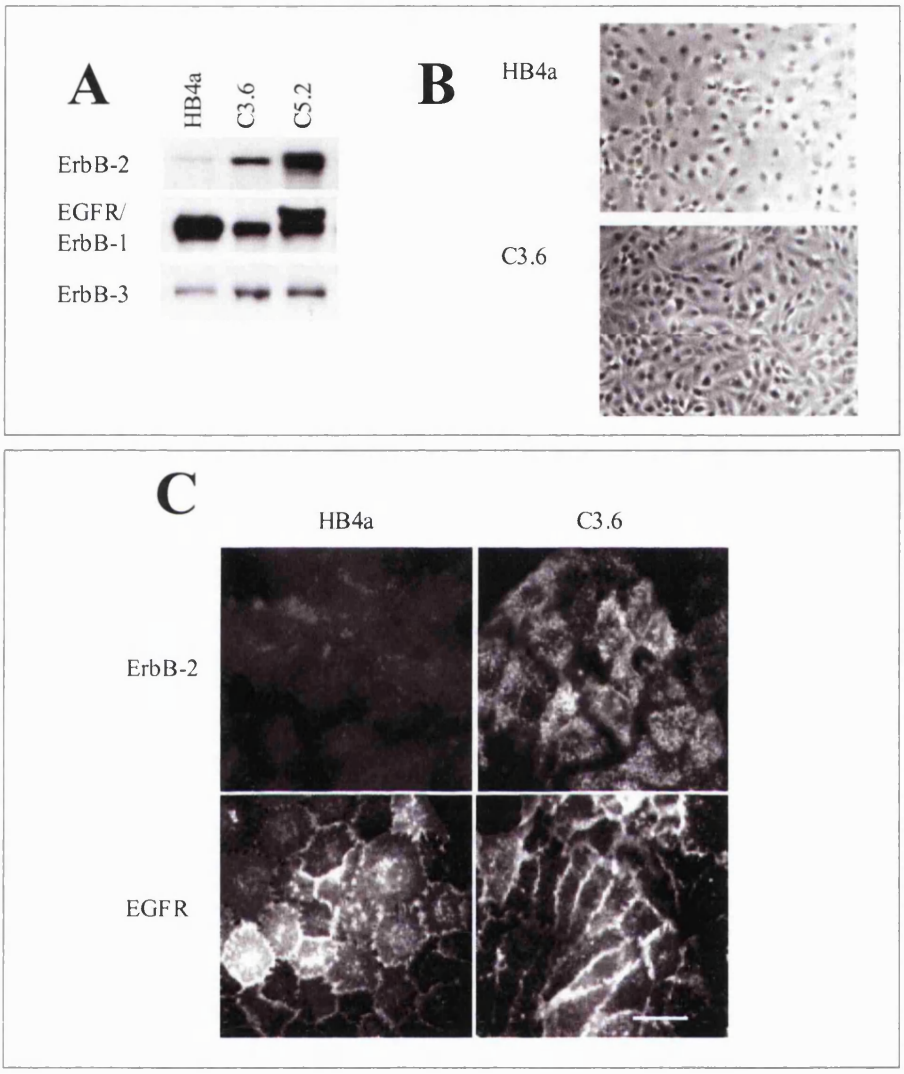


Figure 4.1: Characterisation of HB4a and C3.6 cell lines. A) ErbB-receptor tyrosine kinase expression levels in the parental cell line HB4a and the ErbB-2 overexpressing clones, C3.6 and C5.2 were estimated by immunoblotting with specific antibodies. ErbB4 was not detectable in these cell lines (data not shown). B) The morphology of the HB4a and C3.6 cells was compared by phase contrast microscopy. C) The cellular localisation of ErbB-2 and EGFR was assessed by immunostaining of HB4a and C3.6 cell lines (bar 150 μm). Due to low expression levels in the HB4a cells, ErbB-2 could not be detected.

High levels of EGFR were detected in the parental cell line, which is representative of HLMEC cells as previously described (Bates *et al.* 1990; Burke and Wiley 1999). Importantly, the levels of EGFR and ErbB-3 were respectively down-regulated and up-regulated in the ErbB-2 overexpressing cell lines, suggesting that ErbB-2 overexpression could alter the levels of other ErbB receptors. As discussed in Chapter 1, these changes are likely to influence the heterodimerisation pattern of the ErbB family and hence the specificity of downstream signalling in the two cell lines. To further clarify this point, the mechanisms of regulation of the ErbB family members were further investigated and results are shown in Chapter 6.

Cells were also imaged by phase contrast microscopy and their morphological characteristics established. The C3.6 cells were found to be more elongated and tightly packed than the HB4a cells (Figure 4.1B). These data suggest that ErbB-2 overexpression results in an altered cellular phenotype, manifested as a change in cellular morphology and possibly in altered cell-cell interactions. This correlates with previous reports of ErbB-2 overexpression where proteins involved in adhesion to the matrix and other cells (integrin α2 and E-cadherin) were found down-regulated in fibroblastic clones (Baeckstrom *et al.* 2000). Immunofluorescence was also carried out to assess the cellular distribution of ErbB-2 and EGFR in the HB4a and C3.6 cells (Figure 4.1C). HB4a and C3.6 cells were serum-starved and fixed cells were permeabilised and stained with anti-ErbB-2 and anti-EGFR antibodies and with TRITC-coupled mouse secondary antibody. Images were captured by fluorescence microscopy (Chapter 2). This showed that ErbB-2 was highly expressed in the C3.6 but could not be detected in the parental line. This receptor was mainly localised in the microvilli of the plasma membrane. On the other hand, EGFR was localised at the cell-junction with most receptors expressed at the membranes. This distribution of EGFR correlated with a previous report and is characteristic of a luminal epithelial cell phenotype (Burke and Wiley 1999). The cellular distribution of each receptor will be further analysed and discussed in Chapter 6.

4.3 2D-DIGE experimental design

A differential global protein expression analysis of HB4a and C3.6 was carried out by 2D-DIGE to establish ErbB-2-dependent changes in protein expression that may account for changes in cellular phenotype. The effect of ErbB-2 overexpression in luminal epithelial cell lines was first investigated under basal conditions. This was carried out under serum-starved conditions to optimise the synchronisation of the cells in the cell cycle. It had previously been shown however that even under these conditions, the two cell lines show differences in morphology and proliferative capacity (Harris *et al.* 1999), suggesting that differential protein expression occurs in cells maintained in the absence of serum.

In these experiments, serum starvation was also a means of reducing sample contamination with bovine serum proteins, particularly bovine serum albumin (BSA) which may mask endogenous proteins in 2DE gels. For reproducibility, samples were run in triplicate and reciprocally labelled with each cyanine dye (Figure 4.2). Additionally, the replicate samples were prepared at different times to minimise artefactual differences due to sample preparation or cell growth confluency.

Differences could be detected visually by overlaying the paired fluorescent images produced from single gels. This was carried out using Adobe Photoshop image software by pseudo-colouring, layering and multiplying the exported tiff versions of the gels. Figure 4.3 shows a merged view of the protein spots separated on a 2D gel which was then used as a master throughout this study. In these merged images, proteins of similar expression level appear black and proteins with differential expression appear in red (decrease in C3.6 vs. HB4a) or in blue (increase in C3.6 vs. HB4a). This method of comparison, although rapid, did not provide information regarding the relative expression levels of the proteins visualised but only gross differences such as the presence versus absence of a protein. More importantly, such comparisons can be misleading as the images are not properly normalised. It was therefore important to evaluate further the changes in protein expression using appropriate software enabling the normalisation of gel images and the accurate matching and comparison of protein features taking into consideration their pixel values.

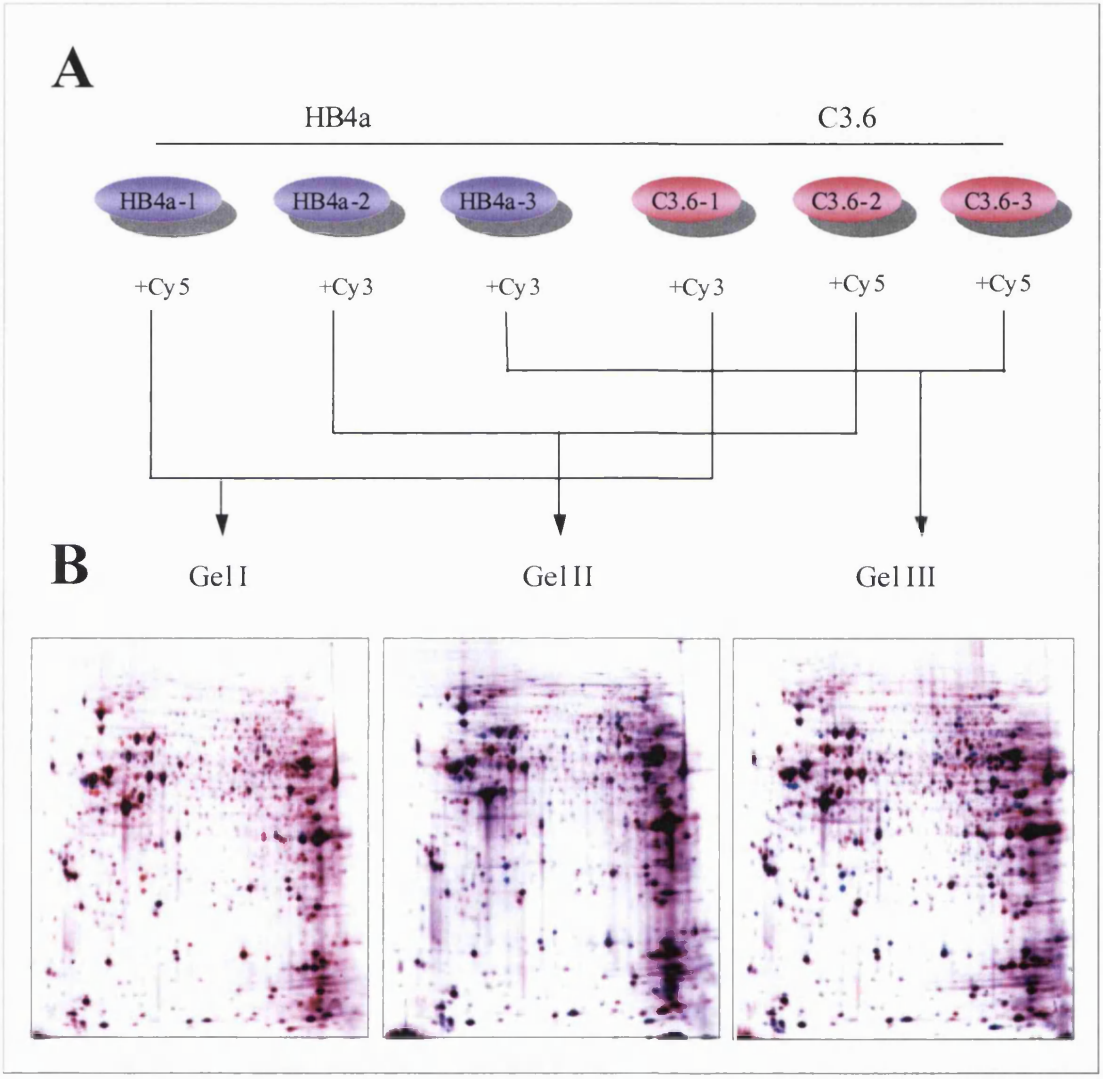


Figure 4.2: 2D-DIGE study of ErbB-2 overexpression in HB4a and C3.6 cell lines. A) A triplicate reciprocal labelling experiment was carried out with Cy3 and Cy5 for increased reproducibility in the comparison of HB4a and C3.6 cells under serum-starved conditions. 100 μg of total cell lysate from HB4a and C3.6 was labelled with Cy3 and Cy5 in different combinations and paired samples were run on the same 2DE gel. B) 2DE gels were imaged using the 2920 2D-Master Imager and paired images were merged using the Adobe Photoshop software. Here, samples labelled with Cy3 are shown in red and samples labelled with Cy5 are shown in blue.

Image analysis was conducted using Melanie III software (Swiss Institute of Bioinformatics, Geneva, Switzerland) and DeCyder software (Amersham Biosciences, UK). These programs use different algorithms for protein spot detection, although in both methods, spot boundaries are defined and the gel images are filtered by background subtraction and normalised. Comparison of gel images is then performed using internal landmarks comprising abundant protein features present in all gel images. Subsequently, statistical analysis can be performed and protein features falling outside a set threshold of protein fold change, and with statistical relevance (p <0.05) can be taken into account.

The best differential analysis was carried out using the DeCyder software, because it had been designed for the analysis of DIGE experiments. Paired images produced from the same 2D gel could be automatically curated and matched using the differential in-gel analysis (DIA) module of DeCyder since the images are directly superimposable. The matched pairs were then compared across gels using the biological variation analysis (BVA) module. Statistical validation was carried out and protein features displaying reproducible changes (n=3, p <0.05) were selected. Reduced reproducibility was evident for the lower-intensity spots, supporting previous observations relating accuracy to protein abundance (Tonge *et al.* 2001). This quantitative image analysis and statistical evaluation of differences between samples revealed 35 protein features that were differentially regulated in C3.6 cells compared to the parental cell line, HB4a and which could be detected in all triplicate gels with average fold ratios +/-1.23 and spot intensities >150,000 pixels to achieve maximum sample recovery for further mass spectrometric identification (Figure 4.3).

As previously discussed (Chapter 3), due to the molecular weight shift caused by dye addition, gels had to be post-stained to define the coordinates of the predominant mass of unlabelled protein and consequently the gels were directly post-stained with SYPRO Ruby fluorescent stain. Most proteins could be matched between the DIGE images and the post-stained SYPRO Ruby images and although all 35 proteins could not be matched with the post-stained image, 27 proteins were assigned with confidence and picked using an Ettan Syprot automated spot picker for subsequent protein identification.

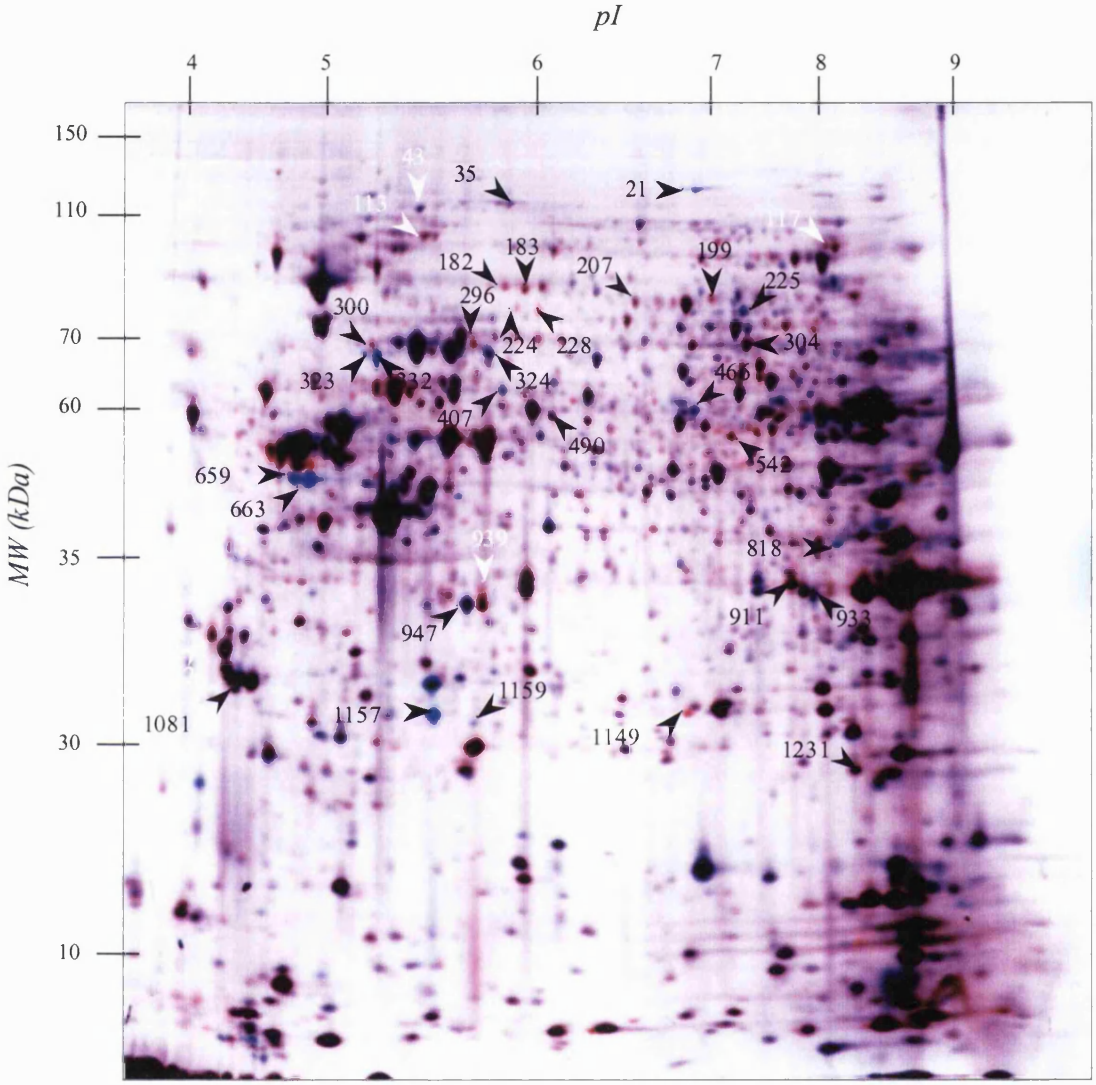


Figure 4.3: Master gel image of protein expression differences between serum-starved HB4a and C3.6 cells. 100 μg of total cell lysate was labelled with HB4a-Cy3 (red) and C3.6-Cy5 (blue) respectively and resolved by 2DE. The image shown was merged using Adobe Photoshop software and differentially expressed proteins are visualised as red or blue and marked with spot identity numbers (black arrows). Unidentified proteins are shown with white arrows. Images from triplicate gels were curated and analysed. The proteins showing significant differential expression are reported in Table 4.1 with their spot identity number and protein identity.

For automated spot picking, the gels had been bonded to one of the glass plates when they were cast to prevent movement of the gel during picking (Chapter 2). A picking list was generated in the DeCyder software from the SYPRO Ruby post-stained image using two fluorescent reference markers stuck to the glass plates to define spot coordinates for the spot picker. Subsequently, differentially expressed proteins were directly analysed by mass spectrometry without running parallel preparative gels, as was performed in other studies (Zhou *et al.* 2002).

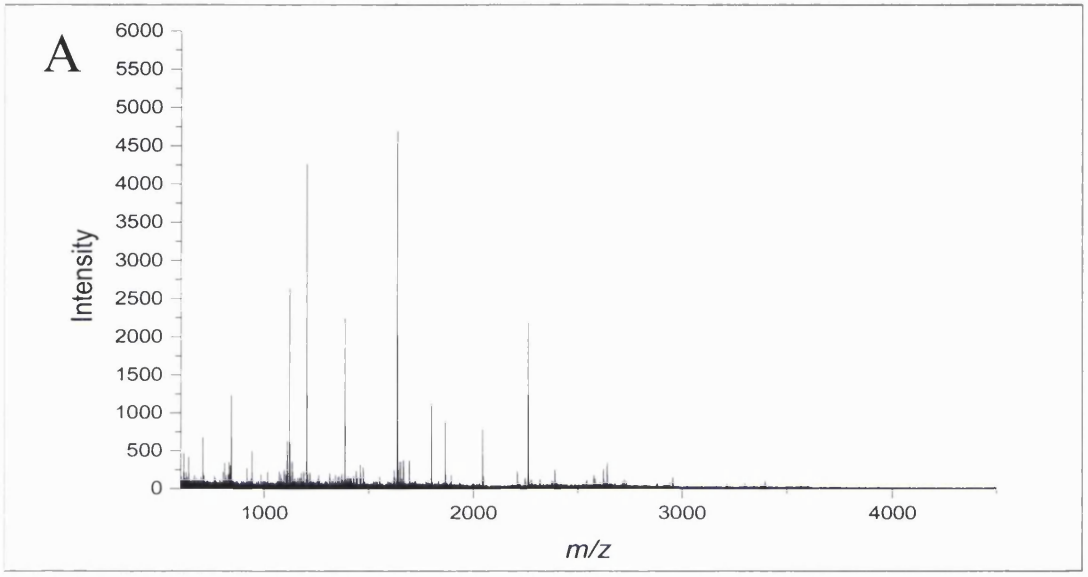
4.4 Protein identification by mass spectrometry

One of the most sensitive methods for protein identification relies on mass spectrometry. Improved mass spectrometers and sample preparation methods now enable the detection of low femtomole amounts of gel-loaded proteins with a high accuracy of identification. Several approaches can be used for protein MS-identification although the most commonly used relies on peptide mass mapping by matrix assisted-laser desorption/ionisation-mass spectrometry (MALDI-MS) (Chapter 1).

In this work, a modified porcine trypsin was used to generate tryptic peptides from gel-separated proteins. This enzyme cleaves at the C-terminal side of lysine and arginine residues without giving rise to excessive autolysis peaks. Protein identification was carried out on samples from triplicate gels and the same protein identifications were found in separate runs, giving increased confidence in the protein identities. Proteins were digested using a modified procedure (Chapter 2) and the tryptic digest mixture was analysed by MALDI-MS using a Bruker Reflex III instrument. The mass spectra obtained were well resolved and showed high signal to noise ratios with high mass accuracy (within +/- 50 ppm) and the picked proteins could be identified with confidence (Table 4.1). Moreover, the peptide mixtures did not require any preliminary cleanup for salt removal or other contaminants. A typical example of a peptide mass spectrum and database search results is shown in Figures 4.4, 4.5 and 4.6.

spot ID	Average Fold difference	UniGene Name	Acc. Number	Mr	pI	% Coverage	Peptides Matched
1157	32.48	ubiquitin carboxyl-terminal esterase L1 (UCH-L1)/ PGP9.5	AAD09172	23026	5.3	51	16
1159	18.60	Heat shock protein 27 (Hsp27)	P04792	22327	7.83	48	11
663	12.69	keratin 17	NP_000413	48105	4.97	63	31
21	7.60	carbamoyl-phosphate synthetase 1 (CPS1)	XP_010819	131192	5.79	38	42
332	5.67	lymphocyte cytosolic protein 1 (L-plastin)	XP_007122	70288	5.29	47	25
225	2.87	radixin	NP_002897	68564	6.03	46	21
818	2.81	3-hydroxyisobutyryl-Coenzyme A hydrolase	NP_055177	42907	8.34	48	23
466	2.00	glutaminase C	AAD47056	65474	8.09	45	23
324	1.96	lymphocyte cytosolic protein (T-isoform)	P13797	70436	5.52	48	27
1081	1.93	14-3-3 beta	NP_003395	28082	4.76	85	32
407	1.89	copine III	NP_003900	60131	5.6	40	22
947	1.75	inorganic pyrophosphatase	NP_066952	32660	5.54	58	18
35	1.32	leucine-rich PPR-motif containing	P42704	145202	5.5	37	42
933	1.26	aldo-keto reductase family 1, member B1 (aldose reductase)	493797	35706	6.55	32	17
300	-1.23	lamin B1	NP_005564	66408	5.11	58	36
490	-1.25	FK506 binding protein 5	AAA86245	50378	5.8	53	26
207	-1.27	villin 2 (ezrin)	T47177	73936	6.46	56	38
304	-1.44	thyroid autoantigen 70kD (Ku antigen)	NP_001460	69843	6.23	27	13
182	-1.49	Ku 80 auto antigen, DNA repair protein (acidic isoform)	NP_066964	82705	5.55	54	43
911	-1.58	PDZ and LIM domain 1 (elfin)	NP_066272	36071	6.56	65	18
296	-1.61	Heat shock 70kD protein 2 (Hsp70)	AAD11466	69995	5.56	49	36
1231	-1.66	superoxide dismutase 2, mitochondrial (MnSOD)	XP_004242	24750	8.34	68	13
183	-1.70	Ku 80 auto antigen, DNA repair protein (basic isoform)	NP_066964	82705	5.55	28	40
199	-1.73	DNA replication licensing factor, MCM7 (CDC47 homolog)	P33993	81281	6.08	51	28
542	-1.77	RuvB-like 1 (E. coli)	NP_003698	50228	6.02	45	17
1149	-2.19	triosephosphate isomerase 1	999892	26538	6.51	72	14
228	-7.26	MxA, interferon-inducible protein p78	P20591	75534	5.6	66	37

Table 4.1: Summary table of 27 proteins differentially regulated in serum-starved HB4a and C3.6 cells identified by peptide mass mapping by MALDI-MS. Average fold-differences were obtained from gel image analysis of triplicate gels using DeCyder. All *p*-values were <0.05. Spot identifications relate to protein coordinates as shown in Figure 3. The percentage coverage represents the amino acid sequence coverage from the identified protein and the peptides matched represent the number of peptides which matched the theoretical masses.



B

MOWSE Score	#/21(%) Masses Matched	% Cov	% TIC	MeanData Err ppm	Tol ppm	MS-Digest Index #	Protein MW (Da)/pI	Accession #	Species	Protein Name
3.960e+005	21 (100)	45.0	100.0	8.81	65.3		60131/5.6		UNREADABLE	gi|4503015|ref|NP_003900.1| copine III [Homo sapiens]
2.422e+004	17 (80)	35.0	81.0	9.70	67.0		60171/5.7		HOMO SAPIENS	similar to copine III
2862	12 (57)	33.0	57.1	6.66	70.2		59585/5.5		UNREADABLE	gi|20822051|ref|XP_131305.1| copine III [Mus musculus]
222	8 (38)	30.0	38.1	-1.87	52.3		43494/6.3		UNREADABLE	gi|17542784|ref|NP_500737.1| Peptide chain release factor (43.5 kD) [Caenorhabditis elegans]
185	4 (19)	28.0	19.0	-4.57	80.6		25905/5.0		UNREADABLE	gi|23134729|ref|ZP_00116503.1| hypothetical protein [Synechococcus sp. WH 8102]

Figure 4.4: Typical mass spectra and database searching for protein identification by peptide mass fingerprinting. Proteins were picked and proteolysed with trypsin (Chapter 2). The peptide mixture generated was then analysed by matrix assisted laser desorption/ionisation-MS (MALDI-MS). A) A typical mass spectrum was acquired on a Bruker Reflex III. B) Peptide mass mapping was performed by database searching using the Protein Prospector (version 3.4.1), MS-Fit search tool (UCSF, USA). Identical search parameters were used for all the proteomic analysis.

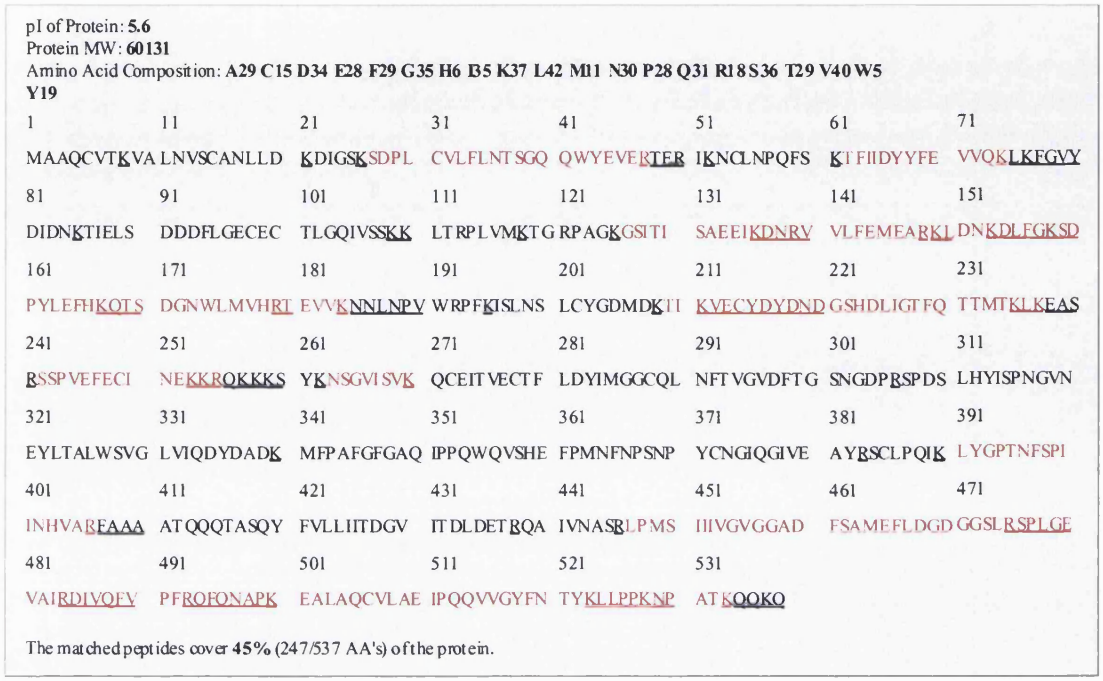
Detailed Results

1. 20/69 matches (28%).
Acc. #: **Species:** UNREADABLE **Name:** gi|4503015|ref|NP_003900.1| copine III [Homo sapiens]
Acc. #: **Species:** HUMAN **Name:** Copine III
Index: **MW:** 60131 Da **pI:** 5.6

m/z Submitted	MH+ Matched	Delta ppm	Modifications	Start	End	Missed Cleavages	Database Sequence
803.4502	803.4627	-16		263	270	0	(K) (Q)
832.4231	832.4317	-10		494	500	0	(R) (E)
941.5584	941.5420	17		476	484	0	(R) (D)
1078.6886	1078.6625	24		524	533	1	(K) (Q)
1093.6181	1093.5716	43		140	148	0	(R) (K)
1109.6175	1109.5665	46	1Met-ox	140	148	0	(R)VVLFEMEAR(K)
1120.6698	1120.6155	48		485	493	0	(R) (Q)
1177.7027	1177.6581	38		149	158	2	(R) (S)
1664.9328	1664.8576	45		62	74	0	(K) (L)
1691.9831	1691.9155	40		140	153	2	(R) (D)
1722.9319	1722.8485	48		242	255	2	(R) (Q)
1799.0142	1798.9604	30		391	406	0	(K) (F)
2016.0323	2015.9973	17	1Met-ox	168	184	1	(K)QTSDGNWLMVHRTEVVK(N)
2043.1727	2043.1391	16		476	493	1	(R) (Q)
2541.1402	2541.2084	-27		27	47	0	(K) (T)
2576.1694	2576.1992	-12	1Met-ox	159	179	1	(K)SDPYLEFHKQTSDGNWLMVHR(T)
2623.2631	2623.3401	-29	1Met-ox	126	148	2	(K)GSITISAEEIKDNRVVLFEMEAR(K)
2641.2744	2641.3336	-22		501	523	0	(K) (L)
2956.3144	2956.4436	-44	2Met-ox	447	475	0	(R)LPMSIIIVGVGGADFSAMEFLDGDGGSLR(S)
2956.3144	2956.4436	-44	1MetO+32	447	475	0	(R)LPMSIIIVGVGGADFSAMEFLDGDGGSLR(S)
3212.3921	3212.3375	17	1Met-ox 1MetO+32 3PO4	447	475	0	(R)LPMSIIIVGVGGADFSAMEFLDGDGGSLR(S)
3212.3921	3212.4876	-30	1Met-ox 1PO4	159	184	2	(K)SDPYLEFHKQTSDGNWLMVHRTEVVK(N)
3393.4653	3393.5983	-39		209	237	2	(K) (E)

The matched peptides cover **45%** (247/537AA's) of the protein.
Coverage Map for This Hit (MS-Digest index #):

Figure 4.5: Protein identification by database searching. The list of peptide masses was matched against the NCBI theoretical peptide fragment database using the MS-Fit program from Protein Prospector (version 3.4.1) (UCSF, USA). The search criteria used allowed +/- 50 ppm mass error and up to 2 missed cleavages. A) The reported hit summarises the list of matched submitted masses (here 20 out of 69 peptides matched the theoretical database).


pI of Protein: **5.6**
Protein MW: **60131**
Amino Acid Composition: **A29 C15 D34 E28 F29 G35 H6 I35 K37 L42 M11 N30 P28 Q31 R18 S36 T29 V40 W5 Y19**

1	11	21	31	41	51	61	71
MAAQCVTKVA	LNVSCANLLD	KDIGSKSDPL	CVLFLNTSGQ	QWYEVERTER	IKNCLNPQFS	KTFIIDYYFE	VVQKLKFGVY
81	91	101	111	121	131	141	151
DIDNKTIELS	DDDFLGECEC	TLGQIVSSKK	LTRPLVMKTG	RPAGKGSITI	SAEEIKDNRV	VLFEMEARKL	DNKDLFGKSD
161	171	181	191	201	211	221	231
PYLEFHKQTS	DGNWLMVHRT	EVVKNNLNPV	WRPFKISLNS	LCYGDMDKTI	KVECYDYDND	GSHDLIGTFQ	TTMTKLKEAS
241	251	261	271	281	291	301	311
RSSPVEFECI	NEKKRQKKKS	YKNSGVISVK	QCEITVECTF	LDYIMGGCQL	NFTVGVDFTG	SNGDPRSPDS	LHYISPNGVN
321	331	341	351	361	371	381	391
EYLTALWSVG	LVIQDYDADK	MFPAFGFGAQ	IPPQWQVSHE	FPMNFNPSNP	YCNGIQGIVE	AYRSCLPQIK	LYGPTNFSPI
401	411	421	431	441	451	461	471
INHVARFAAA	ATQQQTASQY	FVLLIITDGV	ITDLDETRQA	IVNASRLPMS	IIIVGVGGAD	FSAMEFLDGD	GGSLRSPLGE
481	491	501	511	521	531		
VAIRDIVQEV	PFRQFQNAPK	EALAQCVLAE	IPQQVVGYFN	TYKLLPPKNP	ATKQQKQ		

The matched peptides cover **45%** (247/537 AA's) of the protein.

Figure 4.6: Peptide mass mapping and protein coverage. Representation of the protein coverage, where in the case analysed, the matched peptides represent 45% of the sequence of the protein.

Since high mass accuracy and good protein coverage was reached, allowing confident identification of all samples, additional sequence analysis by electrospray ionisation mass spectrometry (ESI-MS) or high-performance liquid chromatography (HPLC)-ESI-MS was not necessary. The 27 identified spots were all cut from gels loaded with only analytical amounts (200 µg) of cyanine dye labelled samples that were post-stained with SYPRO Ruby. This approach was therefore labour efficient as no running of preparative gels was required. Moreover, these results show clearly that the combination of two detection approaches, cyanine dye labelling and SYPRO Ruby post-staining, does not affect the ability to identify proteins by MALDI-MS.

Although approximately 20% of the total amount of each protein should be labelled with the cyanine dyes, no modified peptides were detected by mass spectrometry. Moreover, the extent of missed cleavages observed, which may have

reflected the presence of a dye molecule causing steric hindrance to the proteolysis, was no higher than that usually observed in samples where other methods of detection were used (i.e. the same ratio of cleavages occurring after lysine and arginine residues was observed). This would suggest that the representation of modified peptides over their unmodified form was too low to be detected by mass spectrometry. Although the dye modification of peptides could affect their potential to ionise and be detected during MALDI-MS, a previous study has reported the detection of cyanine-dye labelled lysozyme in-solution, using ESI-MS (Unlu *et al.* 1997). This therefore suggests that peptide modification with the cyanine dyes can be detected by mass spectrometry when large amounts of material are available. The hydrophobicity of the dyes may also account for the apparent absence of detection of labelled peptides, since these peptides may be retained in gel pieces during the protein digestion process or may not crystallise properly on the target. In conclusion, dye labelling did not appear to affect the protein processing for successful mass spectrometric identification and in addition the combination of the two detection methods, DIGE and SYPRO Ruby did not impair the analysis.

4.5 Summary of differentially expressed proteins

All 27 picked proteins could be identified with confidence by MALDI-MS-peptide mass mapping (Table 4.1). These identifications represented mostly different protein species, although several isoforms could be detected, for example in the case of the Ku proteins. Amongst the proteins identified, 14 were up-regulated and 13 were down-regulated to a significant degree in the ErbB-2 overexpressing cell line under serum-starved conditions. The identified proteins are reported on the master gel and their theoretical pIs and molecular weights were verified by comparison with this gel (Figures 4.3 and 4.7). Some of the proteins identified migrated as multiple spots suggesting that post-translational modifications had taken place (MxA, L-Plastin, Ku 80 subunit, cytokeratin 17), although no specific modifications could be detected by MALDI-MS.

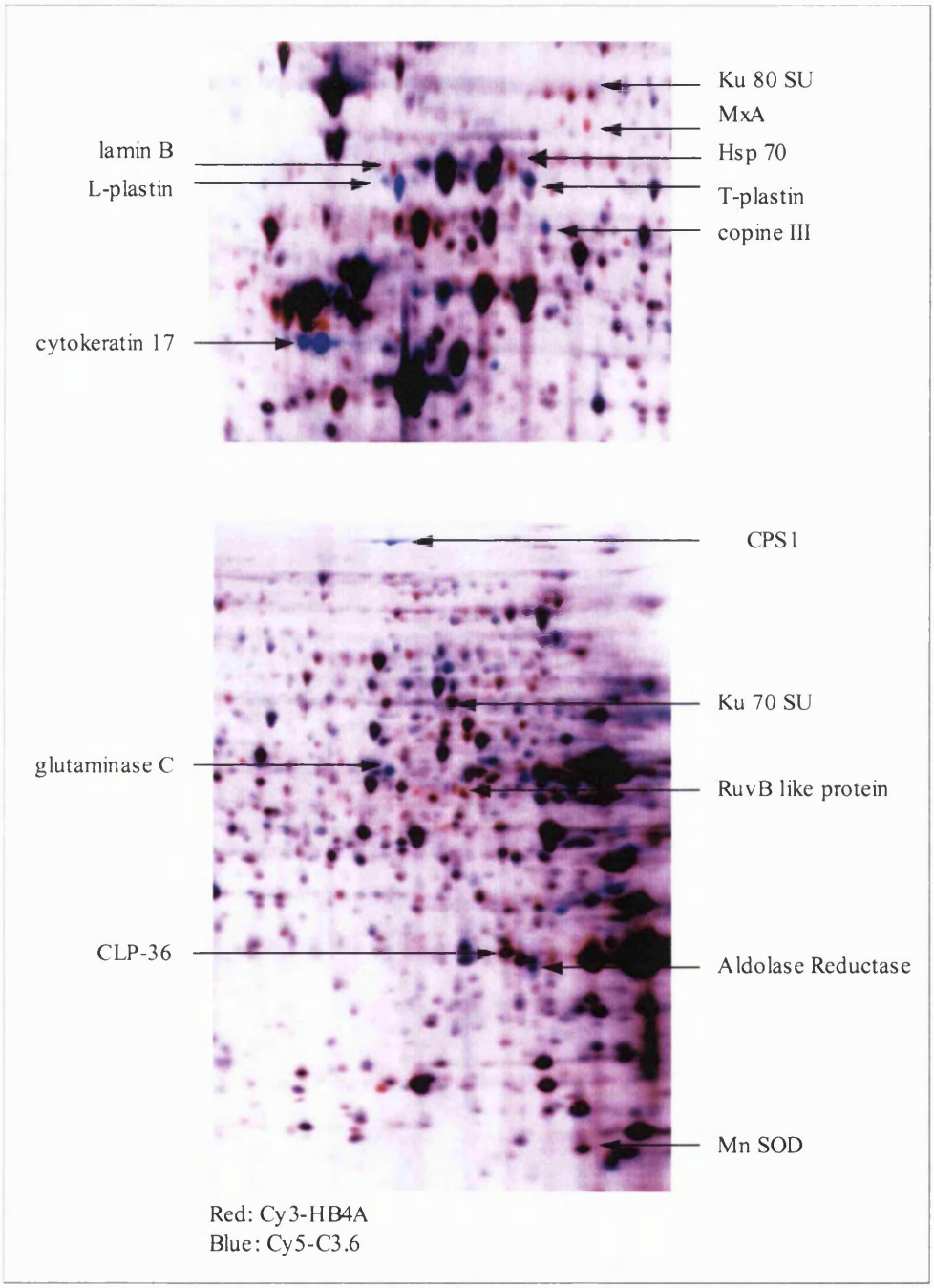


Figure 4.7: 2D-gel representation of 15 differentially expressed proteins identified by 2D-DIGE and MALDI-MS. HB4a and C3.6 cell lines were separately labelled with Cy3 (red) and Cy5 (Blue) respectively and resolved on the same 2DE. Images were then merged using Adobe Photoshop. Proteins detected as up-regulated in the C3.6 compared to HB4a are shown in blue and the proteins down-regulated are shown in red.

Further literature searches were carried out in order to assess the roles of these proteins in ErbB-2-mediated cellular transformation or in cancer. This revealed that some of the proteins identified have been reported to play roles in several different cellular processes including metabolic control, cytoskeletal organisation, membrane trafficking, and protein synthesis. Importantly, it was interesting to find that several of these differentially regulated proteins have already been shown to be deregulated in various forms of cancer. Seven of these proteins (PGP9.5/UCH-L1, Hsp27, cytokeratin17, CPS1, L-plastin, radixin and glutaminase C) showed a 2-fold or greater increase in expression in the ErbB-2 overexpressing cell line (Table 4.1). The protein, which showed the greatest deregulation, and was highly expressed in C3.6, was the deubiquitin enzyme PGP9.5. This protein, also known as ubiquitin C-terminal hydrolase L1 (UCH-L1) was originally reported to be expressed in neuronal tissues. It has however been found overexpressed in several human cancers including lung and colon adenocarcinomas (Hibi *et al.* 1999; Brichory *et al.* 2001). It has been suggested that the uncontrolled deubiquitination of proteins could lead to tumour progression, since normal targets of ubiquitin-dependent proteasomal degradation may not be degraded when PGP9.5 is overexpressed. As yet, there are no reports of PGP9.5 overexpression in breast cancers and the level of expression of PGP9.5 has not been related to other oncogenic markers. Therefore this protein was selected as an important potential target of ErbB-2 overexpression and further work was carried out to validate and characterise it and to assess its role, if any, in ErbB-2-dependent tumour progression (Chapter 6). Similarly, Hsp27 also showed significant increase in abundance in the C3.6 cells compared to the HB4a cells. This chaperone protein has been implicated in various cellular processes including stress responses, apoptosis and actin reorganisation. More importantly, it has been shown to be overexpressed in different human cancers (Sarto *et al.* 2000).

Several of the other identified proteins, which have been previously implicated in cellular transformation and metastasis, include the intermediate filament protein L-plastin (Lin *et al.* 1988; Lapillonne *et al.* 2000), the glutamine-metabolising enzyme glutaminase C (Lobo *et al.* 2000), and the cell signalling modulator protein 14-3-3β (Takihara *et al.* 2000). L-plastin is a member of the plastin family of actin-bundling proteins. Although it

was thought to be specific to heamopoietic cell lineages, it has been found overexpressed in malignant human solid tumours (Lin *et al.* 1993; Park *et al.* 1994) and L-plastin deregulation has also been reported in an expression clustering analysis carried out on a panel of breast cancer tissues and cell lines (O'Hare *et al.*, unpublished data). L-plastin appears to contribute to the regulation of cell adhesion and motility (Jones *et al.* 1998) and the deregulated expression of plastin proteins has been correlated with cancer invasion and metastasis (Otsuka *et al.* 2001). This latter study suggested the potential use of L-plastin as a clinical marker of metastatic cancer (Otsuka *et al.* 2001). Importantly, the functional involvement of L-plastin in prostate cancer proliferation, invasion and metastasis was shown in an antisense study in PC3 prostate carcinoma cells (Zheng *et al.* 1999). Cells transfected with an antisense L-plastin gene, were shown to display suppression of *in vivo* invasion in nude mice. Despite this, the mechanism of action of L-plastin remains unclear and further studies are required to relate the role of its actin-binding properties with the up-regulation of cell proliferation and motility.

The increased expression of the metabolic enzymes carbamoyl phosphate synthetase 1 (CPS1), glutaminase C, 3-hydroxyisobutyryl-Coenzyme A hydrolase, inorganic pyrophosphatase and aldose reductase in the C3.6 cells could be a direct consequence of the enhanced proliferation of these cells in response to ErbB-2 overexpression and the increased energy requirements associated with growth. Similarly, changes in the expression of the identified structural proteins (cytokeratin 17, radixin, ezrin, L-plastin, T-plastin, leucine-rich protein and lamin B) could account, at least in part, for the different morphologies displayed by the two cell types, as was shown previously (Figure 4.1).

Another protein of interest, copine III, was up-regulated in the C3.6 cells compared to the HB4a. This protein is a member of the copine family and represents a new class of C2 domain-containing proteins (Creutz *et al.* 1998), whose exact functions are unknown. The copines are putative calcium-dependent phospholipid-binding proteins and it has been proposed that they may play a role in membrane trafficking (Creutz *et al.* 1998). The overexpression of copine III had not been reported in human cancers or related to ErbB-2 overexpression or growth factor signalling, therefore this protein

represents an interesting target for further characterisation (Chapter 5).

One other protein also up-regulated in the C3.6 cells was 14-3-3β. The 14-3-3 proteins are involved in a vast array of processes such as the response to stress, cell-cycle control, and apoptosis, and they serve as adapters and repressors of cellular processes by sequestering phosphorylated substrates. There are currently 133 full-length sequences available in GenBank for this highly conserved protein family (Mar 2003). This family of proteins is also found to be deregulated in cancers. For example, the expression of 14-3-3τ is frequently lost in breast cancers (Ferguson *et al.* 2000) due to hypermethylation-dependent gene inactivation (Osada *et al.* 2002). However, the up-regulation of 14-3-3β in breast caner has not been reported in conjunction with ErbB-2 overexpression and should therefore be investigated further.

Amongst the significantly down-regulated proteins in the ErbB-2 overexpressing cells, two are interferon regulated proteins, MxA and MnSOD. MxA is an interferon-induced GTPase, and a member of the dynamin family of large GTPases with a role in viral protection (Haller *et al.* 1998). The manganese mitochodrial superoxide dismutase (MnSOD) is an antioxidant enzyme which catalyses the conversion of superoxide radicals ($O_2^{\cdot-}$) to hydrogen peroxide (Weisiger and Fridovich 1973). MnSOD has been shown to play a role in the protection of cells from oxidative stress and also to prevent tumourigenicity (Oberley and Buettner 1979) and this mechanism of protection can be induced by TNFα and also via the activation of interferon signalling (Wong *et al.* 1996). This protein is an interesting marker as it is thought to play a role as a tumour suppressor in breast cancer (Li *et al.* 1995). However, the mechanism involved in the regulation of MnSOD and its role in reduced tumourigenicity are not well understood. Thus two major interferon-regulated genes were found to be down-regulated in the ErbB-2 overexpressing cell line, suggesting suppression of IFN signalling by ErbB-2. As previously reported, interferon treatment has a negative effect on cellular proliferation and has been used for cancer therapy (Maemura *et al.* 1999; Tossing 2001). Thus the suppression of this pathway could be a means by which ErbB-2 promotes cellular proliferation.

Another member of the heat shock protein family, Hsp70, is also found

deregulated in cancers (Sarto *et al.* 2000) and was down-regulated in the ErbB-2 overexpressing cell line. Its exact role in cancer is not well understood, although it could be involved in the modulation of stress responses or anti-tumour immunity (Ito *et al.* 2003). A recent report has implicated the regulatory role of CHIP, a new ubiquitin ligase, in the degradation of ErbB-2. This ligase appears to act by controlling the association of Hsp70/Hsp90 chaperones with ErbB-2. This could correlate with the altered levels of Hsp70 observed by DIGE analysis and could point to a negative regulatory mechanism to keep levels of ErbB-2 overexpressed (Zhou *et al.* 2003).

Interestingly, another protein involved in mechanisms of regulation of cellular damage is the Ku 70/80 which was found to be down-regulated in the C3.6 cells. Ku 70/80 is a helicase involved in the early steps of double strand break recognition. Active heterodimers of the 70 and 80 subunit interact with DNA-PK, a serine-threonine protein kinase which subsequently activates nuclear targets responsible for the processes of DNA repair (Gottlieb and Jackson 1993; Poltoratsky *et al.* 1995). Another DNA helicase, the mini-chromosome-maintenance protein 7 (Mcm7), was also found to be down-regulated in the C3.6 cells. The Mcm proteins form a complex involved in the initiation of DNA replication (You *et al.* 2002). They have been implicated in the replication licensing system, which prevents the re-replication of DNA during the cell cycle (Blow and Laskey 1988; Blow and Hodgson 2002), a regulatory system that is under the control of cyclin-dependent kinases (CDKs) (Dahmann *et al.* 1995; Hua *et al.* 1997). This data therefore suggests that ErbB-2 may impair DNA repair and synthesis through the down-regulation of both Ku 70/80 and Mcm7.

Taken together, these findings showed the use of 2D-DIGE as a valuable tool in the analysis of this model cell system as it enabled the identification of candidate biomarkers of ErbB-2 overexpression. Moreover, this subset of proteins identified had been previously described as potential markers of cancers, validating the HB4a and C3.6 cells as a useful model for the study of ErbB-2-dependent breast cancer. These proteins were shown to be involved in cellular control and proliferation, metabolism, cell-cell interaction and signalling. Thus, the altered regulation of these proteins is likely to be involved in ErbB-2-mediated hyperproliferation and cellular transformation.

4.6 Effect of growth factor treatment by 2D-DIGE analysis

The effect of growth factor treatment on the protein expression profile of both cell lines was further investigated. In a previous study carried out in our laboratory, Timms *et al.* examined the effect of ErbB-2 overexpression on mitogenic signalling in the HB4a and C3.6 cell system (Timms *et al.* 2002). The C3.6 cell line was found to be more responsive to growth stimuli than the parental cell line, but also to show a degree of growth-factor independence. As reported in Chapter 1, although ErbB-2 does not have any known ligand, it is the preferred heterodimerisation partner of the other members of the family and can enhance signalling in response to their specific growth factor binding (Graus-Porta *et al.* 1997; Olayioye *et al.* 1998; Muthuswamy *et al.* 1999; Yarden 2001). HB4a and C3.6 cells were treated with either EGF or Hrgβ1, specific ligands of EGFR and ErbB-3 respectively. Each ligand was shown to have a different potency in inducing downstream signalling. Notably, EGF produced a slight increase in PI3K signalling in HB4a compared to C3.6, whereas Hrgβ1-stimulated PI3K signalling was higher in C3.6, and both ligands triggered an increase in MAPK signalling in C3.6 compared to HB4a (Timms *et al.* 2002). Both ligands were used at physiologically relevant concentrations (1 nM each), which was shown to be sufficient to trigger maximal downstream signalling and cellular proliferation (Timms *et al.* 2002). In the previous 2D-DIGE experiment, cells were compared under serum-starved conditions, which provided information on the basal differential protein expression profile between the cell lines due to ErbB-2 overexpression. In this next experiment, the effect of growth factor treatment on the protein expression profile was investigated in the context of ErbB-2 overexpression.

A time course of EGF and Hrgβ1 stimulation was performed in both cell lines and total cellular extracts were analysed by 2D-DIGE. Four time points were chosen (0h, 4h, 8h and 24h) for treatment with 1 nM Hrgβ1 and triplicate samples from separate experiments were run (Figure 4.8).

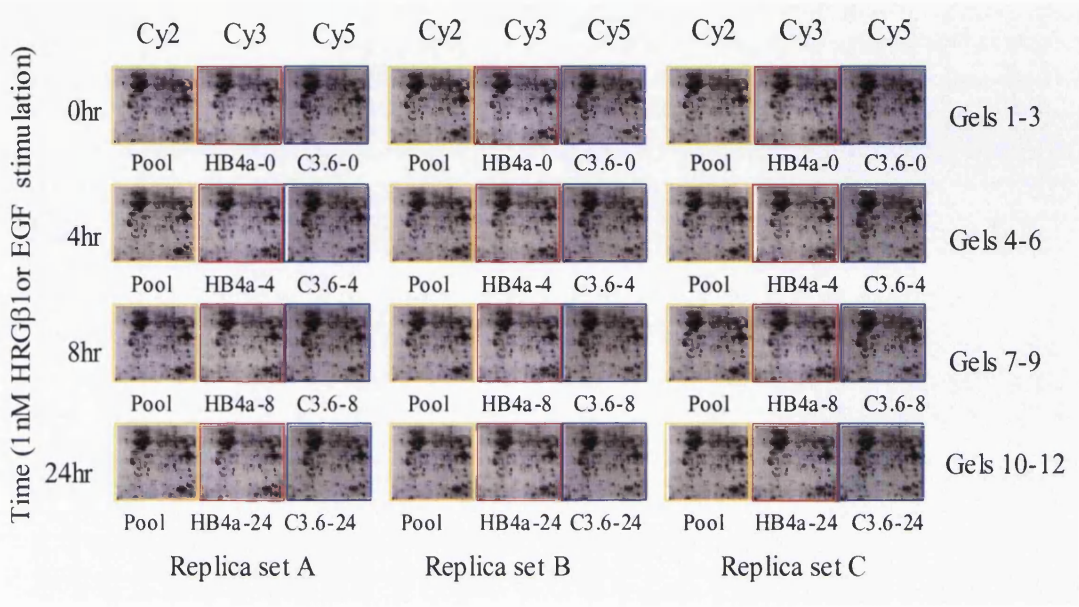


Figure 4.8: Representation of the experimental design in the study of the protein expression profile in response to growth factor treatment. Serum-starved HB4a and C3.6 cells were treated with 1 nM Hrgβ1 (or 1nM EGF) for 4 hrs, 8 hrs and 24 hrs and compared to the untreated samples. HB4a and C3.6 were labelled with Cy3 and Cy5 respectively and compared to an internal standard representing pooled aliquots of each sample labelled with Cy2. Samples were run in triplicate for statistical analysis and to assess reproducibility. This represents 12 gels with 36 images generated. Gel images were processed using the BVA module of DeCyder software (Amersham Biosciences).

The cell lines were compared on the same gels at each time point and HB4a was labelled with Cy3 and C3.6 with Cy5. The third cyanine dye, Cy2, was used to label a pool of samples prepared by mixing equal amounts of protein from all 24 samples. The same amount of this pool (100 μg) was run on all gels, providing an internal standard for normalisation and accurate statistical analysis. Thus, for each growth factor, 12 gels were run in which the HB4a proteome was compared to that of C3.6 at each time of treatment and was also compared to the internal standard, generating 36 images for further

comparison and image analysis. Direct visual comparison of merged gel images enabled the detection of many protein features with altered abundance between the cell lines, but also changes over the time course of treatment with Hrgβ1 (Figure 4.9).

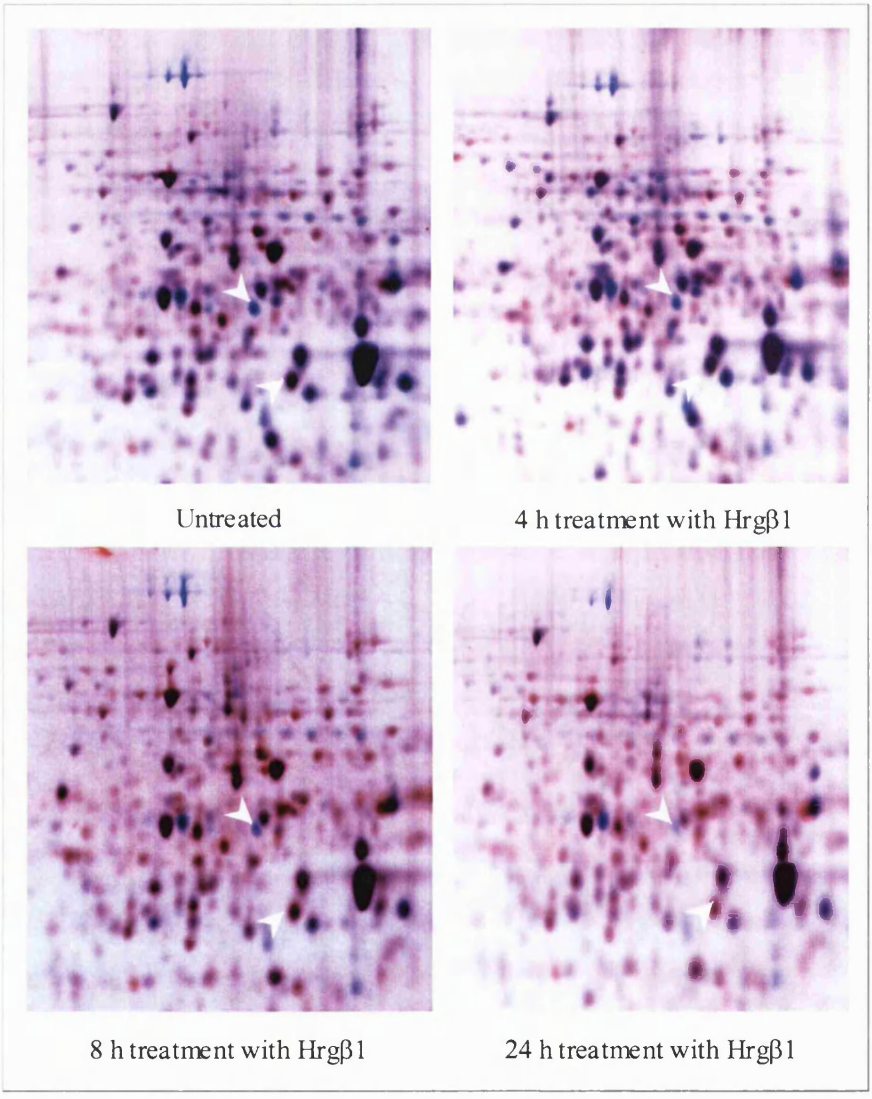


Figure 4.9: Visualisation of differential expression in HB4a vs. C3.6 cells at different time points of stimulation with Hrgβ1. Representative merged images of HB4a and C3.6 labelled with Cy3 (red) and Cy5 (blue) respectively are shown. Some protein features have altered expression across the time course (white arrows).

DeCyder software was then used for semi-automated gel analysis and statistical filtering. A "T test" was performed for which all samples were compared to their internal standard and matched across gels. A description of the software is given in detail in Chapter 2. Briefly, 16-bit tiff files of gel image were exported from the 2D-Master Imager software and entered into the DeCyder batch analysis software as appropriate Cy3/Cy2 and Cy5/Cy2 pairs. At this stage, images are defined into groups based on cell line, time point and replicate experiments. The average number of spots was then estimated to 2500 and automatic spot detection, normalisation and quantitation was performed on the paired images comparing the Cy2-labelled standard to each sample labelled with either Cy3 or Cy5 for each gel. In BVA, spots were matched across gels and improved spot matching was achieved by assigning landmarks on each standard image. The visual display given by the BVA module enabled the comparison of changes on the gel images simultaneously with a three-dimensional (3D) representation of spot intensity and a graphic display of the changes in expression over the time course of growth factor treatment (Figure 4.10). Statistical analysis was then performed and features displaying significant differences in abundance were further checked prior to selection of spots for picking.

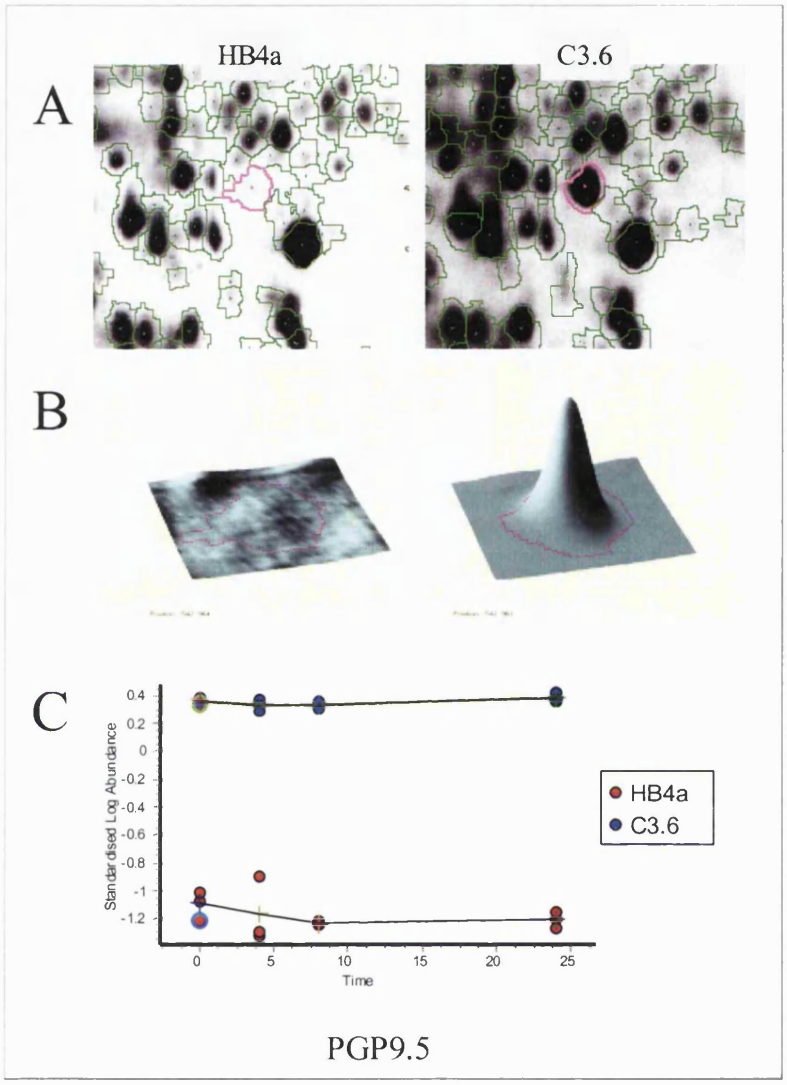


Figure 4.10: DeCyder output showing statistical changes in PGP9.5 expression level between the HB4a and C3.6 cell lines. The expression profiling of PGP9.5 over a time course of treatment with Hrgβ1 is represented. A) Differences in expression level between the cell lines could be observed in the 2DE gel image. B) A three-dimensional representation of the protein abundance was generated enabling a better spot matching and estimation of the changes in abundance between the cell lines. C) A graphic representation shows all Cy3/Cy2 and Cy5/Cy2 ratios for replicate samples and enables the estimation of variations over the time course and between the cell lines. HB4a and C3.6 are shown in red and blue respectively. The values are standardised log of abundance over the standard labelled with Cy2. Each replicate is represented as a single spot at each time point and the lines were plotted using the averaged values.

As a result of this analysis, many significant changes could be observed between the cell lines at each time point of treatment with Hrgβ1 (Table 4.2). Of these, there were more spots up-regulated in C3.6 versus HB4a than were down-regulated.

Time (hrs)	Average Fold-difference	Total no.spots	Decreased in C3.6	Increased in C3.6
T0	>1.3	135	34	101
	>1.5	81	14	67
	>2.0	23	3	20
	>5.0	2	0	2
T4	>1.3	160	53	107
	>1.5	74	22	52
	>2.0	16	2	14
	>5.0	2	0	2
T8	>1.3	159	56	103
	>1.5	68	24	44
	>2.0	20	3	17
	>5.0	3	0	3
T24	>1.3	165	44	121
	>1.5	75	18	57
	>2.0	16	3	13
	>5.0	1	0	1

Table 4.2: Number of protein features showing significant changes in protein abundance between the cell lines at each time point of treatment with Hrgβ1. Here, 2607 matched spots, with spot intensities >100,000 pixels, were analysed. Numbers of significantly different spots at increasing average fold-difference ratios (>1.3 to >5.0) are compared between the cell lines at each timepoint. All average fold differences have *p* values <0.05.

Changes in the number of spots at each time point was indicative of a growth factor-mediated change in the proteome, although relatively few significant differences were detected between time points of Hrgβ1 stimulation (Table 4.3). Despite a decrease

in overall protein expression in both cell lines upon Hrgβ1 treatment, there was an increase in C3.6 after 24 hrs. The subsequent identification of proteins involved in all these changes has not yet been carried out and is awaiting further investigation.

Time (hrs)	Fold-difference	HB4a (down/up)	HBc3.6 (down/up)
T0/T4	>1.3	5/0	6/1
	>1.5	2/0	0/0
	>2.0	1/0	0/0
T0/T8	>1.3	0/1	2/3
	>1.5	0/0	1/0
	>2.0	0/0	0/0
T0/T24	>1.3	3/6	2/11
	>1.5	3/4	2/4
	>2.0	0/0	0/1

Table 4.3: Number of protein features showing significant changes in abundance over time of Hrgβ1 treatment. The number of protein features with average fold differences >1.3, 1.5 and 2 are shown which change significantly between unstimulated cells (T0) and cells treated with Hrgβ1 for 4, 8 and 24 hrs. A total of 2607 matched protein features were used and only features with p values of <0.05 were included in this table.

This analysis enabled the statistical evaluation of the differentially expressed proteins identified in the previous experiment comparing unstimulated cells. The effect of ligand treatment on the expression profile of these proteins was investigated. Many of these proteins identified had similar expression levels at each time point of treatment with Hrgβ1, indicating that Hrgβ1 treatment does not further affect the already altered expression profile of these proteins. However, some proteins displayed altered expression levels across the time course. MxA and MnSOD had reduced protein expression in the HB4a cells after 4 hrs and 8 hrs of treatment with Hrgβ1 (Figure 4.11). Additionally,

copine III, which was constitutively up-regulated in the C3.6 cells, had increased expression upon ligand treatment in the HB4a cells (Figure 4.11). In addition, CPS1 and L-plastin levels were also slightly increased in HB4a in response to Hrgβ1 (Figure 4.11). These data suggest that these proteins are activated downstream of ErbB signalling in the HB4a cells, as their levels begin to approach those observed in response to ErbB-2 overexpression.

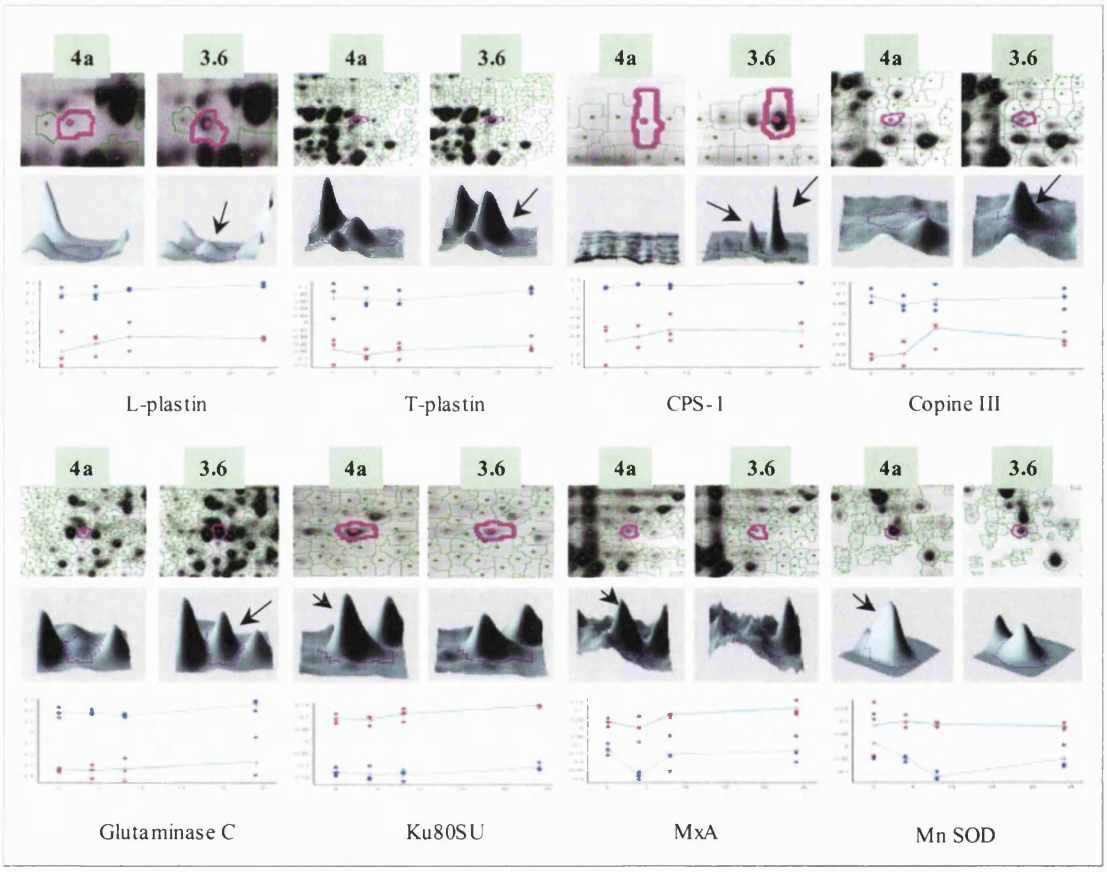


Figure 4.11: DeCyder analysis of statistical changes in protein abundance observed between HB4a and C3.6 for several of the identified proteins. Differences between two samples could be observed simultaneously on the gel image and as a three-dimensional representation of fluorescence intensity. The statistical changes in protein abundance could be evaluated across the period of Hrgβ1 stimulation.

In a similar experiment, cells were treated with 1 nM EGF for 1 hr, 4 hrs, and 24 hrs and compared to the untreated control. The same experimental design was employed and gel images were curated using the DeCyder software. An internal standard pool of all samples labelled with Cy2 was also used for accurate normalisation and quantitation of differences across gels. Under treatment with EGF, more differences could be detected between the cell lines, similar to the changes in protein abundance observed with the ErbB-3 specific ligand, Hrgβ1 (Table 4.4).

Time (hrs)	Average fold ratio	Total No. Spots	Decreased in C3.6	Increased in C3.6
T0	>1.3	191	68	123
	>1.5	113	36	77
	>2.0	30	6	24
T1	>1.3	356	135	221
	>1.5	201	66	135
	>2.0	56	22	34
T4	>1.3	275	87	188
	>1.5	156	46	110
	>2.0	41	12	29
T24	>1.3	239	69	170
	>1.5	139	36	103
	>2.0	44	11	33

Table 4.4: Number of protein features changing in abundance between the cell lines at each time point of stimulation with EGF. DeCyder analysis revealed 2481 spots matched between all 12 gels and 36 images with spot intensities >100,000 pixels. Similarly to the Table 4.2, similarly different spots at increasing fold ratios (>1.3 to >2.0) and with p values <0.05, are compared between the parental and the ErbB-2 overexpressing cell lines.

However, more differences between cell lines were found upon EGF treatment compared to Hrgβ1, perhaps reflecting differences in receptor signalling specificity. In this experiment, an earlier response was investigated, at 1 hr of treatment with EGF. Taken together, this data showed that EGF has an increased potency to induce downstream

signalling (Tables 4.4 and 4.5).

Time (hrs)	Average fold ratio	HB4a (down/up)	C3.6 (down/up)
T0/T1	>1.3	11/16	21/42
	>1.5	4/9	11/17
	>2.0	0/1	2/6
T0/T4	>1.3	9/18	14/18
	>1.5	2/7	7/10
	>2.0	0/1	2/2
T0/T24	>1.3	21/43	47/46
	>1.5	11/15	20/11
	>2.0	0/1	2/2

Table 4.5: Number of protein features with statistically significant changes in abundance over the time points of stimulation with EGF. Similarly to Table 4.3, the protein changes between serum-starved cells (T0) and cells stimulated with 1nM EGF for 1 hr (T1), 4 hrs (T2) and 24 hrs (T3), are compared in each cell line. These were calculated from a total of 2481 matched protein features with spot intensities >100,000 pixels and p values <0.05.

Evaluation of this data therefore suggests that long-term ErbB-2 overexpression has a dramatic effect on the protein expression profile of luminal epithelial cells, as many changes in protein abundance were observed between the cell lines. These changes in protein abundance could account for ligand-independent changes or constitutive ErbB-receptor activation. Growth factor treatment induced less changes in the protein expression profile but these seemed to be specific for the growth factor used (Hrgβ1 or EGF). In addition, this could also be attributed to the level of detection of proteins, as ligand-induced protein changes could occur below the threshold of detection of this technique. Taken together, this analysis has provided important information regarding the growth factor regulation of the identified proteins.

4.7 Comparison of protein expression and mRNA profiles

In parallel with the proteomic study described here, a gene transcription profiling of the HB4a and C3.6 cells was performed in our laboratory (White *et al.*, submitted manuscript). This experiment was carried out under identical growth conditions to those used for 2D-DIGE protein expression profiling. Total mRNA was isolated and reverse transcribed from cells treated with Hrgβ1 for 0, 1, 4, 8, 12, 18 and 24 hrs. These were then labelled with Cy3 and Cy5 and hybridised to microarray slides with probes consisting of a redundant group of 9932 PCR-derived, sequence-verified cDNA probes (representing approximately 6000 unique genes). Both cell lines were directly compared at each time point using a reciprocal duplicate labelling strategy. It is important to note that microarray analysis is not a quantitative approach and no internal standards were used in this experiment. Thus the analysis only provides relative ratios of mRNA expression associated with ErbB-2 overexpression at each time point, and not a measure of the actual up- or down-regulation between time points.

An extensive list of differentially regulated genes was generated for each time point (White *et al.* submitted manuscript). Overall, 667 probes displayed a greater than two-fold differential expression between C3.6 and HB4a, at one or more time points and substantially more genes were differentially expressed at 0 hr, 18 hrs and 24 hrs. Notably, four functional groups of genes displayed differential mRNA expression between the cell lines and/or over the time course of treatment with Hrgβ1. These genes were involved in cellular adhesion, AP-1 mediated transcription, bone remodelling and interferon signalling. The identity of these genes supports previous data showing that ErbB-2 overexpression alters cellular adhesion and mitogenic signalling via the MAPK pathway. In addition, these data suggest that ErbB-2 overexpression may promote bone metastasis, which occurs at higher incidence in breast cancer patients. Interestingly, numerous genes involved in interferon signalling were found to be down-regulated in the ErbB-2 overexpressing cell line. Previous reports have shown that interferon signalling has an inhibitory effect on cell growth (Harvat and Jetten 1996; Iacopino *et al.* 1996), although there are no reports linking ErbB-2 overexpression to the down-regulation of interferon-regulated genes.

Protein expression ratios from 2D-DIGE analysis (27 proteins; Table 4.1) and from previous western blotting analyses (16 proteins; Timms *et al.* 2002 and unpublished data) obtained by densitometry of films were compared to mRNA expression ratios from microarray analysis, representing a total of 43 genes which were present on the microarray chips. A statistically significant relationship between protein ratios and mRNA ratios was observed with a correlation coefficient of 0.81 (p <0.001) and a linear regression of 0.66 (p <0.001) (Figure 4.12). This high correlation shows that a large number of these proteins are transcriptionally regulated.

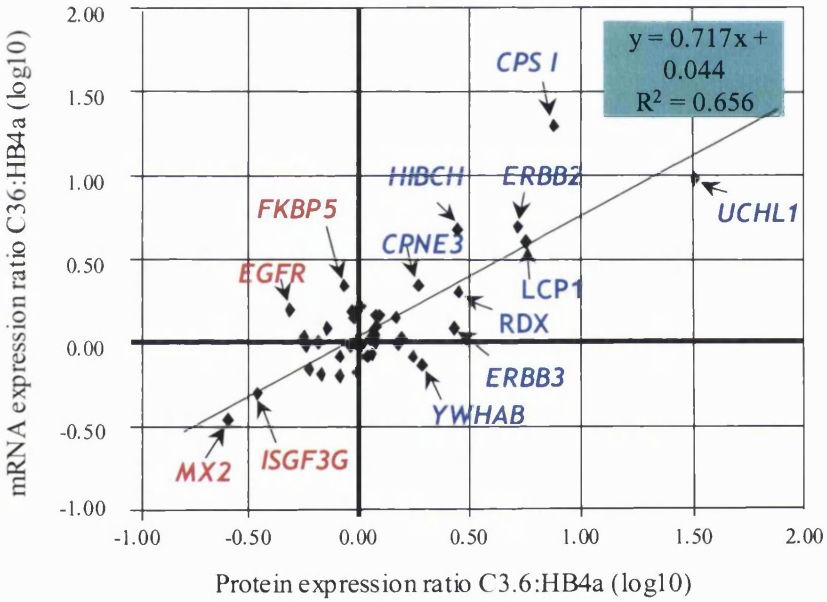


Figure 4.12: Comparison of protein and transcriptional profiling. Correlation between protein ratios and mRNA ratios of 43 genes in serum starved HB4a and C3.6 cells. Protein fold ratios were obtained from the 2D-DIGE analysis and from 1D-western blotting data. The mRNA data was derived from microarray experiments described in more detail by White *et al.* (submitted data). Names correspond to gene card names. Some of the genes listed were also found differentially regulated at the protein level by 2D-DIGE analysis of serum-starved HB4a and C3.6 cells. These are: PGP9.5 (UCH-L1), copine III (CPNE3), carbamoyl-phosphate synthetase (CPS1), L-plastin (LCP1), radixin (RDX), FK 506 binding protein 5 (FKBP5), 14-3-3 β (YWHAB), 3-hydroxyisobutyryl-CoA hydrolase (HIBCH), MxA (Mx2).

Some genes were constitutively up-or down-regulated at all time points in the microarray experiment, and 20 genes were found to be up-regulated and 27 genes down-regulated in the C3.6 compared to the HB4a cell line. Amongst these, many were also found to be differentially up- or down-regulated at the protein level in the proteomics analysis conducted in the 2D-DIGE analysis. These were the metabolic enzyme carbamoyl-phosphate synthetase 1 (CPS1), the actin bundling protein, L-plastin (LCP1), PGP9.5 (UCHL1) and copine III (CPNE3). Moreover, the finding that a significant group of genes involved in interferon signalling had altered mRNA expression (down-regulated in the C3.6 cells), correlated with the decreased protein levels of MxA (or Mx2, an interferon induced viral resistance protein) and ISGF3γ, the interferon-responsive gamma subunit of the ISGF3 transcription complex, which has been reported to play an important role in the transcriptional regulation of many interferon-regulated genes. The significance of these changes is currently being investigated in our laboratory (Bertani M. and Timms J.F.).

Global changes in protein features and gene expression over the time course were also compared, and surprisingly an inverse correlation between protein levels and mRNA levels was found, as cells progressed through the cell cycle (Figure 4.13). Indeed, whereas more proteins were up-regulated in the C3.6 cells, more mRNAs were down-regulated in these cells. Although this data appears to be contradictive, the rate of protein turn-over in both cell lines needs to be further investigated. Amongst the changes observed between the cell lines, some were affected by growth factor treatment, suggesting a role downstream of ErbB receptors. In addition, the transcriptional ratio did not always correlate with protein expression ratios. For instance, EGFR and ErbB-3 were down- and up-regulated respectively at the protein level, but not at the mRNA level. This could suggest that these proteins are not regulated at the transcriptional level, but probably post-translationally and a differential protein turn-over in response to ErbB-2 overexpression could be involved. Further work was carried out to examine changes in ErbB receptor levels and processing (described in Chapter 6).

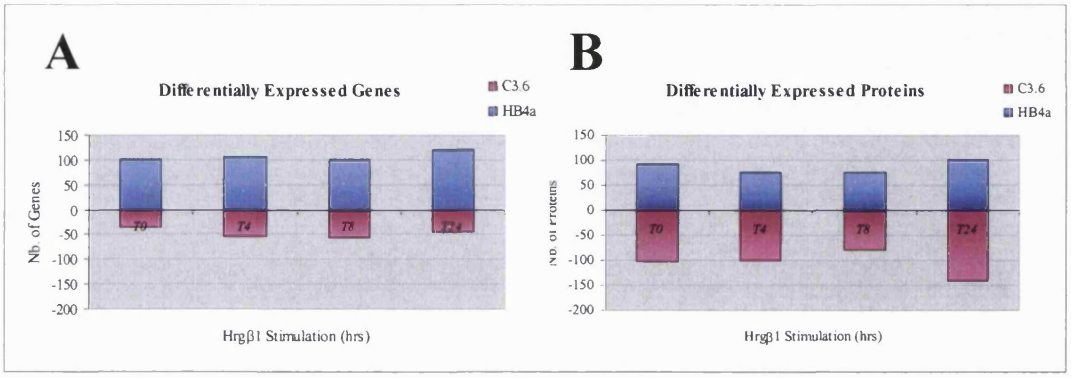


Figure 4.13: Comparison of expression profiles of proteins and mRNA. A) Number of protein features with a greater than 1.3 fold difference in abundance between HB4a and C3.6 at each time point of stimulation with Hrgβ1. Only average values from triplicates with *p* values <0.05 were included (see Table 4.2). B) Number of differentially expressed genes at all time points of stimulation with Hrgβ1. Ratios of mRNA levels in C3.6 versus HB4a with a geometric mean of >2.0 or <0.5 were included.

Taken together, combining both gene expression and protein expression data has provided important information concerning protein post-transcriptional regulation and function. Generating tools to filter this complex data is becoming essential for the future of functional genomics and proteomics.

4.8 Conclusions and Discussion

During the time frame of these experiments, there were few reported biological studies using DIGE analysis (Tonge *et al.* 2001) and it was therefore important to validate this technique as a useful functional proteomic tool. The evaluation of 2D-DIGE for the expression profiling of a model cell system of ErbB-2 overexpression has been investigated in this Chapter. This work has shown that DIGE analysis could provide sensitive, accurate and quantitative analysis of differential protein regulation in a model

cell system of breast cancer. Reproducible and significant differences could be detected both between the parental and the ErbB-2 overexpressing cell lines and in response to Hrgβ1 or EGF stimulation.

The major advantage given by DIGE analysis was that this approach enabled the study of multiple samples in a reproducible and statistical manner. A statistically meaningful analysis of eight different samples (two cell lines and four growth conditions) run in triplicate could be achieved in a single 12-gel experiment. This is superior to conventional 2DE expression analysis, both in terms of accuracy, speed and workload. Here, a time course study of the effects of growth factor treatment on the ErbB-2-mediated protein expression profile was investigated. A pool of all samples labelled with Cy2, representing an average of all samples analysed, was used as an internal standard to be run on each gel and used for cross-gel comparison and quantitation on a spot-by-spot basis. Two conditions were thus compared (the cell lines and the time points) in a statistical analysis using a T-test, allowing significant differences to be assessed. In this study, the number of spots that could be matched between gels was high, which showed the high potential of 2D-DIGE as a reproducible method. Alban *et al.* recently carried out a validation of DIGE for statistical analysis using DeCyder and confirmed the findings observed in this study (Alban *et al.* 2003).

Proteins found differentially expressed between the cell lines under serum-starved conditions were further analysed by peptide mass mapping using MALDI-MS. This data revealed the compatibility of cyanine dye covalent labelling with subsequent identification by MALDI-MS, as the proteins picked could be identified with confidence. Additionally, this showed that post-staining with SYPRO Ruby was compatible with downstream protein identification in support of previous work (Lopez *et al.* 2000). Due to the high protein coverage and good statistical results obtained during database searching, no additional sequence analysis by LC-ESI-MS/MS was required and the proteins herein were identified solely by MALDI-MS peptide mass fingerprinting. All these identifications will nonetheless require further validation work, in order to be fully confident about the identities of the proteins as well as the changes in abundance observed. This validation process and further characterisation of the proteins will be

discussed in the following Chapter.

A subset of proteins identified in this study suggested that the model cell system employed was a useful model of cancer progression. Indeed, as discussed previously, amongst the proteins analysed, many had been previously reported to be deregulated in different types of cancers, implicating their role in transformation. Of the 27 proteins identified, 8 had already been shown to be implicated or differentially regulated in cancers. Therefore, although the cell lines cannot form tumours in nude mice, and hence do not fulfil the profile of fully transformed cells, they do display a protein expression pattern which is representative of cancerous cells, perhaps at an intermediate stage in the tumourigenic programme.

This work revealed that there were more differences between the un-stimulated cell lines than could be observed between the time points of stimulation with Hrgβ1 (or EGF). This could suggest that the level of sensitivity of 2D-DIGE analysis was not sufficient to enable the analysis of relevant proteins expressed at low levels such as many signalling molecules. Unfortunately, this is also the case with many 2DE gel-based approaches which are limited by the amount of proteins which can be reproducibly and effectively resolved, and by the sensitivity limits of current protein detection methods. Moreover, as previously discussed, changes of proteins of low abundance volumes were more prone to variations and therefore more difficult to statistically validate. They would have therefore been omitted during the statistical filtering. Despite this, changes were detected which may have an impact on the understanding of ErbB-2-dependent transformation.

The parallel comparison by proteomic and microarray analysis showed that although the list of differentially expressed genes was much larger than for the proteomic analysis (667 genes compared to 27 proteins), an overlapping subset of genes was found to be differentially regulated between the cell lines, reflecting a high correlation between the transcription profile and the protein expression profile. Constant effort is being applied to combine the complex data generated from both proteomic analyses and microarray technology (Gygi *et al.* 1999; Chen *et al.* 2002; Patel *et al.* 2002). Some reports have shown a discordance between both mRNA and protein levels. Chen *et al.*

reported in a large-scale study of protein expression and mRNA expression in lung adenocarcinomas, that there was no significant correlation found (correlation coefficient = - 0.025) (Chen *et al.* 2002). In the study carried out here, a limited number of proteins was identified and quantitated (43), which were also represented on the microarray chips. However, a good correlation (correlation coefficient = 0.81; p value <0.001) between protein expression ratios and mRNA ratios was observed in this experiment. This indicated that many of these proteins are regulated at the transcriptional level. However, many changes observed by microarray could not be detected by 2D-DIGE and vice versa, which could be due to low level of protein detection or absence of the gene of interest on the cDNA chips. Some proteins showed an inverse correlation between average protein ratios and gene expression ratios, suggestive of differential protein turn-over between HB4a and C3.6. Preliminary work is underway to explore the degradation routes taken by activated receptors and other proteins that display differential expression profiles between the two cell lines. Due to the up-regulation of a ubiquitin hydrolase (PGP9.5), in the C3.6 cells, the ubiquitin-proteasome pathway was also selected for further investigation, with particular focus on the ErbB receptor regulation (Chapter 6).

Taken together, this data showed that the altered phenotype (morphology, proliferation and adhesion) observed between HB4a and C3.6 could be represented at the protein level. Therefore although not all differentially expressed proteins could be identified using this approach, 2D-DIGE analysis represents an important tool for large-scale protein analysis in a cellular system. Combining 2DE analysis with other methods of pre-fractionation or protein separation by liquid chromatography could enable improved understanding of the proteomic impact of ErbB-2 overexpression. Overall, this work has provided key information on the regulation of potential targets of ErbB-2 mediated transformation in an ErbB-2 overexpressing cellular system and points to potential diagnostic and prognostic markers. Moreover, this data also provides new insights into the involvement of ErbB-2 in cellular processes that were hitherto unknown.

Chapter 5 : VALIDATION OF ErbB-2-DEPENDENT DIFFERENTIAL EXPRESSION AND TARGET CHARACTERISATION

5.1 Introduction

The previous Chapter showed the power of 2D-DIGE for comprehensive protein expression profiling in the study of ErbB-2-dependent cell transformation. Differentially expressed proteins with a link to ErbB-2 overexpression were identified in the HB4a and C3.6 model cell system by 2D-DIGE and subsequent peptide mass mapping using MALDI-MS (Chapter 4). Sufficient protein sequence coverage as well as high mass accuracy were achieved, avoiding the need to perform any further analysis by tandem mass spectrometry. Moreover, the identifications made were reproducible since the same proteins were identified from selected spots picked from multiple runs. In addition, all proteins listed in Table 4.1 (Chapter 4) were identified as single component proteins in each spot, improving the confidence in protein identification. However, peptide mass mapping does not provide absolute confidence in protein identification (Chapter 1) and it was therefore important to confirm protein identities and changes in protein abundance using other biochemical tools. To validate the differential expression data and the mass spectrometric identification of proteins and to establish a correlation with altered receptor expression, a series of biological experiments was planned. In this Chapter, the validation of protein identities obtained in the 2D-DIGE comparison of HB4a and C3.6 cells was performed and further functional analysis of potential targets of ErbB-2 overexpression subsequently assessed. Protein identities were examined by conventional western blotting analysis and then, changes in protein expression levels were validated in other HMLEC clones and breast tumour cell lines in order to identify specific differences attributable to ErbB-2 overexpression or cell-type specificity rather than changes associated with clonal variation. The main proteins of interest, which could be validated and with a potential role in ErbB-2 mediated transformation were selected for further characterisation using

other biochemical and cell biology techniques.

5.2 Validation of differences observed by 2D-DIGE

The main approach used to validate protein expression levels relied on the specific detection of proteins by immunoblotting. However, since the systematic generation of new antibodies is time consuming and costly, only proteins for which antibodies were available could be further analysed using this method. Total cellular extracts in NP40 lysis buffer or urea lysis buffer were analysed using a panel of antibodies specific for the proteins of interest (Table 2.2, Chapter 2). Levels were initially compared in the parental cell line HB4a, in the ErbB-2 overexpressing cell line, C3.6, and in another clone, C5.2, that overexpresses even higher levels of ErbB-2 (Figure 5.1). In each western blotting experiments carried out herein, cdk4 was used as a loading control as it was shown that in all conditions tested, levels of cdk4 remain unchanged (data not shown). In addition, antibody specificity was confirmed by the fact that descrete bands were detected at the expected molecular weight.

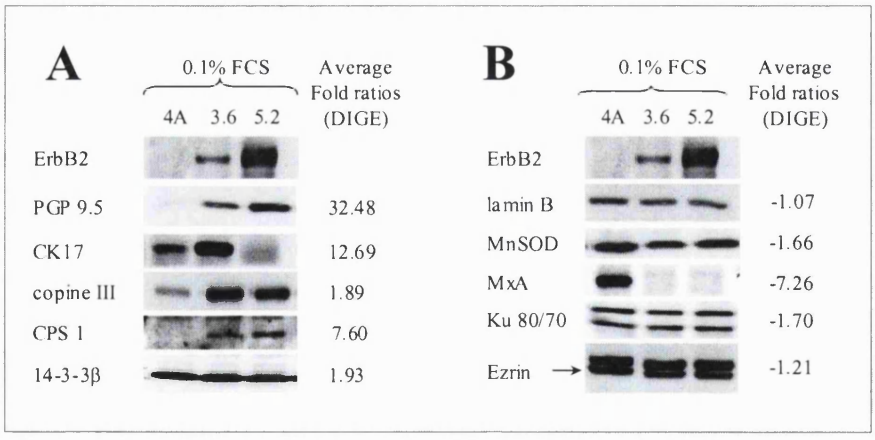


Figure 5.1: Validation of differential protein expression analysis. Western blotting analysis of total cellular extract from HB4a, C3.6 and C5.2 are herein compared by 1D-SDS-PAGE and western blotting with specific antibodies. The protein expression profiles are compared to levels obtained by statistical analysis in the comparison of HB4a and C3.6 under serum-starved conditions by 2D-DIGE. Average fold ratios are from protein expression levels in C3.6 versus HB4a within p values <0.05. A) Proteins which were up-regulated in the C3.6 cells compared to the HB4a cells. B) Proteins which were down-regulated in the C3.6 compared to the HB4a cells. Levels are compared to ErbB-2 levels.

The expression profile under a number of different conditions was compared to the average protein expression ratios obtained by 2D-DIGE analysis between HB4a and C3.6 cells under serum starved conditions. The levels of PGP9.5 (UCH-L1), copine III, cytokeratin 17, lamin B, MnSOD and MxA were monitored and matched to the levels observed in the 2D-DIGE experiment (Figure 5.1). PGP9.5, copine III, CPS1 and CK17, were found to be up-regulated in the C3.6 cells and the increase in expression detected correlated with the fold-ratios observed by DIGE. Additionally, the levels of PGP9.5, copine III, CPS1 were also increased in the C5.2 cells suggesting that the level of expression observed is directly related to ErbB-2 expression levels. However, CK17 was not detected to any significant degree in the C5.2 cells. This may correlate with the fact that these cells have a distinctly altered morphology, compared to HB4a and C3.6, with approximately 20% of cells displaying a fibroblastic phenotype. Furthermore, lamin B1, MnSOD and MxA were shown to be down-regulated in the C3.6 cells, as was observed in the DIGE experiment. The levels were also down-regulated in the C5.2 cells.

Differential protein expression could not be validated for other targets such as the Ku autoantigen, 14-3-3β and the ERM proteins (ezrin, radixin, moesin), which displayed similar levels of expression by western blotting in the three cell lines tested. Several hypotheses could be made to explain this discrepancy. Firstly, the wrong protein could have been picked, possibly due to the shift in molecular weight between the cyanine dye-labelled and un-labelled fraction of proteins. Furthermore, this could be due to false protein identification by peptide mass mapping by MALDI-MS and further protein sequencing data would thus be necessary. However, the differential expression observed for the ERM proteins was also low, with only a -1.27 fold-difference and which although significant in the 2D-DIGE analysis may not be detected using 1DE-western blotting, which has a narrow linear range of detection. In addition, the Ku autoantigen sub-units are nuclear proteins and the differential expression observed by 2D-DIGE could be due to differences in protein solubility in the two different buffer systems used for 2DE and 1DE. In support of this, differences in extraction were also observed for CK17, which could only be extracted in a urea lysis buffer (data not shown). It would therefore be important to validate all these changes by 2DE-western blotting in order to be more

confident about the differential expression observed. The other protein, which showed no changes by western blotting analysis was the 14-3-3β protein. This protein is a member of a highly homologous family of proteins and although the mass spectrometric coverage was high (85% of protein coverage), it was not sufficient to be confident that the isoform picked was 14-3-3β. Finally, the changes observed on a 1DE gel do not necessarily reflect the levels observed by 2DE, since post-translational modifications cannot often be detected on 1DE gels. Indeed, the Ku autoantigen, MxA and L-plastins were detected as multiple spots on 2DE gels, suggestive of post-translational modifications.

5.3 Further analysis of differentially expressed protein targets

Further validation and characterisation was undertaken to establish the link between ErbB-2 overexpression and the observed differential expressions and to address the mechanisms by which ErbB-2 regulates the expression profiles of selected protein targets. First, expression levels were assessed in other cellular clones derived from the HB4a cells which overexpress variable levels of ErbB-2. This was important, as it would verify that the changes observed were real differences due to the overexpression of ErbB-2 and were not due to clonal variations introduced during the selection of ErbB-2 overexpressing clones. In addition, in order to evaluate the effect of serum on the protein expression profiles, cells were compared under serum-starved and randomly growing conditions (10% FCS). Total cellular extracts of each cell line were resolved by 1DE-gel and levels of particular proteins were examined by immunoblotting with appropriate antibodies (Figure 5.2).

For each protein tested, the expression levels observed were similar under both growth conditions tested, which suggests that these differences were not attributable to the state of growth of the cells. Each clone displayed distinct levels of ErbB-2 expression with the C5.2 and C5.5 clones expressing the highest levels of ErbB-2. The expression levels of other members of the ErbB family was also assessed. The levels of EGFR (ErbB-1) and ErbB-3 were variable in the clones with the protein expression level of EGFR displaying an inverse correlation to the levels of ErbB-2. On the other hand, the levels of ErbB-3 were increased in the C3.6 cells compared to the HB4a cells. This would

suggest that ErbB-2 can modulate the expression of other ErbB receptor family members (see Chapter 6). The levels of CPS1 correlated with the levels of ErbB-2, except in the clone C5.5 where the levels were low. Interestingly, the levels of PGP9.5 reflected closely the levels of ErbB-2 in the various clones suggesting that ErbB-2 signalling plays a role in the regulation of PGP9.5 expression. The level of copine III expression was found to generally correlate with the level of ErbB-2 overexpression, and particularly with the levels of ErbB-3. Thus, these differences between HB4a and C3.6 appear to be related to ErbB-2 overexpression, and do not seem to be specific to the C3.6 clone.

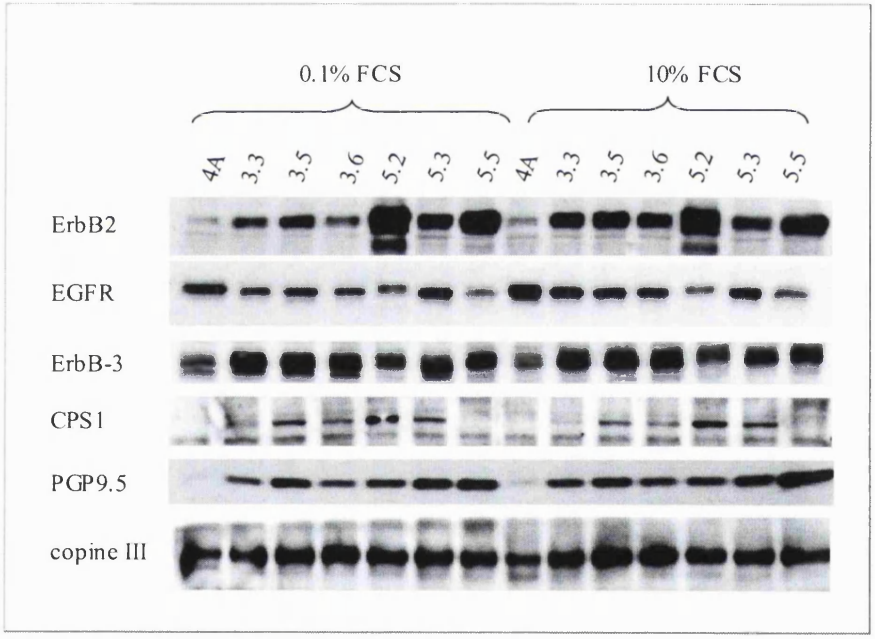


Figure 5.2: Relative protein expression in HMLEC clones. Protein expression levels were assessed in serum-starved (0.1% FCS) cells and randomly growing cells (10% FCS) by immunoblotting. 30 μg of total cellular extract was resolved on a 1DE gel and immunoblotted with specific antibodies (see Table 2.2 Chapter 2). Changes in protein expression were assessed for PGP9.5, Copine III, and compared to expression levels of ErbB-2, ErbB-3 and EGFR.

Subsequently, protein expression levels for several of the proteins were measured in a panel of commonly used breast tumour cell lines, mostly established from pleural effusions of breast cancer patients with metastatic disease. These cell lines are known to be a diverse group which have differential proliferative properties and are morphologically diverse. The protein expression profile of each ErbB family member was assessed by immunoblotting and each cell line displayed a distinct pattern of expression of the ErbB receptors (Figure 5.3). The lum878 cell line which was used as a control normal luminal epithelial breast cell line was also immortalised with SV40 Large T antigen in combination with hTERT. The BT474, SKBr3, MB361 and MB453 breast cancer cell lines expressed high levels of ErbB-2 which were similar to the level of expression detected in the C3.6 and C5.2 clones. In addition, some of these cell lines also expressed relatively high levels of ErbB-3 (SUM185, BT474, SKBr3, MB361, MB435, MB453 and MCF7 cell lines). None of these cell lines expressed substantial levels of EGFR and this high level of expression of EGFR was found to be specific to the luminal epithelial cell lines. These observations were in agreement with the protein levels observed in the HB4a and C3.6 cells, hence supporting the correlation of ErbB-2 expression level with the regulation of other ErbB receptor family members (further discussed in Chapter 6). MxA, which was found to be down-regulated in the C3.6 cell line compared to the parental HB4a cell line, was generally expressed at comparatively low levels in all breast cancer cell lines but not in the normal luminal cell line, lum878. Copine III, which was found to be up-regulated in the C3.6 compared to the HB4a cells was generally shown to be up-regulated in the panel of breast cancer cell lines tested. The levels of copine III were closely correlated to the levels of ErbB-2 and ErbB-3 and relatively high protein expression levels were detected in the SUM185, BT474, SKBr3 and MCF7 cells. This data suggests that copine III overexpression may be related to enhanced signalling through both ErbB-2 and ErbB-3. Although PGP9.5 expression levels were highly correlated to the levels of ErbB-2 in the HMLECs, this protein was not detected to any significant degree in the other breast tumour cell lines except in the control luminal epithelial cell line, lum878. This would suggest that PGP9.5 is a specific marker of ErbB-2 overexpression in luminal epithelial cells.

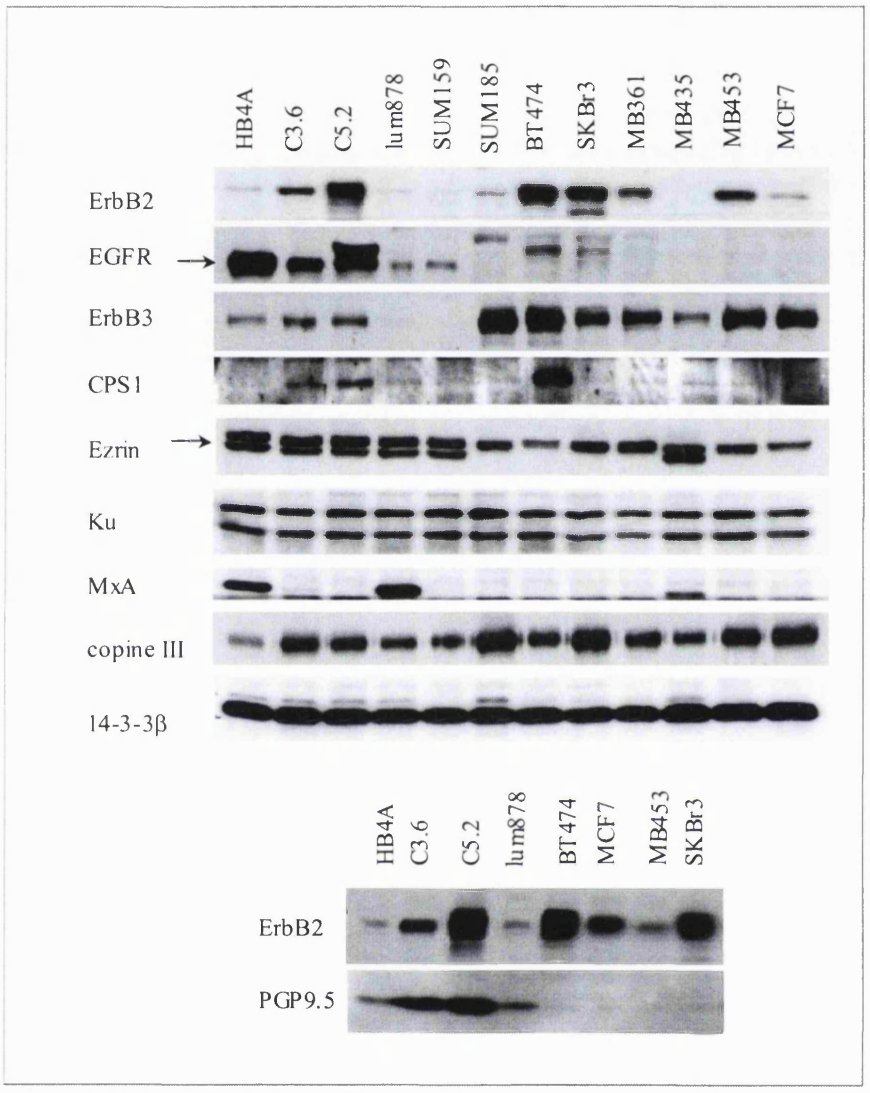


Figure 5.3: Investigation of protein expression profiles in a panel of human breast cancer cell lines. The expression levels of potential targets identified by 2D-DIGE analysis were assessed in a panel of breast cancer cell lines. Cells were grown, serum-starved and total cellular lysates were resolved by 1DE-SDS-PAGE and protein levels were analysed by immunoblotting with specific antibodies.

Further validation of these results should be assessed by immunohistochemistry analysis of primary breast tumours. CPS1, which was overexpressed in the C3.6 and C5.2 cells, was not generally expressed in other breast tumour cell lines tested, apart from the BT474 cells, which also expressed high levels of ErbB-2. Interestingly, although no significant changes in expression of the ERM protein ezrin could be detected between HB4a and C3.6 cells, Ezrin was not detected in the breast cancer cell lines overexpressing ErbB-2 and ErbB-3. The levels of Ku autoantigen and 14-3-3β were roughly equivalent in all the breast cell lines tested.

Taken together, this data enabled the validation of changes observed by 2D-DIGE analysis, some of which were correlated with the levels of ErbB-2 overexpression in HMLEC clones overexpressing ErbB-2 and other breast cancer cell lines.

5.4 Further characterisation of specific target proteins

Further analysis of the proteins of interest was carried out using specific antibodies and the effect of growth factor treatment and specific inhibitors on their protein expression profiles was assessed. In Chapter 4, a large-scale analysis of the effect of growth factor treatment, Hrgβ1 (and EGF), on the proteome and the transcriptional profile of HB4a and C3.6 cells was undertaken. Statistical analysis showed that few differences in protein expression profile could be detected between the time points of treatment with growth factor, and most differences detected were between the different cell lines. According to the combined protein expression and mRNA transcription data and to the complete validation undertaken on the several HMLEC clones and breast tumour cell lines tested, some target proteins with potential interest in the ErbB-2-dependent cell transformation, such as MxA, MnSOD, copine III and PGP9.5, were selected for further characterisation.

5.4.1 MxA, MnSOD and interferon signalling

Investigation of the function of MxA and the impact of its down-regulation in response to ErbB-2 overexpression was undertaken in this study. MxA is a member of the Mx family of proteins which are interferon-induced GTPases belonging to the dynamin superfamily of large GTPases. The Mx proteins have been described as important regulators of the antiviral response (Arnheiter *et al.* 1990) and they were shown to be largely expressed in interferon treated cells (Simon *et al.* 1991). In the 2D-DIGE study, MxA was found to be significantly down-regulated in the C3.6 cells compared to the HB4a cells (-7.26 fold-difference) under serum starved conditions. In addition, treatment with Hrgβ1 further reduced the expression of MxA and protein expression was decreased in the C3.6 and the HB4a cells after 4 hrs of growth factor treatment (Figure 5.4A). This result would suggest that the level of MxA expression could be inversely related to the signalling activity of ErbB-2. MxA protein expression and its mRNA transcription profile were also found to be correlated, indicating that down-regulation of MxA had taken place at the transcriptional level. This data was concomitant with the finding that a significant number of interferon regulated genes were also found to be down-regulated at the mRNA level in the C3.6 cells compared to the HB4a cells (Chapter 4 and White S.L. unpublished data). Due to the negative effect of IFN on cellular proliferation (Stark *et al.* 1998; Borden *et al.* 1999), the down-regulation of interferon-regulated genes is thought of as a mechanism by which ErbB-2 could promote cellular proliferation.

Further investigation to assess the effect of growth factor and interferon (IFN) treatment on the protein expression profile of MxA was therefore carried out. The HB4a and C3.6 cells were treated with EGF, Hrgβ1, IFNβ or IFNγ for 15 min and 4 hrs and the levels of expression of MxA and another family member, MxB, were estimated by 1DE-western blotting (Figure 5.4B). No significant effects of growth factor treatment on MxA expression could be detected in either cell lines with EGF or Hrgβ1, although a slight decrease in expression was observed in the HB4a cells after 15 min and 4 hrs of treatment with Hrgβ1, which is in agreement with the expression profile observed by 2D-DIGE analysis.

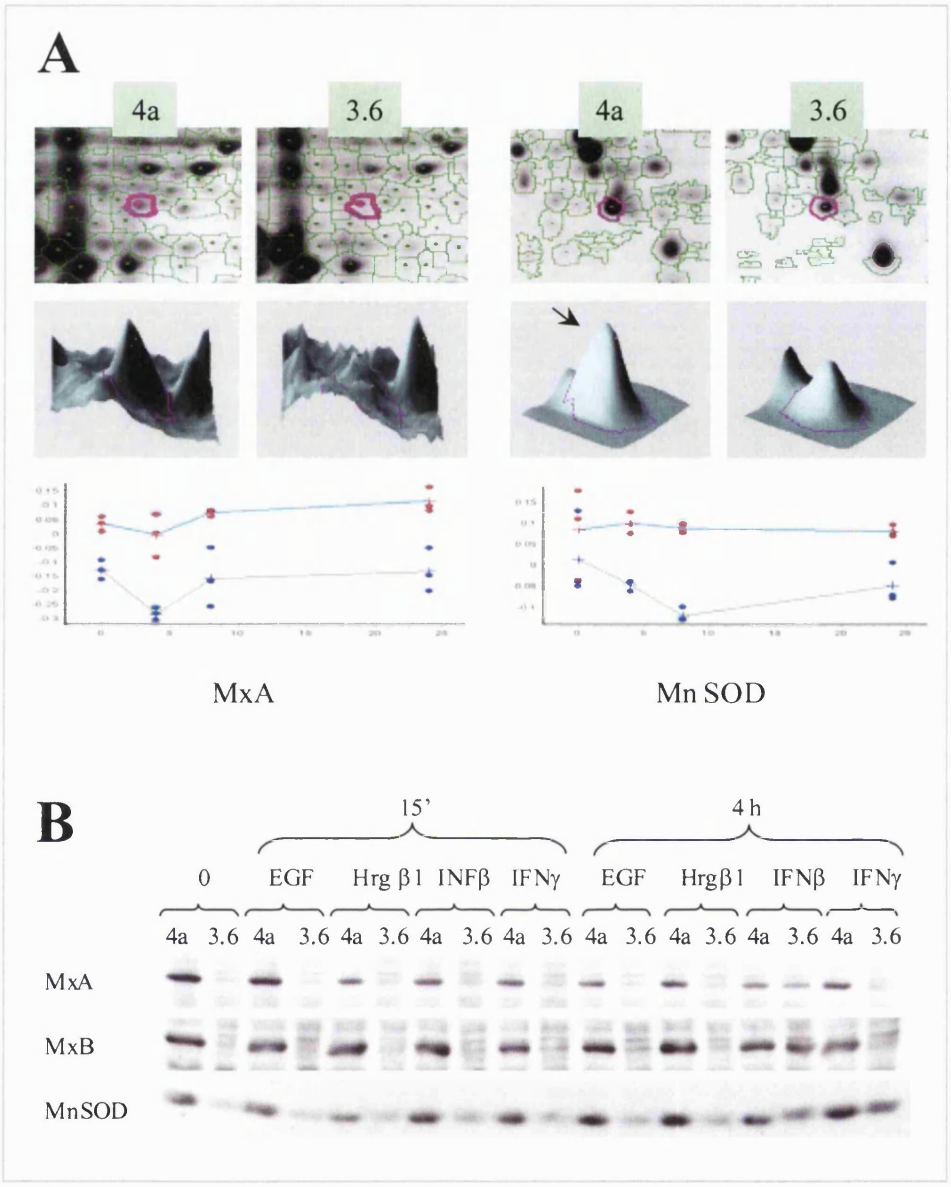


Figure 5.4: Study of the regulation of Mx proteins and MnSOD in HB4a and C3.6 cells. A) The effect of Hrgβ1 treatment on MxA and MnSOD protein expression levels was investigated by large-scale 2D-DIGE analysis. Changes in protein abundance were analysed using the DeCyder software and were plotted for each time point of treatment with Hrgβ1 as shown in the graphic (HB4a is shown in red and C3.6 is shown in blue). B) The HB4a and C3.6 cells were treated with 1 nM of EGF or Hrgβ1 or 1000 U/ml of IFNβ or IFNγ for 15 min or 4 hrs. Immunoblotting was performed with MxA, MxB and MnSOD specific antibodies.

However, the Mx protein expression level was up-regulated in the C3.6 cells but not in HB4a after 4 hrs of treatment with IFN and this activation was specific for the type I interferon (IFNβ), which is in agreement with previously published data (von Wussow *et al.* 1990). Therefore this analysis showed that the protein expression levels of MxA and MxB could be induced in the C3.6 cells and that the altered expression in C3.6 versus HB4a is not likely to be due to the selection of mutations during clone derivation.

In addition, another important protein which was found to be down-regulated in C3.6 compared to HB4a cells, in 2D-DIGE and microarray analyses, was MnSOD (Figure 4.4A and Chapter 4). The expression of this protein has been correlated to the level of oxidative stress in cells but was also shown to be induced by both type I and type II interferons (Raineri *et al.* 1996). Treatment with Hrgβ1 greatly reduced the level of expression of MnSOD in the C3.6 cells after 8 hrs of treatment, but no major effect on the protein profile was detected in the HB4a cells, as revealed by 2D-DIGE analysis (Figure 5.4A). The effect of growth factor (EGF and Hrgβ1) and IFN (IFNβ and IFNγ) treatment on the expression profile of MnSOD was also assessed by western blotting (Figure 5.4B), and similarly to MxA, no significant effect of growth factor treatment could be detected by western blotting. On the other hand, IFN treatment greatly induced the expression of MnSOD in the C3.6 cells after 4 hrs of treatment.

Taken together, this revealed that the MxA and MnSOD gene expression in the C3.6 cells can be stimulated by interferon signalling to levels comparable to the ones observed in the parental cell line, HB4a, directly linking the ErbB-2 overexpression with the IFN signalling. Interestingly, a transcriptional activator of both MxA and MnSOD, was also found to be down-regulated at the mRNA level in the C3.6 cells in the microarray analysis. This protein, ISGF3γ, was also found to be down-regulated at the protein level by western blotting and current work is being carried out in the laboratory in order to further assess the impact of the down-regulation of this protein and the IFN signalling in an ErbB-2 overexpression system. Interestingly, co-stimulation with both growth factor and IFN abrogated the induction of MxA and MnSOD (Bertani M. and Timms J.F., unpublished data), which correlates with the effect of Hrgβ1 observed on the protein expression of these proteins. This observation emphasises the existence of cross-

talk between both IFN and growth factor signalling pathways. Taken together, these findings are in agreement with the up-regulation of signalling pathways which could stimulate high proliferative rate of the ErbB-2 overexpressing cells.

5.4.2 Copine III

Copine III belongs to the copine family composed of seven members. The copines contain two-C2 domains, which are either calcium dependent or independent binding motifs for phospholipids and inositol phosphates (Creutz *et al.* 1998). The copines are therefore thought to bind to phospholipids in membranes through the action of their two C2 domains in their activated calcium-binding form and have been proposed to play a role in vesicular trafficking. The copines also contain a C-terminal A-domain that is found in the extra cellular portion of integrins enabling them to bind extra cellular matrix proteins (Lee *et al.* 1995) (Figure 5.5) and copine III is reported to have an associated kinase activity (Caudell *et al.* 2000). This protein family therefore represents an interesting target to follow up to increase our understanding of ErbB-2 overexpression in cancer progression.

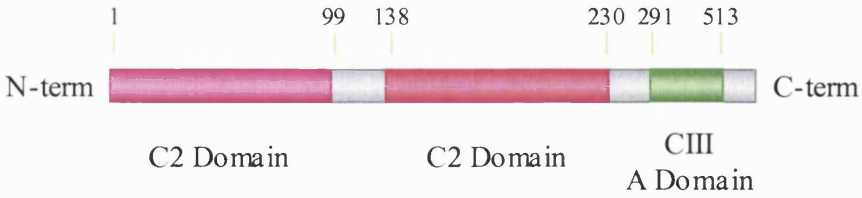


Figure 5.5: Domain composition of copine III. Copine III is a 60 kDa, member of the copine family of proteins. This protein contains two C2 domains and an A Domain and is thought to be phosphorylated on serine and threonine residues, although no site of phosphorylation has yet been found.

As mentioned in Chapter 4, copine III was found to be overexpressed in the C3.6 cell line compared to the parental cell line (Figure 5.6A). The protein expression levels were found to remain generally high in the C3.6 cell line in response to Hrgβ1 with an increase in expression observed after 8 hrs in the HB4a cell line. The levels were also assessed in the HMLEC clones and it was found that the level of copine III generally correlated with the ErbB-2 overexpression (Figures 5.6B and 5.2) although a positive correlation was also found with ErbB-3 expression in the panel of breast tumour cell lines (Figures 5.6C and 5.3). Thus, ErbB growth factor signalling seems to up-regulate copine III expression.

Since copine III was found to possess an associated kinase activity (Caudell *et al.* 2000), the copine III post-translational profile was investigated by 2D immunoblotting with rabbit anti-copine III antibody. The 2D pattern of copine III was compared between the cell lines under serum-starved conditions and after treatment with both EGF and Hrgβ1 for 8 hrs (Figure 5.7A). A train of spots was detected which could possibly be attributed to different post-translational phosphorylation of the protein. The 2DE-western blot images were aligned and the number and intensity of spots compared in each samples (Figure 5.7A). The same number of isoforms was detected in the HB4a and C3.6 cell lines under serum-starved conditions, although C3.6 staining was more intense. Treatment of the cells with either EGF of Hrgβ1 induced the up-regulation of copine III in the HB4a cells whilst levels remained high but unchanged in the C3.6 cells, in support of the results observed by 2D-DIGE analysis. This again shows that copine III is regulated by growth factor treatment (EGF and Hrgβ1) and reveals that several isoforms are expressed in the cell lines and that the expression level of all isoforms is induced similarly when ErbB-2 is overexpressed or when cells are treated with growth factors.

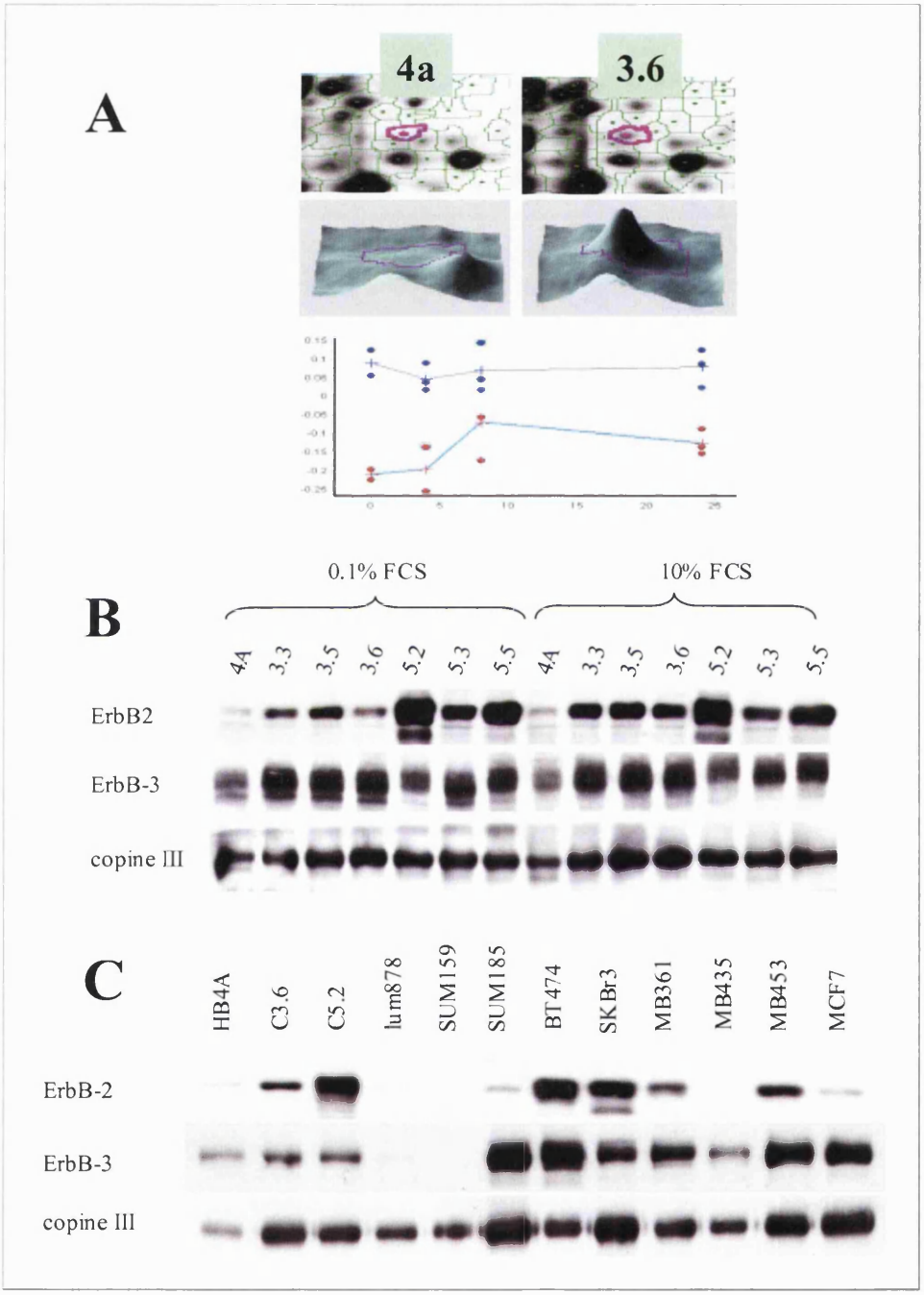


Figure 5.6: Copine III protein expression profile. A) Copine III was found up-regulated in the C3.6 cell line compared to the HB4a with a 1.89 average fold increase in the C3.6 cells. Expression levels were assessed by 1DE-western blotting with specific anti-copine III antibody and compared to levels of expression of ErbB-2 and ErbB-3 in HMLEC clones (B) and in a panel of breast cancer cell lines (C), as previously shown (Figures 5.2 and 5.3).

Although copine III does not possess classical protein kinase motifs, a sequence analysis showed that it displays homology to kinases. Previous work showed that copine III (from endogenous immunoprecipitate, chromatography–purification or recombinant) could induce *in vitro* phosphorylation of the substrate myelin basic protein (MBP) (Caudell *et al.* 2000). In order to further understand the role of copine III in the HMLEC cell lines and to relate ErbB-2 overexpression to copine III activity, endogenous copine III was immunopurified from HB4a and C3.6 cells treated with 1 nM of either EGF or Hrgβ1 using an anti-copine III polyclonal antibody. The immunoprecipitate was split into two fractions. The first fraction was used to validate the specificity of the antibody used, and to compare the levels of copine III in the different immunoprecipitates (Figure 5.7B) and the other fraction was used to perform an *in vitro* kinase assay (Figure 5.7C). Enrichment of endogenous copine III was achieved, although approximately 20% of the total copine III remained in the depleted fraction after the immunoprecipitation. Copine III migrates at approximately 60 kDa and therefore ran very close to the heavy chain of the immunoglobulin (IgG-heavy) used for the immunoprecipitation. It was therefore difficult to clearly estimate the levels of copine III that could be immunoprecipitated. In the *in vitro* kinase assay, incorporation of ^{32}P into the exogenous substrate, MBP, and the 60 kDa band, co-migrating with copine III, confirmed that copine III does indeed have an associated kinase activity (Figure 5.7C). Comparison of *in vitro* kinase assays showed a basally higher activity in the C3.6 cells compared to their parental counterparts. In addition, an induced phosphorylation was observed in the HB4a cells after 8 hrs of treatment with Hrgβ1 and EGF, consistent with the increase in expression levels (Figure 5.7B).

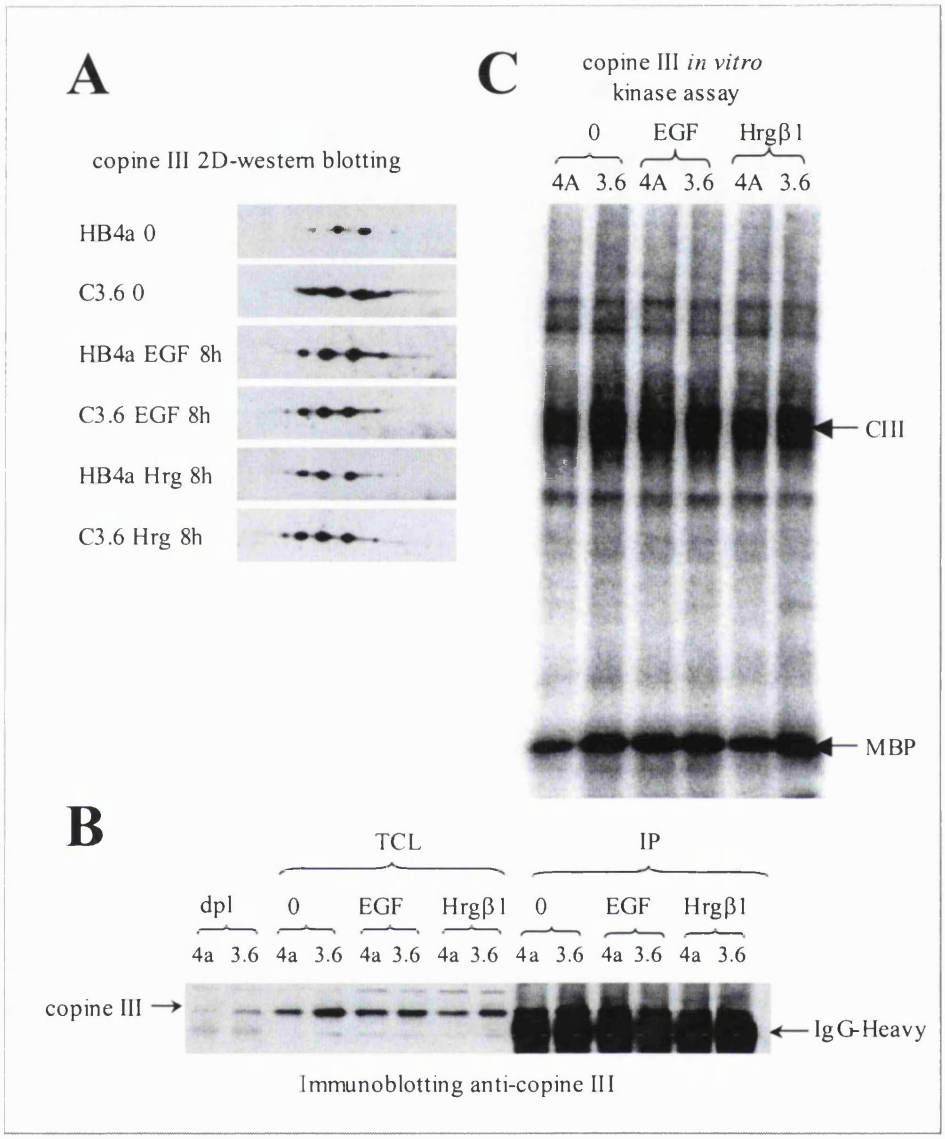
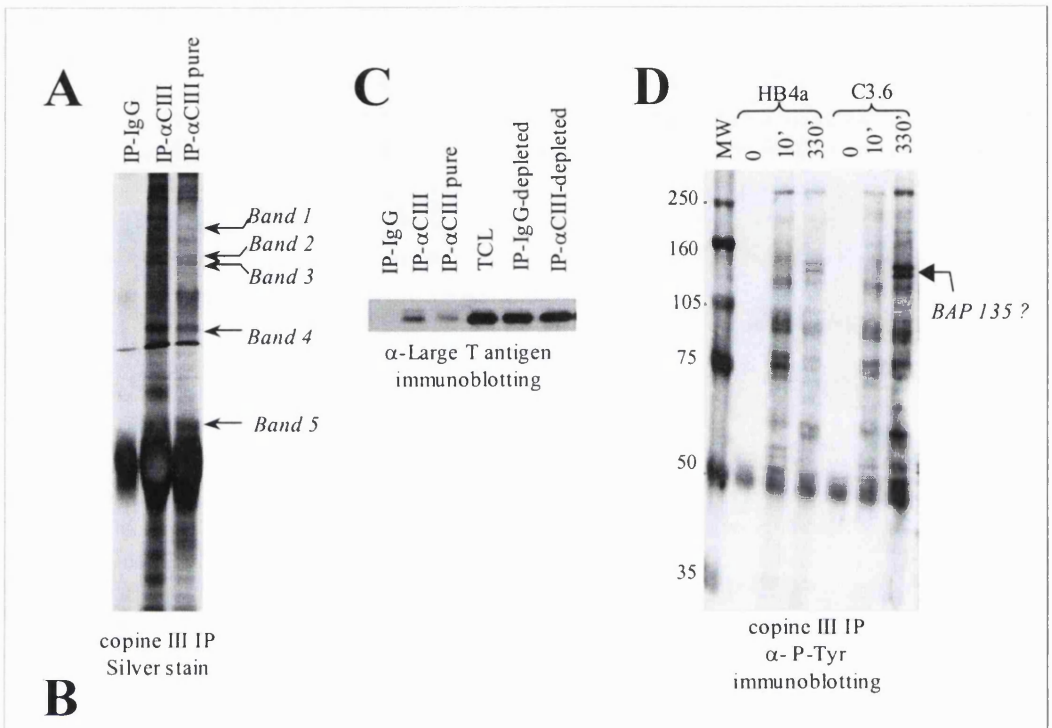


Figure 5.7: Copine III phosphorylation and growth factor regulation. In order to verify the effect of growth factor treatment on the regulation of copine III, cells were treated with 1 nM of EGF or Hrgβ1 for 8 hrs and compared to serum-starved cells. A) Total cellular extracts were resolved by 2DE gel, transferred to PVDF membranes and probed with anti-copine III antibody as described in Chapter 2. B) Endogenous copine III was immunoprecipitated (IP) and analysed by western blotting with anti-copine III antibody and levels were compared to the total cellular lysate (TCL) and the copine III-depleted lysate (dpl). C) An *in vitro* kinase assay was performed using MBP as an exogenous substrate. The assay was resolved by 1DE-SDS-PAGE and transferred to PVDF membrane and phosphorylated proteins were detected using a phosphorimager (as described in Chapter 2).

Therefore the activity observed seems to be constitutive and regulated by the level of copine III expressed. Caudell *et al.* also carried out an in-gel kinase assay showing that the kinase involved migrates at 60 kDa, suggesting that it could be copine III itself (Caudell *et al.* 2000). However, the experiment described here does not definitively show that copine III is the kinase.

In order to identify the putative kinase or other associated proteins, which could be interacting with copine III, a functional analysis was carried out by co-immunoprecipitation of proteins associated with copine III. A large-scale immunoprecipitation of copine III was carried out in the C3.6 cells, since they expressed the highest levels of the protein. Briefly, 4 mg of total cellular extract from C3.6 cells treated with EGF for 8 hrs was incubated with 40 μg of purified copine III antibody. Resolved immunoprecipitates (IP) were then visualised by mass spectrometry-compatible silver staining (Shevchenko) (Figure 5.8A). A control IP was performed with a non-immune rabbit IgG under the same conditions as well as unpurified anti-copine III. Specific bands appearing only in the anti-copine III IP were excised for further mass spectrometric analysis. Similar approaches have been previously employed for the identification of interacting partners of proteins of interest in order to understand the protein network in which a protein is involved. Ten bands were cut and analysed by peptide mass mapping by MALDI-MS. Amongst those, 5 proteins could be identified with confidence (Figure 5.8B) and others were shown to be serum proteins and immunoglobulin fragments from rabbit. Band 5 was identified as copine III, as expected, and the tryptic digest of this protein was further analysed for potential sites of phosphorylation by comparison of theoretical masses by peptide mass mapping (as described below). Four other bands could be identified with confidence as reported in Figure 5.8B. Acetyl CoA carboxylase α could be identified as a copine III interacting protein, however, this protein was later found to be present in other control immunoprecipitates, suggesting that this relatively abundant housekeeping protein interacts non-specifically with copine III (data not shown).



B

Band	Protein Name	Acc. Number	Mr	pI	% Coverage	Peptides Matched
Band 1	acetyl-Coenzyme A carboxylase alpha	XP_037617	265568	5.92	0.28	54
Band 2	Bruton tyrosine kinase associated protein, BAP-135	AAB48826	107970	8.06	0.35	26
Band 3	Bruton tyrosine kinase associated protein, BAP-135	AAB48826	107970	8.06	0.22	22
Band 4	Large tumor T-ag (SV40)	AAC59346	81568	5.83	0.33	28
Band 5	copine III	NP_003900	60131	5.6	0.33	20

Figure 5.8: Functional characterisation of copine III and analysis of co-immunoprecipitating proteins. A) C3.6 cells were treated for 8 hrs with 1 nM EGF and endogenous copine III was immunoprecipitated (IP) from 4 mg of total cellular lysate with α-copine III rabbit serum (α-CIII) or protein A-purified anti-copine III antibody (α-CIII pure). A non-immune rabbit IgG was used as a parallel control for non-specific binding. Immunoprecipitates were resolved on a 10% 1DE gel and proteins were visualised by MS-compatible silver staining. B) Specific bands were excised and further analysed by peptide mass mapping by MALDI-MS. Identified proteins are reported with their accession numbers, and their mass spectrometric match. C) Association of copine III and SV4a large T antigen. Levels of SV40 large T antigen were estimated by western blotting of the copine III IP and non-immune control IP (IP-IgG) and compared to the total cellular lysate (TCL) and depleted fraction (IP-IgG/α-CIII-depleted). D) Cells were treated for 10 min and 6 hrs with a tyrosine phosphatase inhibitor (100 μM pervanadate) and endogenous copine III was immunoprecipitated and analysed by western blotting with an anti-phospho-tyrosine antibody.

Surprisingly, the simian virus 40 (SV40) large T antigen (LT) was identified as a co-immunoprecipitating protein. This was confirmed by western blotting of copine III immunoprecipitate (Figure 5.8C). This interaction appears to be specific as no large T antigen is detected in the control immunoprecipitates; however, a small proportion from the total lysate was detected. As described in Chapter 1, HB4a cells were immortalised with the SV40 large T antigen, and this protein was therefore expressed to a significant degree in this cellular system. This viral protein is a helicase required to initiate the replication of viral DNA, although it also promotes cellular immortalisation by interacting with the p53 and Rb families of tumour suppressor proteins to inhibit their function. The potential interaction of the SV40 large T antigen with copine III was confirmed by western blotting of copine III immunoprecipitates with an anti-large T monoclonal antibody (Figure 5.8A). However, this finding could be non-specific as the SV40 large T antigen has been found mainly expressed in the nucleus (as will be verified later on) and copine III was thought to be involved in membrane trafficking. A recent investigation of SV40 large T-interacting proteins was carried out by yeast two-hybrid screening in rat PNET cell lines (Klawitz *et al.* 2001; Weggen *et al.* 2001). Interestingly, although LT is expressed in the nucleus of the cells, the authors found that SV40 large T antigen could interact with gelsolin, a cytoplasmic calcium-regulated protein which acts as an actin modulating protein. This would therefore suggest that this protein could also translocate to the cytoplasm.

Another potential copine III-interacting protein, which ran as a doublet, was identified as the bruton tyrosine kinase (Btk)-associated protein, BAP-135 (bands 2 and 3). BAP-135 is identical to a putative transcription factor, TFII-I (Roy *et al.* 1997) and was found to be associated *in vivo* with the plekstrin homology (PH) domain of Btk (Yang and Desiderio 1997). This protein is transiently tyrosine phosphorylated upon B cell antigen receptor (BCR) activation in a human B lymphoid cell line, thus suggesting that BAP-135 resides downstream of Btk in BCR-mediated signalling (Yang and Desiderio 1997). Interestingly, Btk is a downstream target of PI3K and the levels of copine III observed in the HMLEC clones correlated closely with the levels of ErbB-3, the only ErbB member to possess p85 binding sites and therefore able to activate

downstream signalling via the PI3 Kinase pathway. Given the role of BAP-135 in phosphotyrosine signalling, the phosphotyrosine profile of copine III immunoprecipiates was assessed by immunoblotting with an anti-phosphotyrosine specific antibody. Immunoprecipitates from cells treated with Hrgβ1 and EGF were analysed for their phospho-tyrosine signal and no significant signals could be detected (data not shown). In order to enhance the phosphorylation signal, cells were treated with a phosphatase inhibitor, pervanadate, for 10 min or 6 hrs prior to copine III immunoprecipitation (Figure 5.8D). A copine III co-immunoprecipitating protein with a similar molecular weight to BAP-135 and also running as a doublet was phosphorylated on tyrosine residues upon treatment with pervanadate. Amongst the overall phosphotyrosine pattern, the 130 kDa protein appeared to be selectively more tyrosine phosphorylated in the C3.6 immunoprecipitate compared to the HB4a cells. This would suggest a differential regulation of protein activity specific to the C3.6 cells, and it could therefore be involved in ErbB-dependent signalling. However, the basal phosphotyrosine activation of BAP135 seems to be of low stoichiometry as phosphorylated proteins could only be detected upon treatment with the phosphatase inhibitor.

Potential phosphorylation sites of copine III were also investigated by peptide mass mapping by MALDI-MS and revealed a potential site of phosphorylation on one serine residue. Indeed, one peptide appeared to be phosphorylated as its measured mass matched the mass of a theoretical serine-phosphorylated peptide. However, further work needs to be carried out in order to validate this potential site of phosphorylation. This could involve sequencing data and phosphatase treatment.

Copine III possesses two C2 domains suggesting a possible capability to interact with phospholipids and inositol polyphosphates (Creutz *et al.* 1998) and that through this mechanism it may localise to membranes. However, due to its potential interaction with SV40 large T antigen, characterised as a nuclear protein, it was interesting to investigate the sub-cellular localisation of copine III. Therefore HB4a and C3.6 cells were analysed by immunofluorescence using antibodies against copine III. Serum-starved cells were treated with 1 nM EGF for 10 min and fixed cells were stained for endogenous copine III (Figure 5.9A).

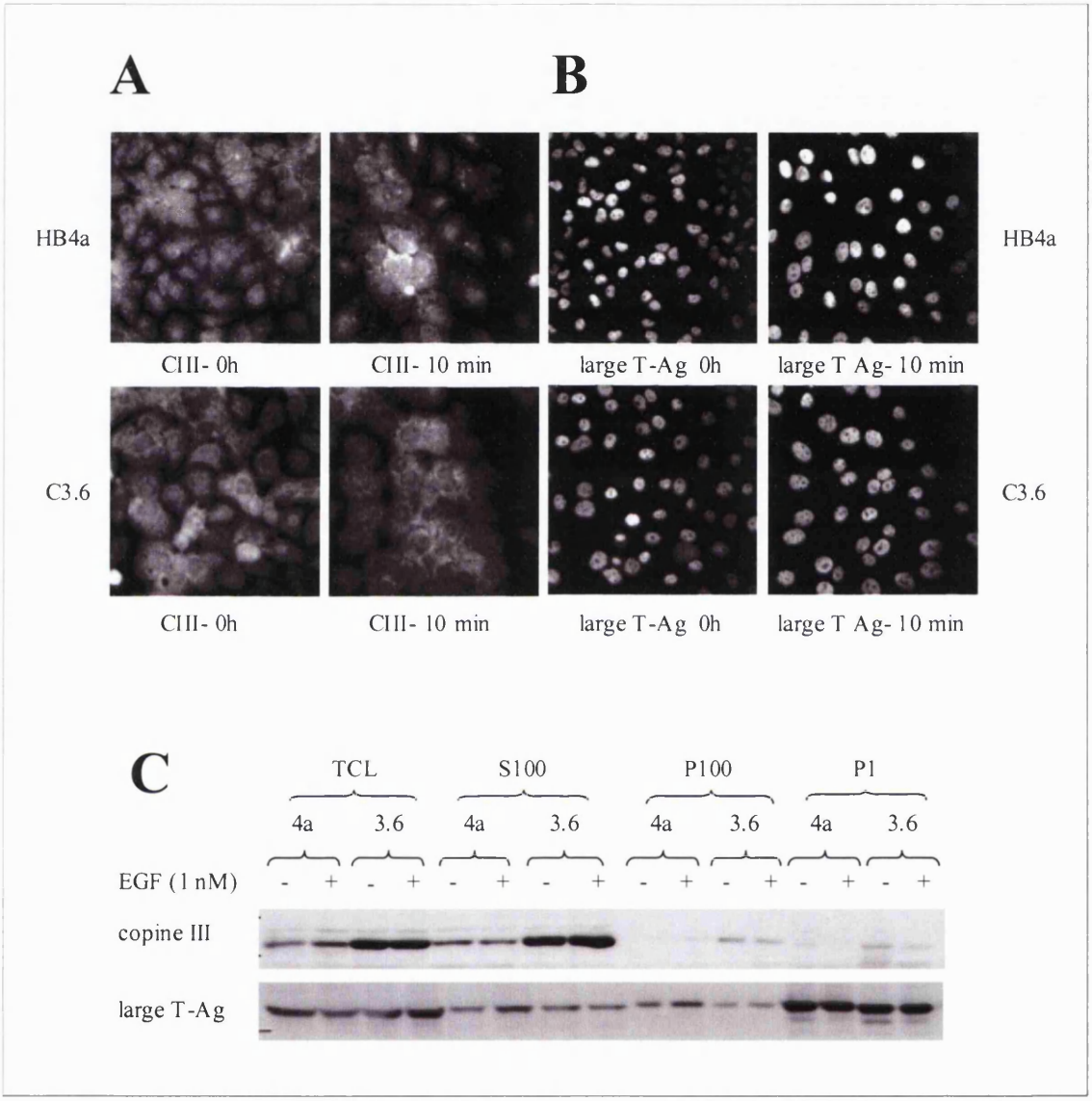


Figure 5.9: Copine III sub-cellular localisation investigated by immunofluorescence and crude cell-fractionation. A) Starved or EGF-stimulated HB4a and C3.6 cells (10 min) were fixed and stained with anti-copine III rabbit polyclonal antibody. B) Additionally, cells were co-stained for large T antigen following the same process. C) The HB4a and C3.6 cell lines were serum-starved for 48 hrs and treated for 4 hrs with 1 nM EGF. A crude sub-cellular fractionation was carried out and protein levels of copine III and large T antigen were measured by 1DE-western blotting in each fraction.

A low intensity signal was detected, which appeared to be specific for copine III (by comparison with a control immunostainning with the secondary antibody alone, data not shown) and localised mainly to the cytoplasm of the cells but with some nuclear staining apparent (Figure 5.9A). Interestingly, after treatment of the cells for 10 min with EGF, copine III seemed to be accumulated at cell-cell junctions. However, this data needs to be confirmed as the sub-cellular localisation observed could also be related to the stage of confluency of the cells. Additionally, immunostaining was carried out for the SV40 large T antigen under the same growth conditions (Figure 5.9B). A strong nuclear signal was detected with no staining in the cytoplasm of the cells. This would suggest that copine III and LT do not co-localise under these conditions or else that the amount of SV40 large T antigen expressed in the cytoplasm is comparatively very low.

The sub-cellular localisation of copine III was also assessed by crude cellular fractionation. Copine III and large T antigen levels were compared in the nuclear (P1) cytosolic (S100) and microsomal (P100) fractions prepared following the protocol given in Chapter 2. The fractions were reconstituted in equivalent volumes of lysis buffer and analysed by 1DE-western blotting with anti-copine III and anti-SV40 large T antigen antibodies (Figure 5.9C). This experiment showed that the copine III was predominantly localised in the cytosolic fraction (S100), which correlated to the sub-cellular distribution observed by immunofluorescence staining. In addition, low levels of copine III were also detected in P1 and P100 fractions in the C3.6 cells, which could be representative of membrane association. SV40 large T antigen was predominantly found localised in the nuclear fraction (P1) but also detected at significant levels in the cytosolic (S100) and the microsomal (P100) fractions. Thus, copine III and the SV40 large T antigen could possibly interact, suggesting that the association observed by co-immunoprecipitation is real. Although a direct involvement of copine III in the ErbB signalling network was not unveiled by this analysis, important insights into copine III function have been revealed. However, further work is necessary to assess the role of copine III up-regulation in ErbB-2-dependent tumourigenesis.

5.5 Conclusions and Discussion

Proteomic analysis enables the large-scale study of a specific biological process and generates a complex unbiased dataset compared to conventional methods. However, further validation of the data and characterisation of targets remains an ambitious task and both time and expertise are essential in the understanding of these complex datasets. In this chapter, the validation of differential protein expression and further characterisation of ErbB-2-dependent differentially expressed targets was carried out. Although mass spectrometric analysis provides direct evidence for the identity of proteins, it remained crucial to validate the quantitative changes in protein abundance observed by using other means of analytical study. For instance, although a single protein match was obtained per protein spot, multiple proteins could fall within the same spot coordinate making *bona fide* determination of differential expression difficult. Here, the specific detection of the proteins of interest was performed by 1DE-western blotting with specific antibodies, enabling increased confidence in protein identification. It was found that conventional 1DE-western blotting could be used to validate most changes observed by 2D-DIGE. Together with the data obtained from transcriptional analysis, this confirmed the differential protein expression profile of PGP9.5, copine III, MxA, MnSOD and CPS-1 in response to ErbB-2 overexpression. However, changes in protein abundance below a 1.5-fold ratio were not reproducible by 1DE western blotting and therefore were not validated as relevant changes (Ku autoantigen, 14-3-3β). Despite this, it could be argued that some of the changes observed on a 2DE gel could be due to post-translational modifications which are not resolved on a 1DE gel. Although this approach of validation successfully enabled the confirmation of most changes observed by 2D-DIGE analysis, this process of validation can be time consuming and somewhat limiting, as it is difficult to obtain good antibodies for all proteins identified by mass spectrometry. In addition, although protein expression ratios could also be estimated by densitometry of immunoblots, such quantitation can vary according to the exposure time and the specificity of the antibody used. Although enhanced chemiluminescence is a sensitive detection method it does not provide a wide dynamic range of detection necessary for accurate protein quantitation, and signal saturation is easily reached.

Finding novel expression changes and protein identities is very important in the characterisation of a cellular process; however, they only become meaningful when their role within a cellular context is understood. For this purpose, proteomic analyses need to be combined with other tools of study in order to generate a complete picture of a biological process. Relating the changes observed in the 2D-DIGE experiment to the overexpression of ErbB-2 in other cell systems was the first step of this functional analysis. In this Chapter, it was shown that the differences in protein expression observed were not solely observed in the comparison between HB4a and C3.6 cells, but could also be detected in other clones derived from the HB4a cells, which overexpress different levels of ErbB-2. This helped to confirm that the changes in protein abundance were not attributable to changes introduced during clonal selection of the cell lines, but related to ErbB-2 overexpression. Similarly, the levels were also checked in a panel of breast cancer cell lines, revealing further information concerning the eventual role of the potential targets in proliferation, morphology and ErbB receptor signalling. In these cell lines, the expression profiles did not perfectly mirror the relative ErbB-2 levels, although some seemed to be correlated to the expression profile of the other ErbB family members (e.g. copine III). This analysis also put forward possible tissue-specific targets such as PGP9.5, which was specifically detected in the luminal epithelial cell lines. Some of the targets tested were responsive to growth factor treatment (MnSOD, MxA and copine III) and their differential protein expression level could be further induced in the parental cell line by growth factor treatment, thus confirming the role of ErbB receptor signalling on their direct regulation. However, the activation of a protein by growth factor treatment cannot always be detected using conventional biochemical approaches or 2DE technique. Hence although a protein could be expressed at similar levels, its activity could be regulated by protein interaction, post-translational modification and differential sub-cellular localisation. It is therefore important in the study of unknown protein targets to further consider the function of a protein and to investigate its protein interaction profile and sub-cellular localisation as well as its levels in expression.

Thus, the next step in this study was the functional proteomic analysis of selected targets. Expression of the Mx family of proteins was further assessed together with a

study of MnSOD. The effect of ErbB-2 overexpression on their expression profile suggested that the down-regulation of MxA and MnSOD in the C3.6 cells was related to enhanced ErbB signalling. Indeed, treatment of the cells with either EGF or Hrgβ1 had a further negative regulatory effect on the regulation of the Mx proteins and on MnSOD. In addition, the expression levels of MxA could be restored in the ErbB-2 overexpressing cell line by treatment with type I IFNs (α/β), and MnSOD expression was induced by both type I and type II IFNs (IFNα/β and IFNγ). Taken together, this data suggested that the levels of MxA and MnSOD observed could be a consequence of a basal suppression of major interferon regulated genes in response to ErbB-2 overexpression. This correlated with the down-regulation of ISGF3γ, a transcription factor involved in the gene expression of MxA and MnSOD, amongst others (Melen *et al.* 2000). There was a significant antagonist effect of growth factor treatments on IFN-stimulated gene expression, thus suggesting cross-talk between growth factor and IFN signalling pathways (Bertani M. and Timms J.F., unpublished data). The cytokine signalling and more specifically the IFN signalling networks are part of a mechanism of protection employed by the cells to respond to viral infections. This in turns induces a blockage in cellular proliferation and ultimately the induction of apoptosis. Therefore the down-regulation of the IFN signalling and IFN-inducible genes in the ErbB-2 overexpressing cells could be a means by which these cells promote proliferation.

Another protein of interest, copine III, was chosen as a potential target for further analysis as little is known about this protein. Copine III was found to be up-regulated in the C3.6 cells compared to their parental line and its expression was found to be regulated by growth factor treatment, suggesting that this protein could be involved downstream of the ErbB signalling network. In addition, copine III was post-translationally modified by phosphorylation and had an associated kinase activity, supportive of previous data (Caudell *et al.* 2000) and the level of phosphorylation and associated kinase activity directly correlated with overall protein expression levels. Analysis of binding partners interacting with copine III revealed its potential interaction with the SV40 large T antigen and BAP-135. This thus suggests a potential role downstream of the Bruton tyrosine kinase (Btk)-family of proteins or in SV40 large T antigen transforming/immortalisation

activity. Further investigation needs to be carried out in order to characterise the functional significance of these findings. Recent work has revealed new insight in the understanding of copine III and showed that this protein could be an important target involved in calcium-dependent intracellular signalling pathways (Tomsig *et al.* 2003). A yeast-two hybrid screen enabled the identification of potential copine III-interacting partners of which MEK1, the protein phosphatase 5 and the CDC42 regulated kinase, are proteins involved in intracellular signalling pathways. This data provides new insight in the understanding of copine III and should be further investigated.

Taken together, this work enabled the validation of most proteins found to be differentially regulated by 2D-DIGE analysis. Additionally, this work showed the high potential of 2D-DIGE in combination with other analytical and biochemical tools for functional protein analysis. The characterisation of potential targets of importance in the study of ErbB-2 overexpression was achieved and revealed a role of a newly described family of proteins, the copine family in the downstream signalling of the receptor tyrosine kinases and also the importance of the down-regulation of IFN-inducible genes in ErbB-2-depedent cellular transformation. In this Chapter, the potential role of ErbB-2 on the differential regulation of the other ErbB family members was also revealed. As previously stated, the ErbB expression profile of a cell type can potentiate its signalling ability and ErbB-2 overexpression seems to play a regulatory role on the other ErbB family members. In addition, a strong correlation between a ubiquitin hydrolase, PGP9.5, and expression levels of ErbB-2 and the ErbB family members was observed. Therefore, further investigation of the role of ErbB-2 on the regulation of PGP9.5 and other ErbB family members will be investigated in the next Chapter.

Chapter 6 : ErbB-2-DEPENDENT EGFR PROTEIN TURN-OVER AND REGULATION OF THE UBIQUITIN/PROTEASOME PATHWAY

6.1 Introduction

The ability of the ErbB receptor family members to be activated by selective heterodimerisation dictate the intracellular sorting of each ErbB receptor (Wang *et al.* 1999; Waterman and Yarden 2001). Although the mechanisms involved in the activation of receptor tyrosine kinases have been extensively studied, the down-regulation of activated molecules is not well characterised. These mechanisms are crucial for the maintenance of the equilibrium within the cell as the quality of a signal also depends on its duration. The down-regulation of receptor tyrosine kinases is achieved in three main steps. These are the dissociation of the ligand, dephosphorylation of activated receptors and their subsequent degradation (Wiley *et al.* 1991; Sorkin and Waters 1993) (Figure 6.1). The rapid internalisation of activated receptors is dependent upon the recruitment of clathrin-coated pits, adaptor molecules and dynamin. Dephosphorylation occurs by activation of intracellular phosphatases, which may also be recruited to the activated receptor, although the defined mechanisms of regulation of dephosphorylation events is poorly characterised. Two main pathways of degradation are known, which involve the proteasome or the lysosome. Generally, the proteasome has been associated with the degradation of polyubiquitinated cytosolic proteins whereas the endosomal-lysosomal system seems mostly to proteolyse complexes associated with receptor trafficking from the plasma membrane, including phagocytosis. Importantly, the ubiquitin-proteasome pathway has also been shown to act as a specific degradation route for key components of the cell cycle, signalling machinery and apoptosis.

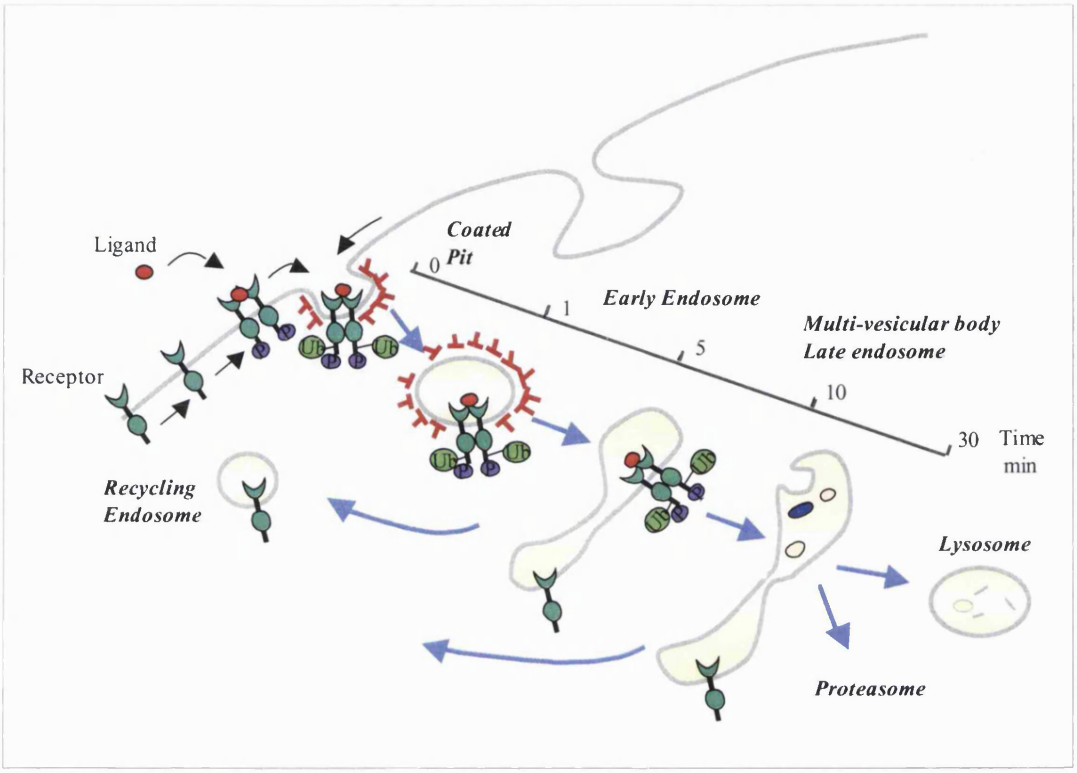


Figure 6.1: Model of endocytosis of receptor tyrosine kinases based on EGFR receptor internalisation and down-regulation (from Waterman and Yarden 2001). Activated receptors trigger the recruitment of effector proteins and adaptor proteins such as dynamin, clathrin coated pits and the AP2 complex. The receptors are then sorted in early endosomes and recycled to the plasma membrane or directed for degradation via the formation of multivesicular bodies (MVB) and into the lysosome or proteasomal degradation pathways.

Ubiquitination of proteins is mainly thought of as a means of adding a targeting signal to send proteins to be degraded and this is supported by the fact that EGFR ubiquitination occurs at the plasma membrane prior to receptor internalisation (Hicke 1997). The division between routes of degradation is however unclear, and for instance EGFR-polyubiquitination has been proposed to induce degradation by proteasomal-mediated targeting to the lysosomal degradative route rather than by proteasome degradation *per se* (Longva *et al.* 2002). c-Cbl, a ubiquitin E3 ligase, was found to be implicated in the ubiquitination of activated EGFR, providing the necessary targeting signals for degradation (Hicke 1997; Levkowitz *et al.* 1998). This mechanism is thought

to be dependent upon receptor tyrosine kinase activation and on the phosphorylation of EGFR at tyrosine 1045, which acts as a docking site for the recruitment of c-Cbl (Levkowitz *et al.* 1996; Levkowitz *et al.* 1999; Longva *et al.* 2002). However, this mechanism of down-regulation is not shared by other ErbB family members, and ErbB-2, ErbB-3 and ErbB-4 are thought to display impaired ligand-mediated receptor endocytosis and degradation compared to EGFR (Sorkin *et al.* 1993; Baulida *et al.* 1996). The ability of ErbB-2 to be ubiquitinated and further degraded is still controversial, and it has been reported that ErbB-2 fails to recruit c-Cbl efficiently (Graus-Porta *et al.* 1997; Lenferink *et al.* 1998). However, since ErbB-2 is the preferred heterodimerisation partner of the other ErbB family members (Graus-Porta *et al.* 1997), it could also play an important role in the modulation of the rate of degradation of other ErbB family members, for instance by favouring receptor recycling (Lenferink *et al.* 1998) (see Figure 6.1). This may explain the potent oncogenic role of ErbB-2, which could act by stabilising ErbB-2/EGFR heterodimers at the cell surface to prolong intracellular signalling (Graus-Porta *et al.* 1997; Tzahar *et al.* 1997; Harari *et al.* 1999).

Variable levels of ErbB receptors were found in the HB4a and C3.6 cells and in particular, EGFR was found to be down-regulated in the C3.6 cells (Chapter 4). These differences in expression were not found at the transcriptional level in the microarray experiment, thus suggesting that ectopically expressed ErbB-2 could negatively regulate EGFR expression by a decrease in the half-life of EGFR protein in C3.6 cells. This hypothesis goes against previously reported data from other cellular models, showing that overexpression of ErbB-2 has a negative effect on the rate of endocytosis and degradation of EGFR (Lenferink *et al.* 1998; Wang *et al.* 1999; Worthylake *et al.* 1999).

Interestingly, the protein found to be the most differentially regulated in the 2D-DIGE analysis, was ubiquitin C-terminal hydrolase UCH-L1/PGP9.5. As previously mentioned, PGP9.5 is a ubiquitin hydrolase which is likely to play a role in the regulation of protein turn-over by ubiquitination. Therefore, a ubiquitin-dependent increase of EGFR internalisation caused by ErbB-2 overexpression in the model cell system studied is a novel and intriguing suggestion that would ultimately explain contradictory results between different cell lines. It was therefore important to confirm these changes and to

understand the possible mechanisms of regulation induced by ErbB-2 overexpression and their effects on cell growth, proliferation and morphology. In this Chapter, differences in EGF-mediated EGFR internalisation and degradation in the HB4a and C3.6 cells were investigated, as well as the roles of the ubiquitin/proteasome pathway. Finally, PGP9.5 was characterised as a potential mediator of receptor down-regulation.

6.2 EGFR receptor turn-over mediated by ErbB-2 overexpression

6.2.1 EGFR, but not ErbB-2 is internalised upon EGF binding in HB4a and C3.6 cells

In order to better understand the differential protein expression of EGFR between the HB4a and C3.6 cells, the distribution of EGFR was assessed by immunofluorescence microscopy. The cells were grown to subconfluency, serum-starved for 24 hrs, and treated for 10 min with 1 nM of EGF. Cells were fixed with 4% paraformaldehyde, permeabilised and stained with specific antibodies to ErbB-2 and EGFR (Figures 6.2). Due to its low expression level in the parental cell line, ErbB-2 could not be detected in these cells. In the C3.6 cells, ErbB-2 was highly expressed and appeared to be localised in microvilli which contain microfilaments associated with bundling proteins at the apical surface of the cells (Figure 6.2A). A relatively high level of EGFR was expressed in both cell lines, correlating with western blotting experiments (Figure 4.1A, Chapter 4) and previous reports showing that EGFR is highly expressed in human mammary epithelial cells (Valverius *et al.* 1989; Burke and Wiley 1999). EGFR was mainly distributed at the plasma membrane and was more condensed at cellular junctions, although a pool of receptors seemed to display perinuclear localisation in the HB4a cells (Figure 6.2B). After 10 min of EGF treatment, ErbB-2 appeared to be mostly localised at the plasma membrane and in particular at the cell-to-cell junctions.

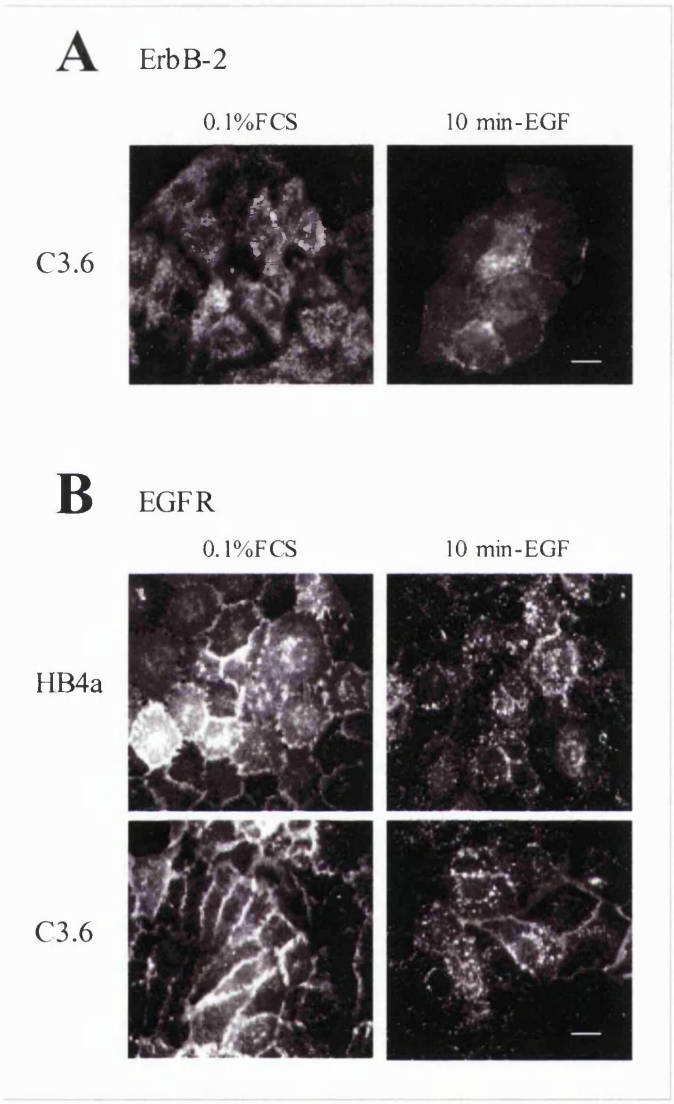


Figure 6.2: Immunostaining of EGFR and ErbB-2 in HB4a and C3.6 cells. The cellular distribution of ErbB-2 and EGFR was investigated in both cell lines under serum-starved conditions and after 10 min of treatment with 1 nM EGF. Fixed cells were stained with anti-ErbB-2 and anti-EGFR monoclonal antibodies and subsequently with FITC-conjugated mouse secondary antibody. Images were captured using a CCD camera mounted over a Zeiss microscope (Chapter 2). Representative images from 6-fields are shown and all images were captured at the same magnification (bar, 50 μm). ErbB-2 was not detected in the HB4a cells due to its low expression level.

No evidence of EGF-dependent ErbB-2 internalisation could be detected (Figure 6.2A). In contrast, a rapid internalisation of EGFR was observed in both cell lines after 10 min of treatment with EGF, and receptors were detected in internal vesicles, probably early endosomes (Figure 6.1), in agreement with previous reports (Lenferink *et al.* 1998; Wang *et al.* 1999; Salazar and Gonzalez 2002).

Therefore this would suggest first that ErbB-2 overexpression does not inhibit EGFR internalisation in the C3.6 cells, and second that EGFR is not internalised in its ErbB-2-coupled heterodimeric state.

6.2.2 Analysis of EGFR degradation and the role of the proteasome

To further assess ErbB-2-dependent changes in EGFR trafficking in these cell lines, and to understand the differential down-regulation of EGFR, the rate of EGF-dependent receptor internalisation and degradation was investigated. The level of EGF receptor localised at the cell surface was monitored by biotinylation with Sulfo-NHS-Biotin, as described previously (Le Bivic *et al.* 1989). Briefly, HB4a and C3.6 cells were treated for 0 hr, 10 min, 30 min, 1 hr and 4 hrs with 1 nM EGF and then incubated for 15 min with NHS-Biotin on ice for 15 min to label EGFR remaining at the surface after ligand-induced internalisation (Figure 6.3). Labelling was performed on ice in order to prevent further internalisation. The fraction of biotin-labelled proteins was then enriched with streptavidin agarose and levels of EGFR and ErbB-2 were assessed by immunoblotting (Figure 6.4 and data not shown). The levels of ErbB-2 expressed at the cell surface were investigated and showed that ErbB-2 expression and subcellular distribution remained unaltered upon treatment with EGF. ErbB-2 was not detected by western blotting in HB4a cells, due to its low expression level (data not shown).

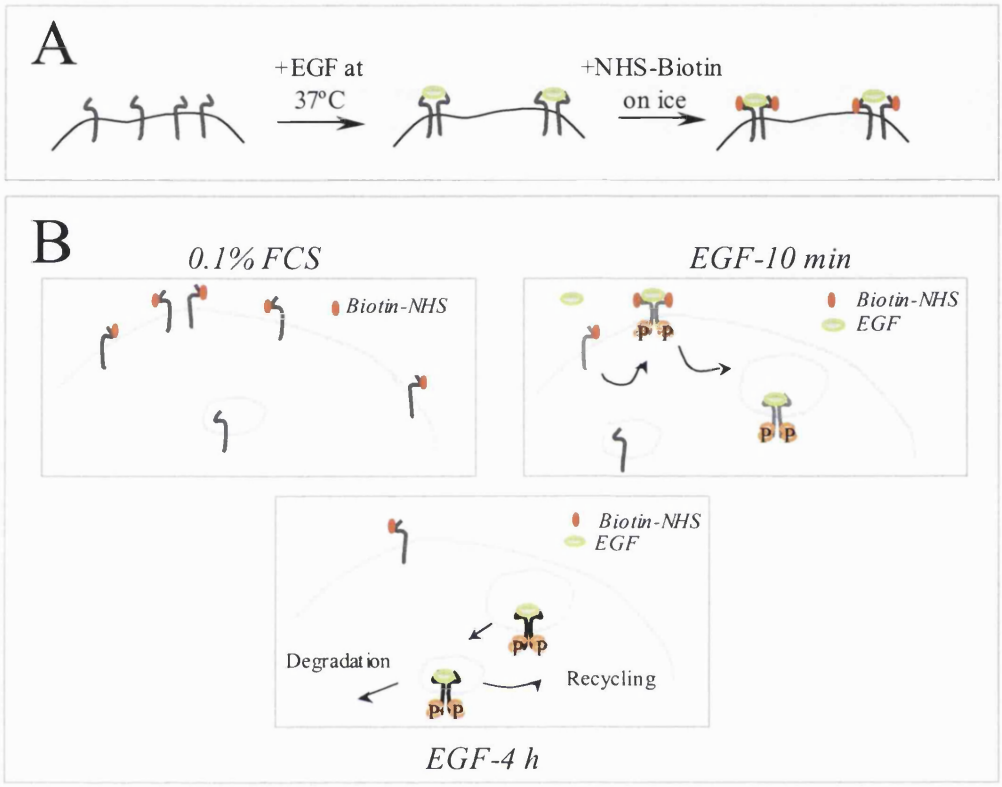


Figure 6.3: Investigation of the level of EGFR at the cell surface and receptor turnover by cell surface labelling with NHS-Biotin. A) Experimental design in the study of EGFR internalisation and down-regulation in C3.6 and HB4a cells. At each time point of treatment with EGF, the cells were washed and labelled with 500 ng/ml of NHS-Biotin on ice for 15 min. The reaction was then quenched with excess BSA and labelled receptors were enriched with streptavidin agarose beads. The enriched fraction was then analysed by western blotting with specific antibodies for EGFR and ErbB-2. B) Expected labelling pattern at each time point of stimulation with EGF.

Under serum-starved conditions, the total EGFR expression level was higher in HB4a cells compared to C3.6, as previously observed. However, analysis of several experiments revealed that the receptor expression level at the cell surface at time 0 was similar between HB4a and C3.6 cells. This result, is in agreement with the previous immunolocalisation experiment, and suggests the existence of an internal perinuclear pool of receptor in the HB4a cells, whereas in the C3.6 cells, EGFR seems to be mostly confined to the cell surface (Figure 6.2B). Since ErbB-2 has been proposed to induce EGFR recycling to the cell surface (Lenferink *et al.* 1998; Worthylake *et al.* 1999), this observation could also suggest that constitutively internalised EGFR might be recycled faster to the cell surface in the C3.6 cells. However, this alteration of EGFR trafficking would not explain why there is less total receptor. Upon EGF treatment, the levels of EGFR expressed at the cell surface decreased steadily after 10 min of stimulation in both cell lines (Figure 6.4A). Decrease of surface EGFR was faster in the C3.6 cells compared to the parental line, with difference in levels maximal after 30 min of treatment with EGF (40% versus 90% of un-stimulated levels in C3.6 versus HB4a, Figure 6.4B). Interestingly, comparison of levels of surface labelling suggested that EGFR internalisation was delayed for the first hour of growth factor treatment in the HB4a cells.

In order to assess the role of the ubiquitin/proteasome pathway on EGF-induced internalisation and degradation of EGFR, cells were also pre-treated with PS341, a specific inhibitor of the proteasome (Adams and Elliott 2000; Aghajanian *et al.* 2002). The level of expression of total and cell-surface EGFR expressed in the HB4a and C3.6 cells was assessed by western blotting, as before (Figures 6.4A and 6.5A). Pre-treatment of the cells with PS341 diminished degradation of EGFR although it was less potent in the C3.6 cells (Figure 6.5). The fact that EGFR was degraded in the C3.6 cells after treatment with PS341, suggests that these cells can degrade EGFR via another route, possibly via lysosomal degradation. In addition, this difference was also apparent when the cell surface EGFR was examined (Figures 6.4A and 6.5B).

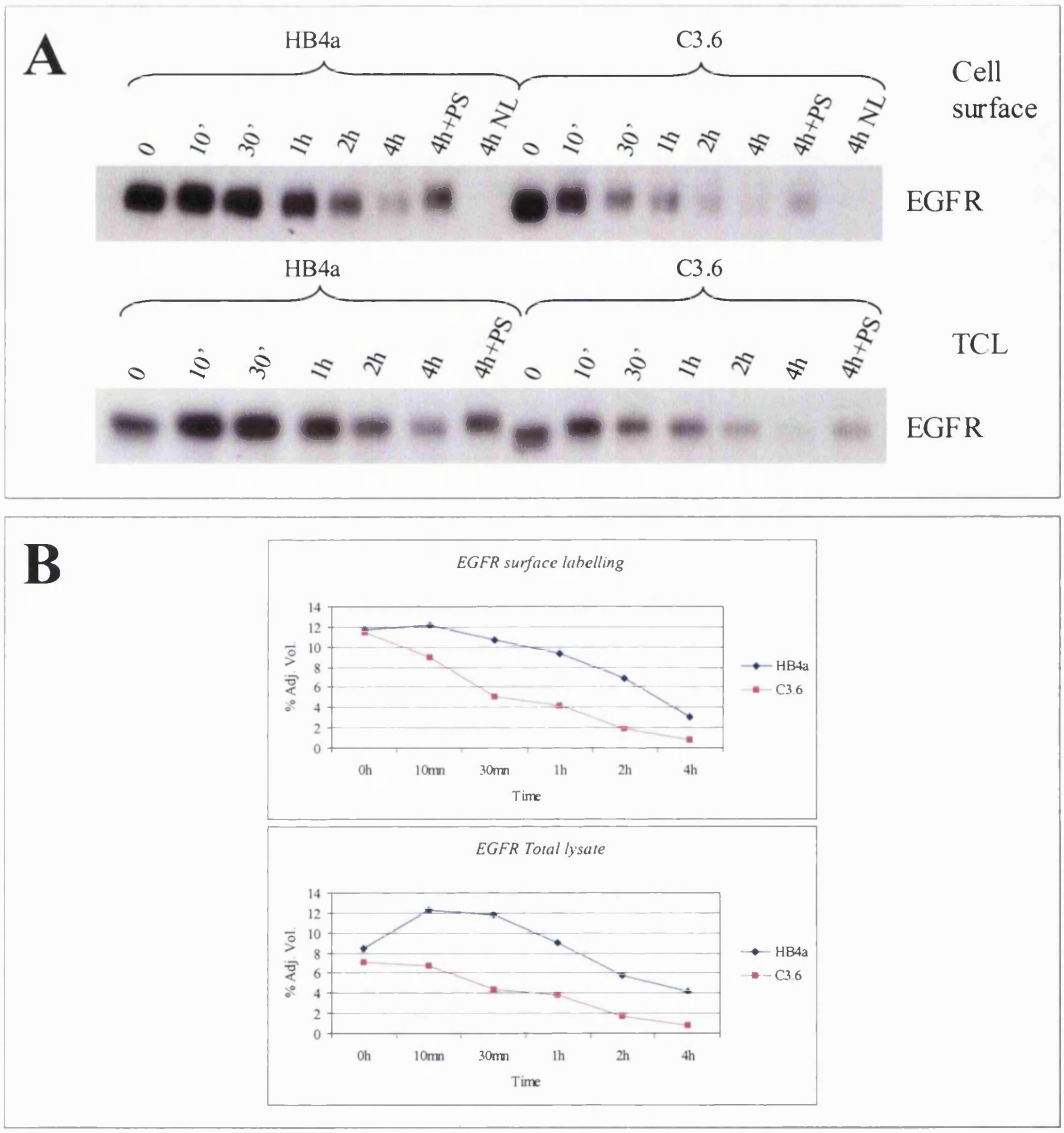


Figure 6.4: Rate of EGFR receptor internalisation in the HB4a and C3.6 cells in response to EGF. Cells were treated with EGF for various periods of time and labelled with NHS-Biotin as described (Figure 3). A) The enriched biotinylated fraction was resolved on a 1DE-SDS-PAGE gel and analysed by western blotting. The levels of EGFR expressed at the cell surface were compared to levels of EGFR internalised, obtained post-biotin enrichment (TCL) by immunoblotting with specific antibodies. NL: non-labelled; PS: PS341 pre-treatment 30 min before EGF stimulation. B) The rate of receptor internalisation was estimated by densitometry of western blots using QuantityOne software (BioRad) and similar results were obtained in three independent experiments. The levels at the cell surface were compared to total levels.

Taken together, this data indicates that EGF treatment triggers EGFR internalisation and degradation in both cell lines, but with increased rates of internalisation and degradation in the ErbB-2 overexpressing cells, which reflects the decreased overall expression level. Notably, the differential effect of PS341 on receptor down-regulation in the cell lines also suggests a distinct role for the ubiquitin/proteasome pathway in the regulation of EGFR receptor turn-over between the HB4a and C3.6 cells.

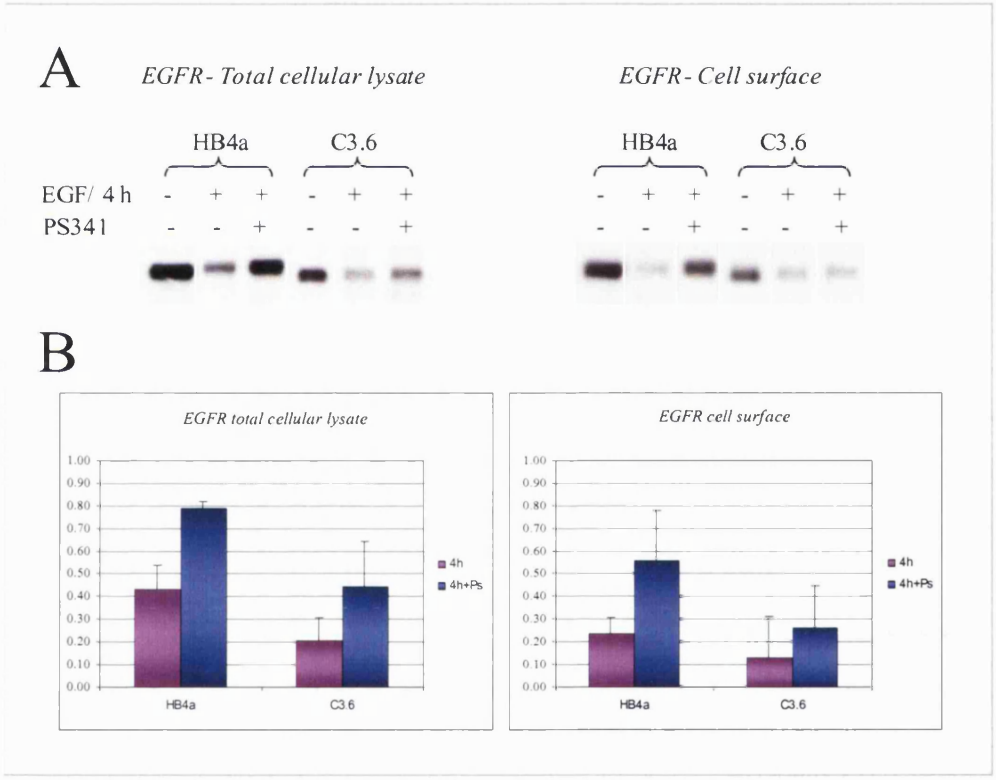


Figure 6.5: Assessment of the rate of EGFR degradation in HB4a and C3.6 cells. Cells were treated for 4 hrs with EGF with or without PS341 pre-treatment. Cells were subsequently labelled with NHS-Biotin and the streptavidin-enriched fraction was analysed by 1DE-western blotting. A) EGFR receptor levels at the cell surface and in the total cellular fraction, representative of a triplicate experiment. B) Ratios of EGFR abundance for each treatment were calculated against the untreated (HB4a-0 hr or C3.6-0 hr) sample following densitometry of immunoblots.

6.2.3 *Mechanisms of EGFR internalisation and degradation*

It has been proposed that EGF-induced autophosphorylation of EGFR induces its stable poly-ubiquitination and directs further receptor sorting to the degradation pathway (Longva *et al.* 2002). As previously reported, this event requires the phosphorylation of EGFR at tyrosine 1045 and the subsequent recruitment of c-Cbl (Levkowitz *et al.* 1999). Therefore to address the role of receptor tyrosine kinase activity on the differential degradation of EGFR in the HMLEC system cells were treated with AG1478, a receptor tyrosine kinase inhibitor specific for both ErbB-2 and EGFR (Egeblad *et al.* 2001). Cell treatment with 5 μM AG1478 results in a blockage in receptor tyrosine phosphorylation and downstream signalling via the MAPK and PI3K/Akt signalling pathways (Chan and Timms, unpublished data). In addition, EGFR protein turn-over was also investigated in the presence of an inhibitor of protein synthesis, cycloheximide (CHX) to control for any alterations in receptor synthesis. A parallel experiment was then carried out to examine the expression level of EGFR (assessed by immunoblotting) and its sub-cellular localisation (assessed by immunofluorescence) in the absence and presence of AG1478, CHX and PS341 to block kinase activity, *de novo* synthesis and degradation, respectively. Total cellular levels of EGFR were analysed by 1DE-western blotting and compared to levels of c-Cbl (Figure 6.6). In this experiment, EGFR levels were more similar in the HB4a and C3.6 cells, which is not representative of previous experiments. However, this could be due to different stages of confluency compared to previous experiment. After 4 hrs of stimulation with EGF, EGFR degradation was more advanced in the C3.6 cells, compared to the HB4a cells, and again, PS341 blocked receptor degradation more efficiently in HB4a compared to C3.6 cells. Pre-treatment with AG1478 appeared to abrogate receptor tyrosine phosphorylation in both cell lines, as shown by blockage of the molecular weight shift induced by EGF. More importantly, kinase inhibition almost completely blocked EGF-induced degradation in the HB4a cells but surprisingly, degradation was still apparent in the C3.6 cells. Thus, there appears to be a proteasome and kinase independent mechanism of EGFR degradation that is enhanced by ErbB-2 overexpression. Finally, pre-treatment of the cells with CHX showed an inhibition of EGFR degradation in HB4a cells, but did not significantly affect

the level of degradation in C3.6 cells. This data suggests that the mechanism of degradation of EGFR requires protein synthesis and tyrosine phosphorylation in HB4a but not in C3.6 cells. Interestingly, c-Cbl was up-regulated in the C3.6 cells and its expression increased when cells were pre-treated with PS341 and to a lesser extent with AG1478. As previously reported, ErbB-2 does not efficiently recruit c-Cbl, However, the increased level of ErbB-2 could account for higher phosphorylation of EGFR at the tyrosine 1045 which would promote receptor internalisation and ubiquitin-dependent degradation in the C3.6 cells through increased c-Cbl recruitment.

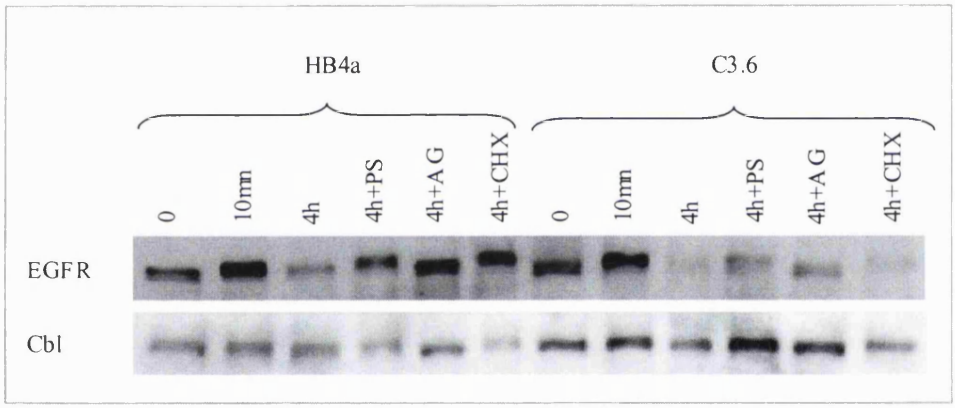


Figure 6.6: Investigation of the regulation of EGFR degradation upon treatment with inhibitors of the proteasome, ErbB kinase activity and protein synthesis. HB4a and C3.6 cells were serum-starved and treated with EGF for 10 min and 4 hrs or pre-treated with 1 μM PS341, 5 μM AG1478 or 10 μg/ml CHX for 30 min prior to EGF treatment. Total cellular extracts were resolved by 1DE and levels of EGFR and Cbl were estimated by western blotting.

Receptor down-regulation and degradation was further investigated by immunofluorescence analysis. Cells were treated using the same conditions as described above and fixed cells were subsequently immunostained for EGFR. As before, EGF-mediated receptor internalisation and degradation could be observed in both cell lines and was indistinguishable (Figure 6.1). Notably, PS341 reduced receptor internalisation and degradation and the effect was greater in the parental cell line.

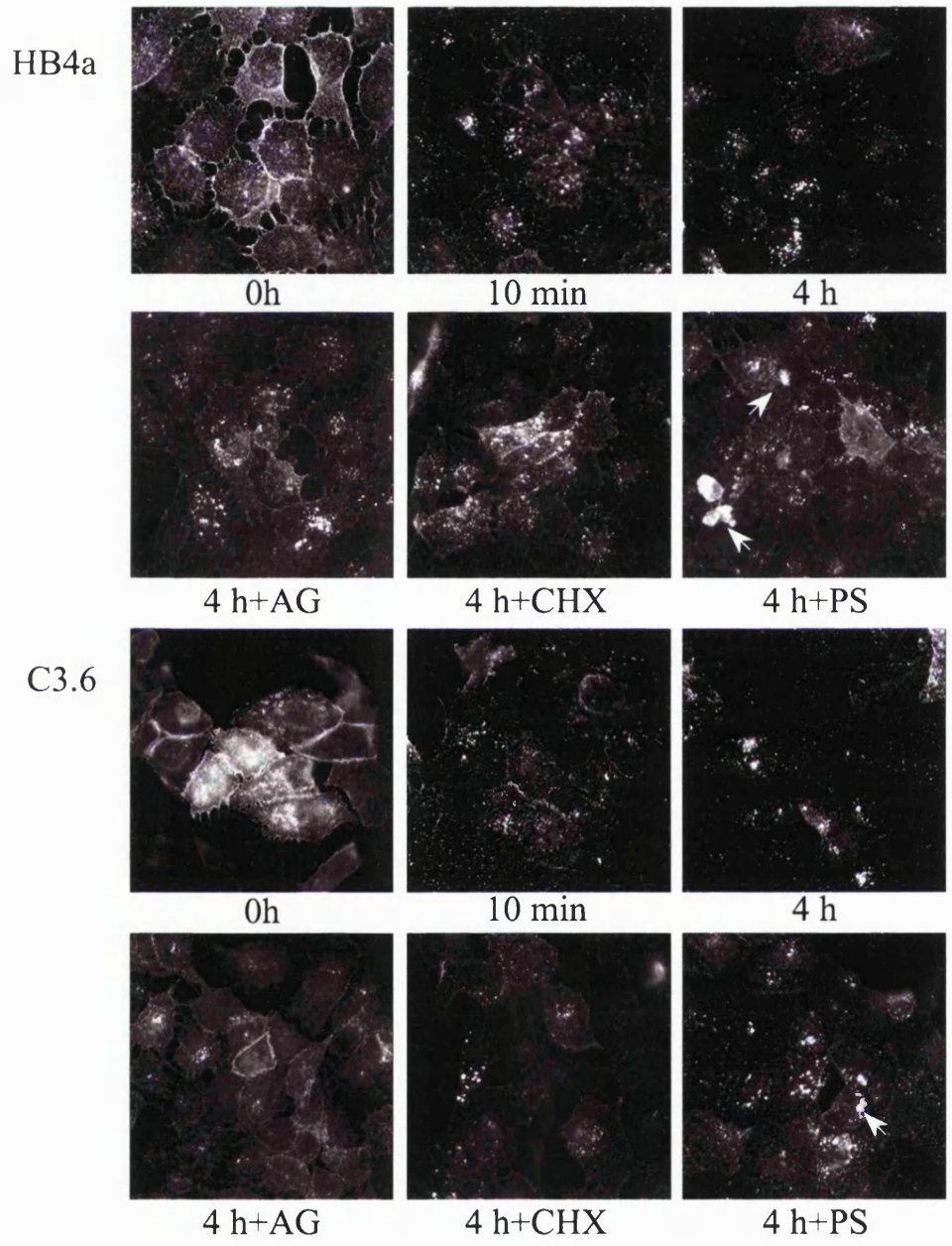


Figure 6.7: Effect of PS341, CHX and AG1478 on EGF-dependent receptor internalisation and degradation. Cells were treated following the same conditions as previously described (Figure 6.6) and fixed cells were immunostained for EGFR using a mouse monoclonal antibody specific for EGFR. Staining was revealed using a mouse secondary antibody conjugated with TRITC and imaging was performed using a CCD-camera mounted over an Axiophot microscope with a x63 immersion oil objective. EGFR was detected in large intracellular vesicles in PS341-treated cells, as shown by the white arrows.

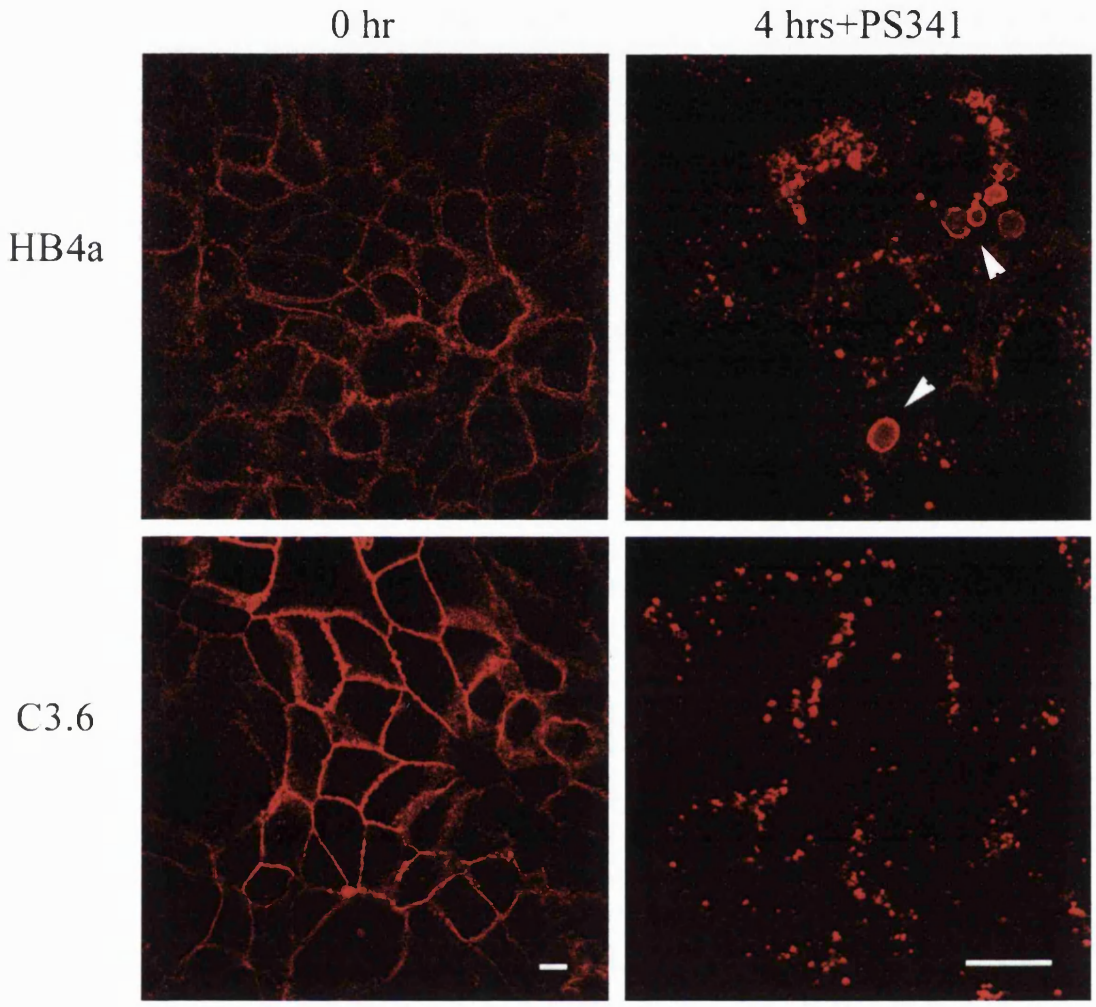


Figure 6.8: Effect of PS341 treatment on EGF-dependent EGFR receptor internalisation and degradation. Cells were prepared as previously described (Figure 6.7). Fixed cells were permeabilised and immunostained with anti-EGFR specific antibody and with TRITC-conjugated anti-mouse secondary antibody. Images were captured using a confocal microscope (details in Chapter 2). Receptors could be detected in sub-cellular compartments similar to multivesicular bodies (MVBs) as shown by the arrows (bar, 50μm).

In addition, internalised receptors seemed to accumulate in the cells in presence of PS341 (Figure 6.7 "white arrows") and this effect was better observed by confocal microscopy (Figure 6.8). Internalised receptors appeared to accumulate in sub-cellular compartments which resembled multivesicular bodies (MVBs), which are secondary lysosomal compartments containing small vesicles. This vesicular accumulation of EGFR was not so prominent in the ErbB-2 overexpressing cells. This data therefore supports the results observed by western blotting analysis and agrees with the hypothesis that ErbB-2 overexpression plays a role in regulating EGFR receptor down-regulation by directing EGFR through a degradation route different to the parental cell line, probably via the lysosomal degradation route.

EGF receptor distribution was also analysed after pre-treatment of cells with AG1478 or CHX (Figure 6.7). Generally, the staining intensity of EGFR reflected levels observed by western blotting. Indeed, treatment of the cells with CHX seemed to reduce the rate of receptor degradation upon treatment with EGF in the parental cells and no effect was observed in the C3.6 cells. This would suggest that the mechanism of receptor down-regulation is independent on newly synthesised proteins in the ErbB-2 overexpressing cells. Furthermore, EGFR internalisation occurred even when cells were pre-treated with AG1478 in HB4a cells with partial inhibition observed in the C3.6 cells. This indicates that the mechanism of receptor internalisation is partially independent of EGFR tyrosine kinase activation, although receptor degradation was affected by treatment with AG1478.

6.3 Characterisation of PGP9.5 and its role in ErbB-2-mediated transformation

The increased expression level of PGP9.5 observed in the C3.6 cells was further characterised to correlate this level of expression with the potential differential degradation route observed for EGFR in the HB4a and C3.6 cells. PGP9.5 is a deubiquitin enzyme and family member of the ubiquitin carboxy-terminal hydrolases (UCH). Although the process of protein ubiquitination is well studied, little is known about the regulatory function of deubiquitinating proteins (DUBs). There are two known

classes of DUBs, the UCHs and the ubiquitin isopeptidases (ubiquitin-specific processing proteases UBPs). The DUBs are involved in the recycling of ubiquitinated proteins, but they also act as enhancers of the ubiquitin/degradation pathway by regenerating the pool of free ubiquitin (Larsen *et al.* 1998; Swaminathan *et al.* 1999). Three known isoforms of the UCH family have been identified in mammalian cells, UCH-L1 (PGP9.5), UCH-L2 and UCH-L3. PGP9.5 was found widely expressed in neuronal tissues and throughout all stages of development (Day and Thompson 1987; Wilkinson *et al.* 1989). However, it was also found overexpressed in patients with advanced pancreatic cancer (Tezel *et al.* 2000) and non-small lung cancer (Hibi *et al.* 1999), and has been associated with tumour progression in colorectal cancer (Yamazaki *et al.* 2002). Although PGP9.5 was found up-regulated in various cancers, the defined mechanism of this enzyme and its role in cancer progression are not yet understood. Thus, pursuing further the analysis of PGP9.5 up-regulation in response to ErbB-2 overexpression remained an important goal.

As reported in the Chapters 4 and 5, PGP9.5 was the protein found to be the most differentially regulated between the HB4a and C3.6 cell lines with an average 32.5-fold increase in protein abundance in the C3.6 cells compared to the HB4a cells. Further validation in a panel of HMLEC clones expressing different levels of ErbB-2 revealed a close correlation between the levels of expression of ErbB-2 and the levels of PGP9.5 (Figures 6.9A and 5.2). However, the expression of PGP9.5 was specific to luminal epithelial cell lines and was detected in a control luminal cell line, Lum878, with no significant expression detected in other breast tumour cell lines (BT474, MCF7, MB453, SKBr3) (Figure 6.9B and 5.3).

Interestingly, PGP9.5 was shown to be a potential interaction partner of JAB1, a co-activator of c-jun and component of the signalosome (Claret *et al.* 1996; Seeger *et al.* 1998), by yeast-two hybrid experiment (Caballero *et al.* 2002). Further biochemical and biological analysis demonstrated that PGP9.5 could also interact with $p27^{kip}$, a cyclin-dependent kinase inhibitor member of the kinase inhibitor protein (KIP) family and key negative regulator of the cell cycle. PGP9.5 was shown to play a putative role in the trafficking of $p27^{kip}$ from the nucleus to the cytoplasm and degradation together with JAB1 (Caballero *et al.* 2002). Interestingly, ErbB-2 overexpression can potentiate cellular

proliferation by mediation of the degradation of $p27^{kip}$ (Lane *et al.* 2000). $p27^{kip}$ down-regulation is thought to occur via ubiquitin-dependent protein degradation and/or relocalisation to the cytoplasm (Lane HA 2000; Timms *et al.* 2002). Therefore this is potentially interesting, since the function of PGP9.5 has not been previously described and this ubiquitin hydrolyse could play a role in the modulation of cell cycle regulators in ErbB-2 dependent transformation. The level of expression of $p27^{kip}$ was estimated by immunoblotting with specific antibodies in the HB4a and C3.6 cells and also in other HMLECs overexpressing ErbB-2 and a panel of breast cancer cell lines (Figure 6.9).

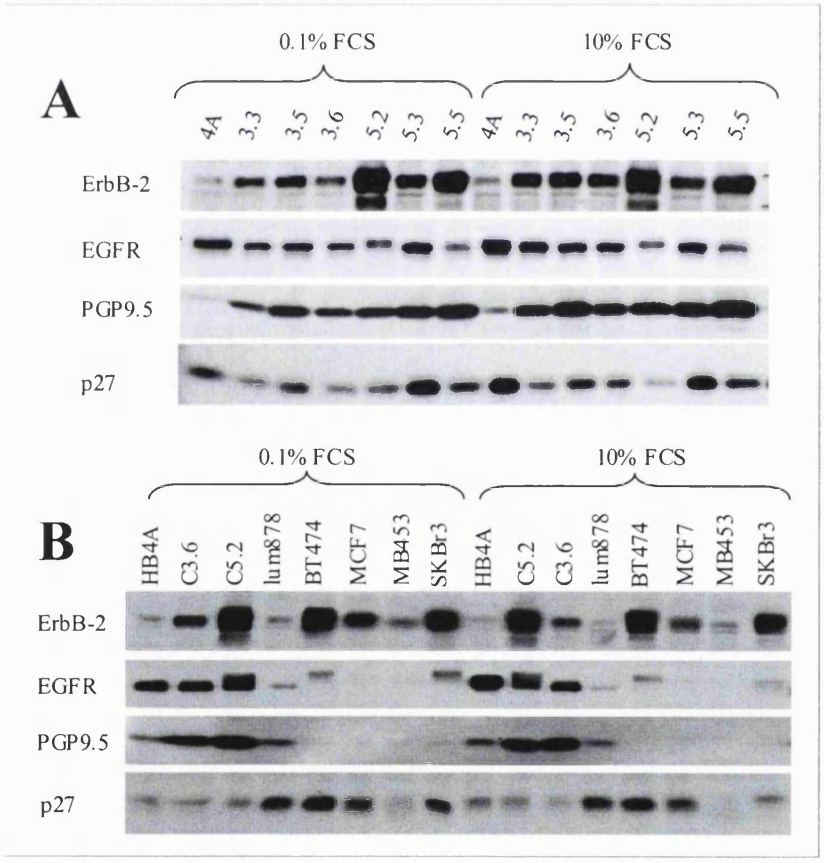


Figure 6.9: PGP9.5 protein expression levels in HMLEC clones and in a panel of breast cancer cell lines. Levels of PGP9.5 were determined by immunoblotting and were correlated to expression levels of other key proteins ($p27^{kip}$, ErbB-2, EGFR). These data were previously shown in Chapter 5 (Figures 5.2 and 5.3), except $p27^{kip}$.

This data revealed that $p27^{kip}$ protein expression was inversely correlated with PGP9.5 and ErbB-2 expression levels. This could be observed in the HMLECs cell lines but was more striking in the comparison of the panel of breast cancer cell lines.

To further examine the biological role of ErbB-2 in the regulation of PGP9.5, a possible protein-protein interaction between these proteins was investigated by co-immunoprecipitation and co-localisation assay. Endogenous ErbB-2 was immunoprecipited and the co-precipitation of PGP9.5 was tested by western blotting and PGP9.5 was not detected (data not shown). In addition, the putative co-localisation of ErbB-2 and PGP9.5 was assessed by double immunostaining. HB4a and C3.6 cells were serum-starved for 24 hrs and treated with EGF for 10 min and 4 hrs to check if this cellular localisation was growth factor treatment dependent (Figure 6.10). Due to the low expression level of ErbB-2 in the parental cell line, this experiment could only be carried out in the ErbB-2 overexpressing cell line, C3.6. PGP9.5 was detected in the cytoplasm and the nuclei of the cells whilst ErbB-2 remained expressed at the cell surface in microvilli (0 hr) and in area of cell-cell contact (10 min). Under the conditions tested, ErbB-2 and PGP9.5 did not appear to co-localise (Figure 6.10).

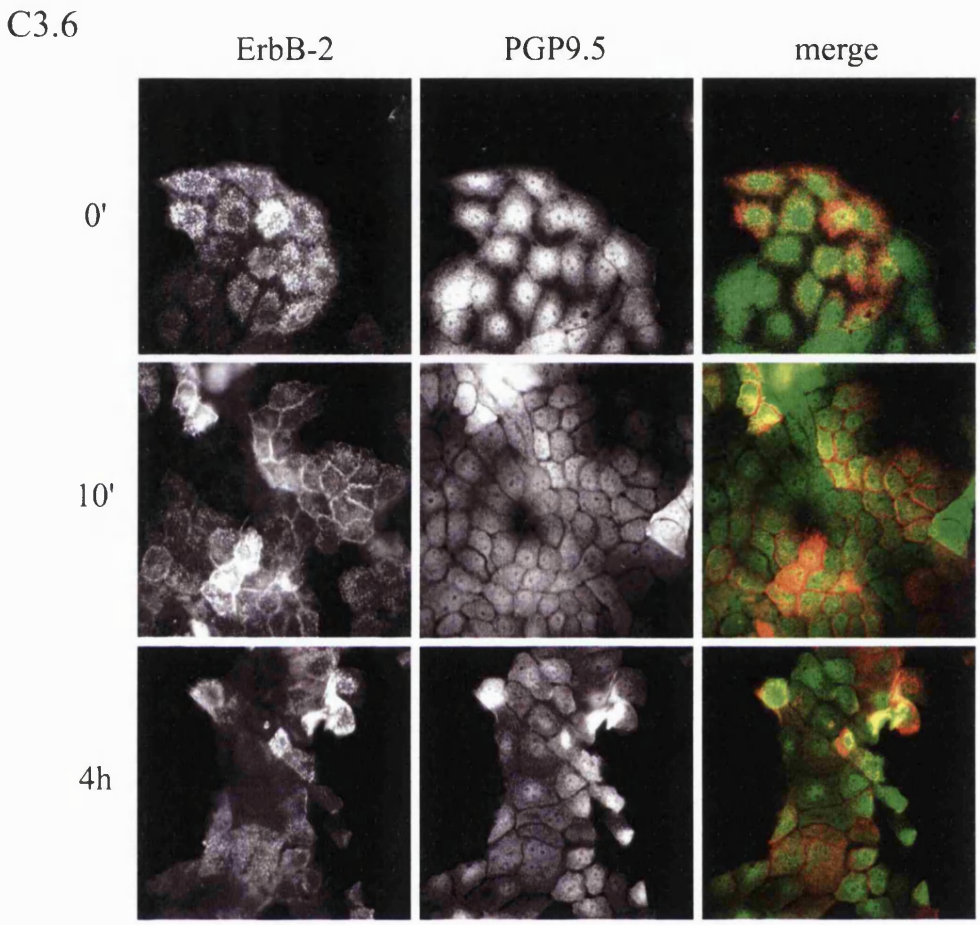


Figure 6.10: ErbB-2 and PGP9.5 sub-cellular localisation investigated by immunofluorescence. Cells were treated with EGF for 0 hr, 10 min or 4 hrs, fixed, permeabilised, and double stained for ErbB-2 and PGP9.5 using a mouse monoclonal anti-ErbB2 and a rabbit polyclonal anti-PGP9.5 antibodies, respectively, and the corresponding TRITC-conjugated anti-mouse and a FITC-conjugated anti-rabbit secondary antibodies. Images were acquired using a CCD-camera mounted over a Zeiss microscope and were overlaid using Adobe Photoshop software.

PGP9.5 and EGFR sub-cellular distribution was also examined by double immunostaining of HB4a and C3.6 cells serum starved or treated for 10 min with 1 nM EGF (Figure 6.11). As previously reported, EGFR undergoes a rapid internalisation after 10 min of treatment with EGF in both cell lines. PGP9.5 was distributed throughout the cell in the cytosol and in the nucleus of the C3.6 cells, however, it was found mostly in the nuclei of the HB4a cells.

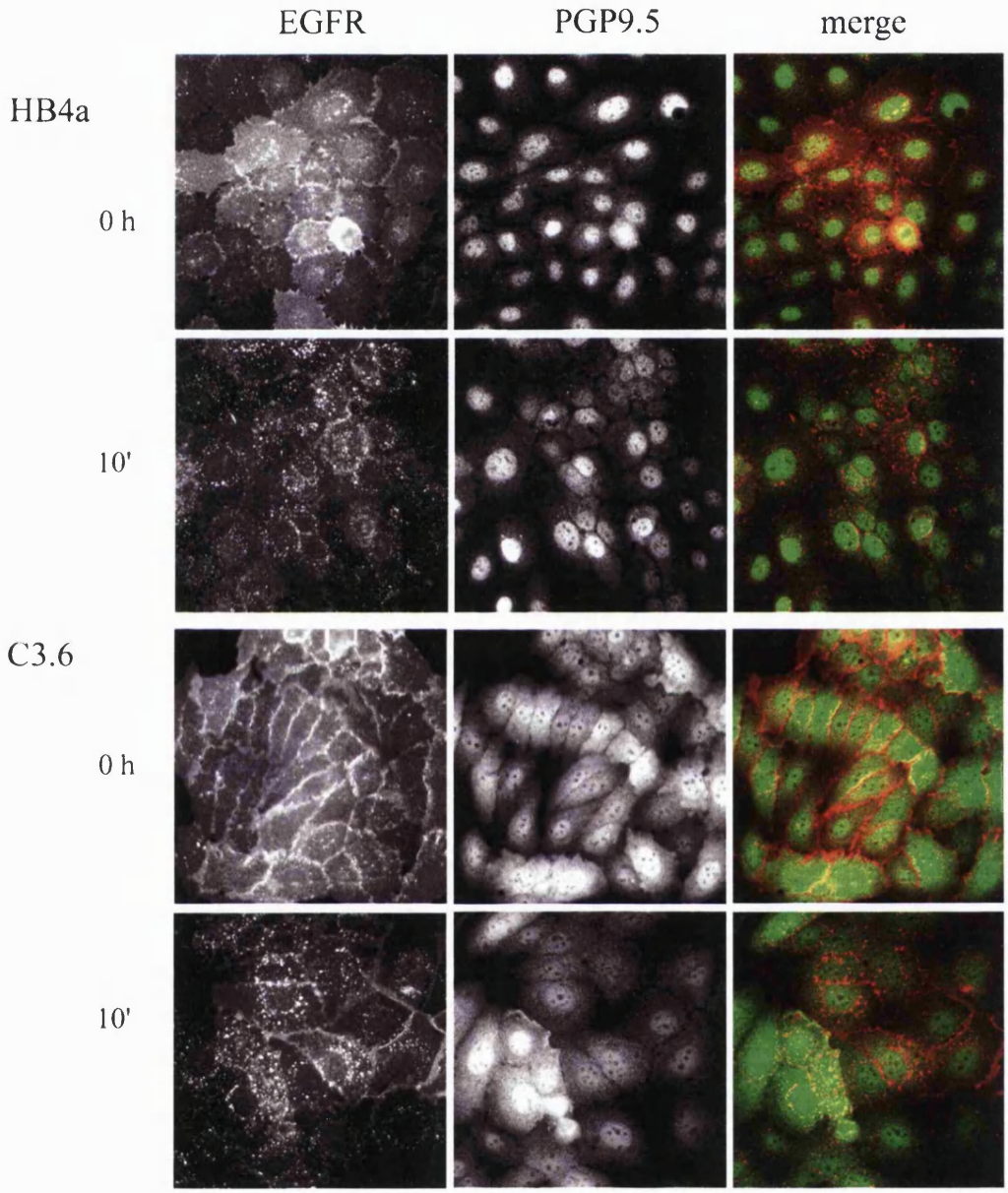


Figure 6.11: PGP9.5 and EGFR sub-cellular distribution in HB4a and C3.6 cells. Cells were serum-starved and treated for 10 min with 1 nM EGF. Fixed cells were permeabilised and double stained for EGFR and PGP9.5 (as described in Figure 6.10). Images were acquired using a CCD-camera mounted over a Zeiss microscope and were overlaid using Adobe Photoshop software.

Merged images were compared and PGP9.5 did not appear to co-localise with EGFR under the growth conditions tested including the vesicles and perinuclear areas (Figure 6.11).

The sub-cellular localisation of PGP9.5 was also assessed by crude cellular fractionation. HB4a and C3.6 cells were serum-starved and stimulated with 1 nM EGF for 4 hrs. Cells were lysed in hypotonic buffer and nuclear (P1), cytosolic (S100) and microsomal (P100) fractions were separated by differential centrifugation steps as described in Chapter 2. Cell equivalents of each fraction were then resolved on a 1DE gel and protein expression levels were assessed by western blotting (Figure 6.12). The levels of EGFR and ERK1/2 were estimated in each fraction in order to validate their purity. Indeed, EGFR was not detected in the cytosolic fraction (S100) although low levels were detected in the membrane fractions (P100) and (P1). Thus the cytosolic fraction is not contaminated with membrane proteins, but some losses are observed in the membrane fraction. ERK1/2 was detected mostly in the cytosolic fraction but also in the nuclear fraction. This could be a specific cellular localisation, as activated ERK1/2 are known to translocate to the nucleus of the cells (Aplin *et al.* 2001). The level of PGP9.5 in each fraction was then examined and compared between the cell lines. PGP9.5 was detected in the cytosol of the ErbB-2 overexpressing cells, but was not detected in the parental cells. In addition, no significant level of PGP9.5 could be detected in the microsomal (P100) or nuclear (P1) fractions. In addition, PGP9.5 remained in the cytosolic fraction after EGF treatment for 4 hrs. This data did not correlate with the cellular localisation found by immunostaining, where PGP9.5 was also detected in the nucleus of the cells. This could be attributed to the insolubility of nuclear proteins under the conditions of cellular extraction used and losses could occur in the extraction of the nuclear fraction. However, the nuclear staining observed could be unspecific although previous literature supports the localisation of PGP9.5 in the nucleus (Caballero *et al.* 2002).

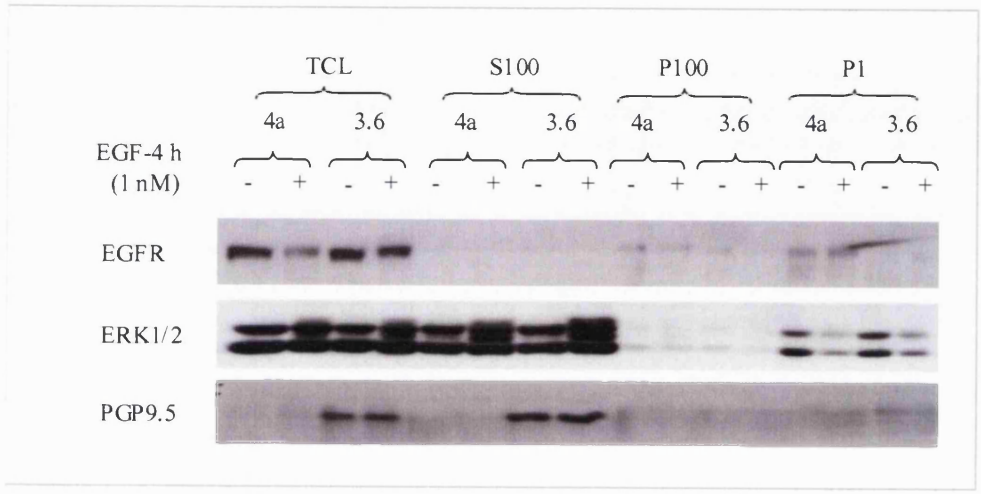


Figure 6.12: Analysis of PGP9.5 sub-cellular localisation upon treatment with EGF by sub-cellular fractionation. Cells were serum-starved and stimulated with 1 nM EGF for 4 hrs. Nuclear (P1), cytosolic (S100) and microsomal (P100) fractions were extracted following the described protocol (Chapter 2). Identical cell equivalents were loaded on a 1DE gel and resolved proteins were analysed by western blotting using anti-EGFR, anti-ERK1/2 and anti-PGP9.5 antibodies. The extra doublets of ERK1/2 are activated phosphorylated forms following stimulation.

The effect of Hrgβ1 and EGF treatment on the protein expression profiles was previously assessed and this showed that neither growth factor had a significant effect on the expression of PGP9.5 (Figure 4.9, Chapter 4). In order to further understand the role of ErbB-2 overexpression in the up-regulation of PGP9.5 in the C3.6 cells, the effect of several specific inhibitors on its expression profile was assessed by immunoblotting and immunofluorescence. Expression levels of PGP9.5 were compared to levels of EGFR and the pool of free ubiquitin was estimated using an anti-ubiquitin antibody (Figure 6.13). PGP9.5 was highly expressed in the C3.6 cells, as previously observed, but treatment with inhibitors AG1478 and CHX had no significant effect on PGP9.5 expression levels. This suggests that ErbB-2 mediated induction of PGP9.5 are not triggered by EGF treatment. However, a slight increase in PGP9.5 expression was detected in C3.6 cells pre-treated with the proteasome inhibitor PS341 but not in the HB4a cells. Interestingly, the C3.6 cells expressed higher levels of free ubiquitin compared to HB4a cells, and

PS341 treatment also increased the expression level of free ubiquitin in the C3.6 cells. From this, it is tempting to speculate that PGP9.5 may play a role in the regeneration of the pool of free ubiquitin.

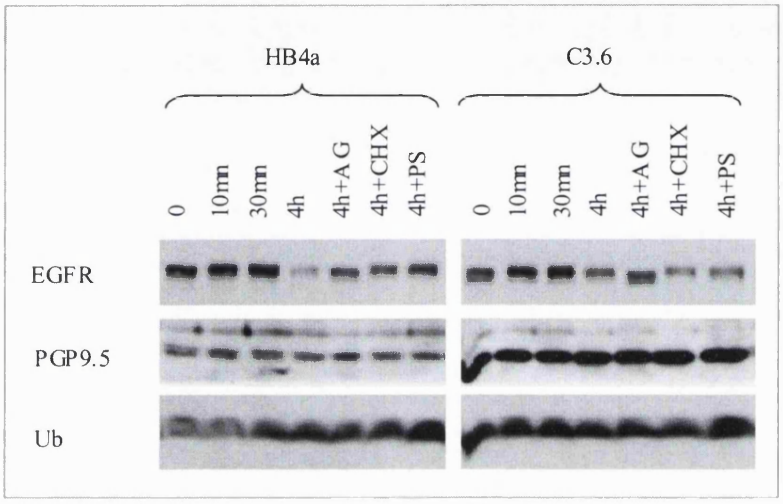


Figure 6.13: Regulation of PGP9.5 expression. HB4a and C3.6 cells were treated for 10 min, 30 min and 4 hrs with 1nM EGF or pre-treated for 30 min with 5 μM AG1478, 10 μg/ml of CHX or 1 μM PS341 prior to 4 hrs treatment with EGF. Total cellular lysates were resolved by 1D-SDS-PAGE and protein levels were analysed by western blotting with EGFR, PGP9.5 and free-Ubiquitin specific antibodies.

PGP9.5 sub-cellular distribution upon co-treatment with EGF and specific inhibitors was investigated by immunostaining (Figure 6.14). PGP9.5 remained expressed in the nucleus of the HB4a cells at low levels although treatment with all three inhibitors increased its expression level. After 4 hrs of treatment with EGF, some dotted PGP9.5-staining was observed in the peri-nuclei of the HB4a cells and to a lesser extent in the C3.6 cells. This phenomenon was also observed in presence of the proteasome inhibitor PS341. A slight nuclear and cytoplasmic staining was observed in C3.6 cells, as previously found, although an increase in the nuclear staining was apparent in AG1478 and CHX-treated cells. This data suggests that the sub-cellular localisation of PGP9.5 may be regulated downstream of ErbB receptors, although its functional significance remains unclear.

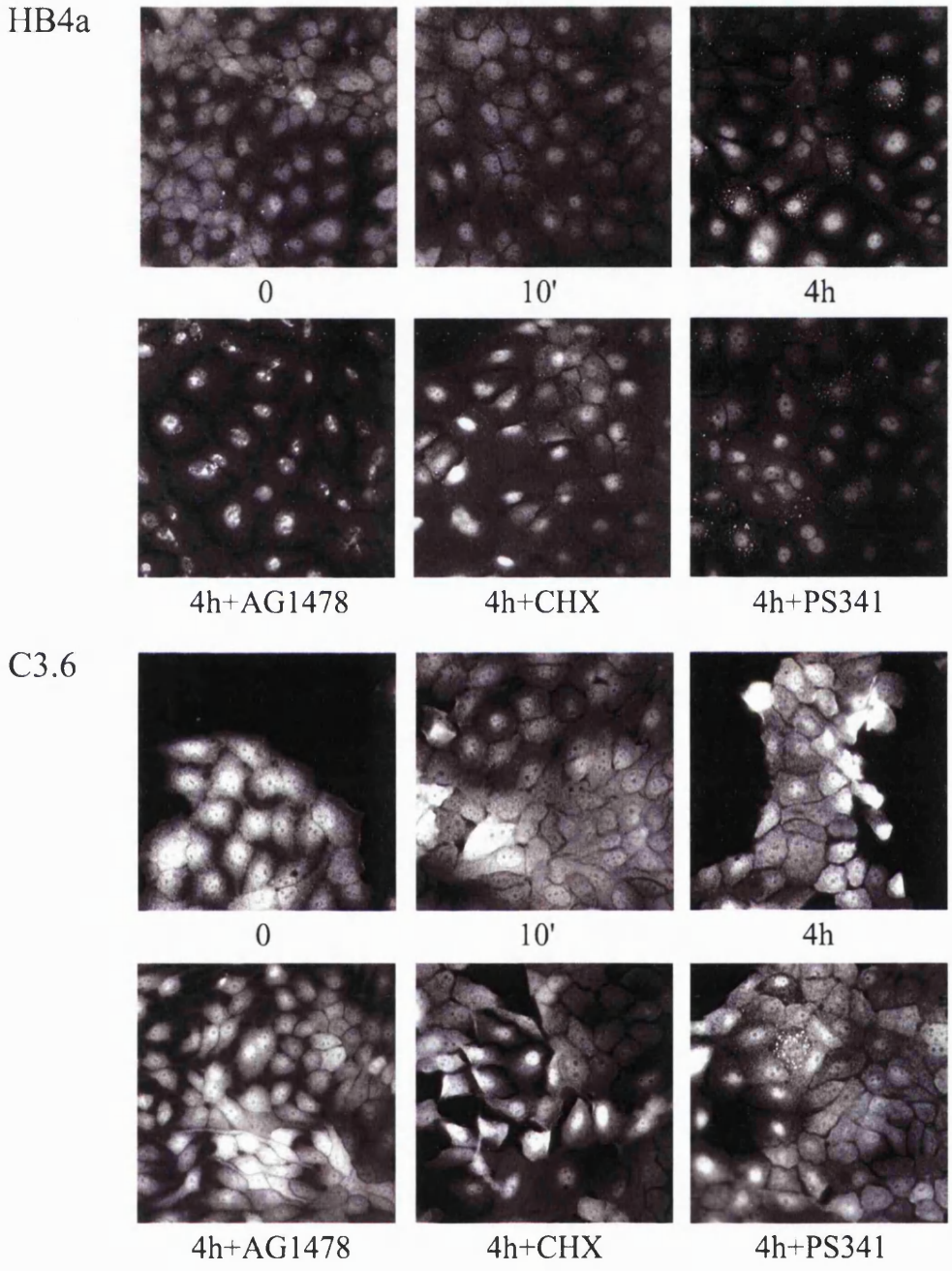


Figure 6.15: PGP9.5 cellular distribution in the HB4a and C3.6 cells following EGF and inhibitor treatment. Cells were grown to sub-confluency and serum starved for 24 hrs and treated with 1 nM EGF for 10 min and 4 hrs with or without prior treatment with AG1478, CHX or PS341. Fixed cells were permeabilised and stained with anti-PGP9.5 specific antibody and FITC-conjugated rabbit secondary antibody. Images were acquired using a CCD-camera mounted over a Zeiss microscope.

On the other hand, the sub-cellular distribution of PGP9.5 was not altered upon treatment with PS341, suggesting that it is not dependent upon proteasomal activity.

Taken together, these data suggest that PGP9.5 may be involved in the regeneration of the pool of free ubiquitin, since increased level of free ubiquitin were detected in C3.6 cells versus HB4a cells. In addition, the differential expression profile observed between HB4a and C3.6 cells could be attributed to a difference in sub-cellular localisation and the function of PGP9.5 could reside in its sub-cellular distribution. Further experiments are required to confirm the dynamic nuclear redistribution of PGP9.5.

6.4 Conclusions and Discussion

By using 2D-DIGE analysis and western blotting, a set of proteins involved in membrane modelling, trafficking and ubiquitin-dependent protein degradation were found to be deregulated by ErbB-2 overexpression. The differential expression level of EGFR observed in the HB4a and C3.6 cells, and the differential protein expression found by 2D-DIGE led to the investigation of the role played by ErbB-2 in EGFR down-regulation. Due to its favoured role as a heterodimerisation partner of EGFR (Graus-Porta *et al.* 1997) and its inability to recruit c-Cbl or other known ubiquitin ligases (Lenferink *et al.* 1998), ErbB-2 is thought to stabilise dimers at the cell surface and therefore prolong cell signalling. In order to further understand the mechanism of ErbB-2-mediated cellular transformation, similarities and differences in the EGFR degradative pathway in the HMLECs cell lines were therefore investigated. Firstly, EGFR was found to be expressed at lower levels in the C3.6 cells. Secondly, the rate of receptor internalisation was assessed between the cell lines and showed that the C3.6 cells had an increased rate of internalisation. However, ErbB-2 was not internalised and remained at the cell surface, suggesting that heterodimerisation is not maintained during internalisation. Furthermore, the ability of the cells to degrade activated receptors was assessed by comparison of total levels of EGFR at different time points of treatment with EGF. According to this experiment, most of the receptor was degraded after 4 hrs of treatment with EGF although the rate of protein degradation was enhanced in C3.6 cells compared to the

HB4a cells. Thus, ErbB-2 overexpression affects EGFR down-regulation by enhancing receptor degradation upon EGF treatment. Accordingly, immunofluorescence analysis confirmed that EGFR was being degraded at 4 hrs, since at that time point most of the receptors were localised in peri-nuclear vesicles resembling multivesicular bodies.

The mechanism of ligand-induced degradation was subsequently assessed using inhibitors of the proteasome, ErbB tyrosine kinase activity and protein synthesis. This data showed that the proteasome inhibitor PS341 significantly reduced the level of EGFR degradation in the HB4a cells, but had little effect in the C3.6 cells. This therefore suggests that ErbB-2 overexpression promotes degradation of EGFR through a proteasome-independent route of receptor degradation, possibly via the lysosome. The role of the receptor tyrosine kinase activity in the mechanism of receptor degradation was also investigated using AG1478 to specifically inhibit EGFR and ErbB-2 tyrosine kinase activity. This treatment suggests that the degradation of EGFR is also dependent to a degree on receptor tyrosine kinase activation although this blockade was less prominent in the C3.6 cells. In addition, AG1478 treatment appeared to only partially block the internalisation of the receptors, as visualised by immunofluorescence. Thus, EGFR receptor kinase activity is necessary but not sufficient for complete ubiquitin-dependent internalisation and degradation, and can also occur in a kinase-independent manner. Furthermore, EGFR degradation in the C3.6 cells seemed to require protein synthesis, as shown by treatment of the cells with cyclohexemide (CHX). Taken together, this data shows that HB4a and C3.6 cells employ differential but overlapping modes of EGFR receptor down-regulation, and the C3.6 cells seemed to be more potent at inducing EGF-dependent receptor degradation. This is likely to have an effect on the duration and specificity of growth factor signalling.

A putative component of the ubiquitin proteasomal pathway, PGP9.5, was found to be highly up-regulated in the C3.6 cells compared to the HB4a cells. This deubiquitin enzyme (DUB) was also found to be up-regulated in several human cancers tested (Hibi *et al.* 1999; Yamazaki *et al.* 2002), however, little is known about the direct implication of the up-regulation of this protein on the development of cancers. Interestingly, the ubiquitin/proteasome pathway has been shown to play an important role in the neoplastic

growth and metastasis, and the proteasome inhibitor, PS341, is already in clinical trials as a cancer therapy (Aghajanian *et al.* 2002). Therefore, further analysis was carried out in order to assess the function of PGP9.5.

The level of expression of PGP9.5 was closely correlated to the level of ErbB-2, and PGP9.5 appears to be a useful maker of ErbB-2 overexpression of luminal epithelial cells. Despite this, there was no evidence of their co-localisation by either immunoprecipitation or immunofluorescence. Some results did however suggest a differential sub-cellular localisation between the HB4a and C3.6 cells lines as PGP9.5 was found to be expressed in the nuclei of both cell lines, but also in the cytoplasm of the C3.6 cells by immunostaining. The sub-cellular localisation of PGP9.5 was also assessed by cell-fractionation. This showed that the majority of the pool of protein was localised in the cytosol and no significant amount of PGP9.5 could be detected in the nuclear fraction of the cells. Additional experiments must be carried out to address this apparent contradiction between both approaches.

Recent reports have shown that PGP9.5 can interact *in vitro* with JAB1 and $p27^{kip}$ (Caballero *et al.* 2002). A comparison of expression levels of PGP9.5 and $p27^{kip}$ in the HLMEC clones and the panel of breast cancer cell lines showed an inverse correlation in their expression profile. Interestingly, $p27^{kip}$ was shown to be down-regulated by a process of poly-ubiquitination, although the defined mechanism of regulation is not well understood. Furthermore, $p27^{kip}$ expression is often down-regulated in ErbB-2 overexpressing cells, as demonstrated here. It has also been shown that PGP9.5 was differentially localised according to the state of growth of the cells and this cellular distribution correlated with an inverse cellular staining of $p27^{kip}$ (Caballero *et al.* 2002). This would therefore suggest that PGP9.5 could regulate $p27^{kip}$ by altering its sub-cellular localisation. Furthermore, this also correlated with previous reports on the cellular distribution of $p27^{kip}$, which was found predominantly expressed in the nucleus of the HB4a cells, but at low levels and in the cytoplasm of the C3.6 cells (Timms *et al.* 2002). In parallel to the biochemical studies, the effect of specific inhibitors on the regulation of PGP9.5 was also assessed by immunostaining. Treatment with the three inhibitors tested increased nuclear staining, although the significance of this possible redistribution is not

known.

Taken together, this work suggests that ErbB-2 overexpression can alter the ubiquitin-dependent degradative pathway and therefore affect overall protein turn-over. This could explain the global differences in protein expression observed between the parental cell line and the ErbB-2 overexpressing cells (Chapter 4). The up-regulation of PGP9.5 mediated by ErbB-2 overexpression might play a pivotal role in the alteration of regulatory process of protein turn-over. Firstly, it could facilitate the regeneration of the pool of free ubiquitin, which in turn enables increased rate of degradation. This could also correlate with the observation made where EGFR internalisation and degradation is increased in the ErbB-2 overexpressing cell lines, which in turn possess an increased pool of free ubiquitin. Secondly, a potential role of PGP9.5 in protein trafficking between the nuclear and the cytosolic compartments has been reported (Caballero *et al.* 2002) and will be further assessed. The increased expression of PGP9.5 may play a role in the down-regulation of $p27^{kip}$ and this appears to be an important mechanism employed by the cells to regulate their rate of proliferation (Timms *et al.* 2002). Such hypotheses need to be further tested and additional experiments beyond the scope of this thesis are required.

Chapter 7 : CONCLUSIONS AND FUTURE PROSPECTS

The family of ErbB receptor tyrosine kinases has been extensively studied over the past years and many complex signalling events have been deciphered. ErbB-2 receptor was shown to be overexpressed in 10 to 30% of invasive breast cancers (Slamon *et al.* 1987) and this receptor represents an important target in breast cancer therapy. Indeed, a humanised antibody targeted against ErbB-2, Herceptin (trastuzumab) (Carter *et al.* 1992), has already been approved for clinical use to treat patients with breast cancer (Horton 2002). However, the exact role of ErbB-2 overexpression in the development of cancer and the mechanisms by which Herceptin inhibits specific growth of ErbB-2 overexpressing cells *in vivo* are not well understood. The signalling pathways activated downstream of ErbB-2 are complex as multiple ligands can trigger different dimer combinations between each ErbB family member, and importantly ErbB-2 is the preferred partner in these dimerisation/activation events (Graus-Porta *et al.* 1997). In addition, cross-talks between signalling pathways can also occur, therefore generating an extremely complex network of signals. This has the potential to activate and repress the expression of a diverse set of genes which determines the functional proteome of a cell.

The focus of my thesis has been the investigation of the effects of ErbB-2 overexpression on the protein expression profile of a model cellular system of breast cancer in order to identify new components of this complex machinery, and ultimately to identify biomarkers of breast cancer for further targeted therapies. Several fields of research have been combined, providing important information on major kinetic events which rule the behaviour of a cell. To approach this complex machinery from a wide angle, a large-scale investigation was preferred, and this was carried out by proteomic analysis. Major improvements in proteomics have enabled the achievement of large-scale protein expression profiles such as the proteome of *P. falciparum* (Florens *et al.* 2002), *S. cerevisiae* (Perrot *et al.* 1999) or other organisms with small genomes. In addition, proteomic analysis has been applied to cancer research in the search for new biomarkers to improve diagnosis, define new therapeutic targets and eventually deliver better

prognosis for the patient (Hanash 2003).

Several techniques have been used for protein expression profiling, involving gel-based and non-gel based approaches. In this study, a new strategy of two-dimensional (2D) gel analysis was employed, two-dimensional difference gel electrophoresis or 2D-DIGE (Unlu *et al.* 1997). This protein detection method uses a panel of three fluorescent dyes to compare up to three samples on the same 2D-gel, thus providing increased experimental reproducibility and accuracy for valid statistical analysis of differential protein expression. In addition, given the wide linear dynamic range of the fluorescent signal, this technique provides a better detection strategy for global expression profiling compared to conventional methods. This approach was patented by Minden and co-workers and further commercialised by Amersham Biosciences during the early stages of this work. However, although this method was very promising, at the time when I started this study, no comprehensive biological 2D-DIGE analysis had been reported. Therefore I set out to optimise and apply this technique to the protein expression profiling of the cellular system of interest in our laboratory, namely a model cell system of ErbB-2-depedent breast cancer.

Optimisation work reported in this study showed that 2D-DIGE was a sensitive method which provided a level of detection comparable to other conventional methods, such as silver staining and SYPRO Ruby staining. Reversible labelling and comparison of protein detection patterns for each distinct cyanine dye and from replicate experiments was then assessed, validating the use of this method for accurate proteomic analysis of multiple samples. A workflow was established for the analysis of the protein expression profile mediated by ErbB-2 overexpression. Minimum detection thresholds were set and optimum conditions for protein extraction and labelling were established. In addition, 2D-DIGE-compatible post-staining methods were tested for ultimate mass spectrometric analysis of differentially expressed proteins and SYPRO Ruby staining was chosen as the most appropriate 2D-DIGE post-stain.

The main advantage of 2D-DIGE was shown in the study of the effect of ErbB-2 overexpression and growth factor treatment on the protein profile of HMLEC cells. Multiple samples were processed and 2D-DIGE provided good reproducibility in the

analysis as well as quantitative data. The use of a standard, run on each gel together with the analytical samples was shown to improve the comparison of multiple samples run on distinct gels. Indeed, the use of 2D-DIGE was "labour efficient" since this method drastically reduced the number of gels to be run by a factor of three. Importantly, the trend in protein expression over a period of time could be monitored in an accurate but non-complex statistical manner.

To assess the roles of ErbB-2 overexpression in breast cancer, two HMLECs model cell lines, HB4a and C3.6, were compared by 2D-DIGE. These cell lines are the appropriate cell type from which most breast cancers arise. In addition, C3.6 cells overexpress levels of ErbB-2 similar to that seen in carcinomas (Harris *et al.* 1999). Thus this cellular system was a good model for the study of ErbB-2-dependent breast cancer. Cells were compared by 2D-DIGE under serum-starved conditions, and 27 proteins with significant differential expression were selected and identified by MALDI-MS. Several of those have already been shown to be differentially regulated in various human cancers and could therefore play an important role in processes of human luminal epithelial cell transformation. These are PGP9.5, Hsp27, cytokeratin17, CPS1, L-plastin, radixin and glutaminase C.

Within the set of proteins identified, a subset of structural proteins was characterised, these being L-plastin, ezrin, radixin, cytokeratin 17, leucine-rich protein and lamin B. Such proteins could play a role in cell shape and adhesion. The altered expression profile of this set of proteins could correlate with the altered morphology of the cells and loss of contact inhibition associated with ErbB-2 overexpression. A second set of proteins found to be up-regulated in the ErbB-2 overexpressing cells belonged to the group of metabolic enzymes (CPS1, glutaminase C, 3-hydroxyisobutyryl-Coenzyme A hydrolase, inorganic pyrophosphatase and aldose reductase). This up-regulation of the metabolic machinery could reflect the hyper-proliferative growth of these cells. The third set of proteins found to be de-regulated in ErbB-2 overexpressing cells and the most valuable for therapeutic targeting, showed expression profiles closely related to the levels of ErbB-2 overexpression and were shown to be regulated by growth factor treatment. These were the interferon regulated genes MxA and MnSOD, the putative membrane

trafficking protein, copine III and the ubiquitin hydrolase, PGP9.5.

Several strategies were employed to validate the proteomic data and further characterise the differentially expressed proteins. Firstly, validating the protein identities and the changes in protein expression observed by 2D-DIGE and mass spectrometry was carried out by immunoblotting using specific antibodies. Secondly, immunofluorescence and cell fractionation approaches were used to assess the cellular distribution of the putative biomarkers of interest, providing possible clues as to site of action or function. Thirdly, immunopurification techniques using specific antibodies were used to characterise protein activity and function. In this study, 1DE-western blotting was employed for the validation of protein identities obtained in the 2D-DIGE comparison of HB4a and C3.6 cells. Once confident about protein expression profiles, further validation and characterisation was carried out in order to address the role of ErbB-2 overexpression on the regulation of these proteins. Consequently, protein expression patterns were characterised in other HMLECs clones and a panel of human breast cancer cell lines, all expressing variable levels of ErbB-2 and the other ErbB family members. In addition, this 2D-DIGE analysis was carried out alongside a microarray transcription profiling analysis (White *et al.*, submitted data) and a high correlation between the two sets of data was observed. Moreover, this analysis provided important information regarding the level of regulation of specific proteins either transcriptionally, or post-trancriptionally through regulation of protein turn-over.

The combination of the transcription and proteome profiling identified a whole family of interferon-inducible genes that were found to be significantly down-regulated in the ErbB-2 overexpressing cells. These include the interferon-induced viral resistance genes, Mx A and B (Haller *et al.* 1998) and a mitochondrial primary oxydant enzyme thought of as a tumour suppressor, superoxide dismutase (MnSOD) (Church *et al.* 1993). This also correlated with a decreased mRNA and protein expression level of ISGF3γ, a key transcriptional regulator of the expression of many interferon-inducible genes, amongst which MxA and MnSOD have been reported (Melen *et al.* 2000). Although the clearly defined role of these proteins in ErbB-2-mediated cellular transformation is still not clear, the work carried out herein showed that the overexpression of ErbB-2 appears

to down-regulate the interferon signalling pathway, which could be a means whereby ErbB-2 promotes cellular proliferation and resistance to interferon (IFN) therapy for breast cancer (Tossing 2001). Therefore, this implicates possible links between interferon signalling and the ErbB signalling networks, and represents a novel and interesting finding which warrants further investigation.

The next protein target which was chosen for further characterisation was copine III. The characterisation of this protein was carried out using biochemical and cellular biology approaches. Firstly, the copine III expression profile was compared in various cell lines revealing that its expression was closely correlated with ErbB-2 and ErbB-3 levels and was responsive to growth factor treatment. Protein interacting partners were also analysed by conventional immunoprecipitation of endogenous copine III. This revealed the putative interaction with the SV40 large T antigen, a protein involved in cellular immortalisation and used to derive the HB4a cells, and the Bruton tyrosine kinase associated protein, BAP-135. This was in agreement with the finding that copine III had an associated kinase activity which correlated with copine III expression levels. Thus further work needs to be carried out in order to address the role of this protein interaction and further understand the impact of ErbB-2 expression levels on the regulation of copine III.

This study also revealed the putative role of ErbB-2 in the regulation of protein turn-over of other ErbB family members. Levels of ErbB receptors were found to be altered in the ErbB-2 overexpressing cells. Indeed ErbB-3 was generally up-regulated and EGFR was down-regulated in the C3.6 cells. These changes in expression between the cell lines were not shared at the transcriptional level, as shown by microarray analysis. This suggested that the level of regulation involved is dependent upon post-translational events and could involve a differential protein turn-over between the cell lines. The rate of receptor turn-over was assessed and revealed that in the C3.6 cells, EGFR was more rapidly endocytosed and further degraded upon ligand treatment than in the HB4a cells. Further investigation showed that both cell lines did not share the same mechanism of receptor down-regulation and that the ErbB-2 overexpressing cells could employ other routes of degradation.

The role of PGP9.5 up-regulation in ErbB-2-mediated cellular transformation was also investigated and although a defined link with ErbB-2-dependent EGFR down-regulation was not unveiled, new information about its function may have been found. PGP9.5 up-regulation correlated with an increase in the pool of free ubiquitin, which in-turn could facilitate protein degradation. This would also support the observation that transcriptional and protein expression profiles were inversely correlated after 24 hrs of treatment with growth factor, suggesting a difference in protein turn-over between the parental and ErbB-2 overexpressing cells. In addition, PGP9.5 could also be involved in the regulation of cell-cycle regulators by triggering the trafficking of proteins from the nuclei to the cytoplasm, as had been observed for the cyclin dependent kinase inhibitor, $p27^{kip}$ (Caballero *et al.* 2002). This result correlates with cellular localisation and degradation of $p27^{kip}$ in the C3.6 cells and with increased cell cycling and proliferation.

Taken together, this 2D-DIGE analysis represents a starting point in the understanding of ErbB-2-dependent cellular transformation. However, further work beyond the scope of this study now needs to be carried out using other strategies. All the characterisation work carried out herein was performed on the endogenous proteins and using constructs which express tagged proteins (such as myc-tag) would greatly improve the specificity of the analysis. Furthermore, experiments of gene silencing to alter the expression of specific proteins of interest, using double stranded interference RNA (Cullen 2002; Famulok and Verma 2002), could be usefully applied in our model cell system. As an inverse strategy, ectopic overexpression of the proteins of interest (such as copine III, MxA or PGP9.5) can also be carried out to assess their function and their putative role in the regulation of the transforming capabilities of ErbB-2 overexpression.

This current 2D-DIGE protein expression profiling of a model cell system of breast cancer was compared to other published proteomic studies of breast cancer and a good correlation in the dataset generated was observed (Hondermarck *et al.* 2001; Vercoutter-Edouart *et al.* 2001; Wulfkuhle *et al.* 2002). A study carried out by Wulfukuhle *et al.*, from laser capture microdissections (LCM) and from whole tissue samples (Wulfkuhle *et al.* 2002) generated similar results to the ones obtained herein. Namely L-plastin, Hsp27 or another member of the 14-3-3 family of proteins, 14-3-

3δ were found to be up-regulated in the panel of LCM tissue samples, thus correlating with the differential proteomic profiling reported herein between the HMLECs parental and ErbB-2 overexpressing cell lines. However, no evidence of statistical reproducibility was reported in this analysis and the global subset of proteins identified appeared to be mostly abundant proteins such as the cytokeratins and heat shock proteins.

Thus I have shown here that the use of 2D-DIGE can enable the identification of important putative protein targets that have potential regulatory roles in ErbB-2-mediated transformation. The proteomic approach undertaken here has provided crucial preliminary information regarding the effects of ErbB-2 overexpression, giving rise to multiple interesting hypotheses to be followed up on.

However, an important issue in the use of 2D-DIGE, or 2DE-based separation methods in general, remains the limitations in sensitivity and proteome coverage. Due to the wide dynamic range in protein expression between the high and low abundant proteins, proteomic analyses favour the characterisation of highly abundant proteins. Therefore the same "usual suspects" often appear as potential targets of differential protein expression, such as the heat shock proteins or the cytokeratins. Although these proteins may represent important markers for diagnosis, these may be secondary effects and therefore may not be suitable therapeutic targets. Improvements have enabled increased sample loads, protein solubility and separation using narrow-range pH strips (Gorg *et al.* 2000; Westbrook *et al.* 2001) or sample pre-fractionation via liquid chromatography (VerBerkmoes *et al.* 2002) or the isolation of organelles (Bardel *et al.* 2002; Schirmer and Gerace 2002). However, other methods of study need to be combined with conventional gel-based methods. Indeed, protein recovery and separation have also been improved using non-gel based methods such as chromatographic separation methods coupled to mass spectrometric analysis. For instance, the multi-dimensional chromatography (MudPIT) (Washburn *et al.* 2001), which can be coupled to differential isotope tagging methods such as ICAT (Gygi *et al.* 1999), or the use of protein chips for a more targeted study of the functional proteome (Jenkins and Pennington 2001) are interesting approaches to overcome the limitations faced by 2DE-based proteomic. These tools are very promising as they can be automated and can make possible the

identification of a wide range of proteins. However, these methods seem to be more adapted for the systematic analysis of all proteins expressed in a cell system, rather than for the differential protein expression profiling of multiple samples. There is therefore a need in proteomic analysis to combine optimum methods for improved sample coverage and functional characterisation of multiple samples.

Optimisation of 2D-DIGE can also be achieved in order to overcome the limitations it currently faces due to the necessary post-staining step and the limits of sensitivity. Therefore research continues on possible improvements to the present 2D-DIGE technique and new dyes are being investigated in our laboratory in order to target other amino acids for improved sensitivity or selectivity. Cyanine dyes have been developed which are targeted to cysteine residues with an iodoacetamide (IA) reactive group and both IA-Cy3 and IA-Cy5 version have already been synthesised (also called ICy3 and ICy5). Although all proteins do not possess cysteines, such targets could provide a means to label proteins to semi-saturation and possibly to enhance the level of sensitivity of detection. In addition, under such labelling conditions, all cystein-containing proteins should be covalently labelled and could therefore be directly picked from the DIGE image without the need to use a post-stain. Current work in our laboratory is being carried out to establish optimum labelling conditions of proteins under native or reduced conditions. Furthermore, multiple applications can also be applied for the use of these fluorescent tags such as the native labelling of proteins under various growth conditions to assess cysteine-accessibility and thus protein activity. These dyes could also be combined to label specific subset of proteins, such as tagging the phosphoproteome by 2D-DIGE. The ICy Dyes could be used to label phosphospecific peptides post β-elimination and Michael addition, as was proposed with the PhIAT approach (Goshe *et al.* 2001).

Since the beginning of this work, a small number of new reports on the validation of 2D-DIGE for statistical analysis and comprehensive biological proteomics have been published. These reports emphasised the power of the method and optimisation for each experimental design showed the flexibility of 2D-DIGE for the analysis of various biological samples. 2D-DIGE has already been applied to the study of the mouse heart

(Kernec *et al.* 2001), mouse liver (Ruepp *et al.* 2002), laser capture microdissection of esophageal carcinoma (Zhou *et al.* 2002) or even rat brain, platelets or organelles (work in progress within our laboratory and other collaborations).

Taken together, the work carried out in this study and by other groups show the power of 2D-DIGE as a sensitive, reproducible and accurate protein detection method. Importantly, the ability to separate up to three distinct samples on the same 2DE gel does not only reduce the workload, but also provides a gain in reproducibility and quantitation of differential protein expression between multiple samples. The work reported here has revealed the importance of several subsets of proteins which could play important roles in the development of ErbB-2-dependent breast cancers.

REFERENCES

Adams, J. and Elliott, P.J. (2000). "New agents in cancer clinical trials." Oncogene **19**(56): 6687-92.

Adams, J., Palombella, V.J., Sausville, E.A., Johnson, J., Destree, A., Lazarus, D.D., Maas, J., Pien, C.S., Prakash, S. and Elliott, P.J. (1999). "Proteasome inhibitors: a novel class of potent and effective antitumor agents." Cancer Res **59**(11): 2615-22.

Aebersold, R. and Mann, M. (2003). "Mass spectrometry-based proteomics." Nature **422**(6928): 198-207.

Aghajanian, C., Soignet, S., Dizon, D.S., Pien, C.S., Adams, J., Elliott, P.J., Sabbatini, P., Miller, V., Hensley, M.L., Pezzulli, S., Canales, C., Daud, A. and Spriggs, D.R. (2002). "A phase I trial of the novel proteasome inhibitor PS341 in advanced solid tumor malignancies." Clin Cancer Res **8**(8): 2505-11.

Alban, A., David, S.O., Bjorkesten, L., Andersson, C., Sloge, E., Lewis, S. and Currie, I. (2003). "A novel experimental design for comparative two-dimensional gel analysis: Two-dimensional difference gel electrophoresis incorporating a pooled internal standard." Proteomics **3**(1): 36-44.

Aplin, A.E., Stewart, S.A., Assoian, R.K. and Juliano, R.L. (2001). "Integrin-mediated adhesion regulates ERK nuclear translocation and phosphorylation of Elk-1." J Cell Biol **153**(2): 273-82.

Arnheiter, H., Skuntz, S., Noteborn, M., Chang, S. and Meier, E. (1990). "Transgenic mice with intracellular immunity to influenza virus." Cell **62**(1): 51-61.

Baeckstrom, D., Lu, P.J. and Taylor-Papadimitriou, J. (2000). "Activation of the alpha2beta1 integrin prevents c-erbB2-induced scattering and apoptosis of human mammary epithelial cells in collagen." Oncogene **19**(40): 4592-603.

Bange, J., Zwick, E. and Ullrich, A. (2001). "Molecular targets for breast cancer therapy and prevention." Nat Med **7**(5): 548-52.

Bardel, J., Louwagie, M., Jaquinod, M., Jourdain, A., Luche, S., Rabilloud, T., Macherel, D., Garin, J. and Bourguignon, J. (2002). "A survey of the plant mitochondrial proteome in relation to development." Proteomics **2**(7): 880-98.

Bartek, J., Bartkova, J., Kyprianou, N., Lalani, E.N., Staskova, Z., Shearer, M., Chang, S. and Taylor-Papadimitriou, J. (1991). "Efficient immortalization of luminal epithelial cells from human mammary gland by introduction of simian virus 40 large tumor antigen with a recombinant retrovirus." Proc Natl Acad Sci U S A **88**(9): 3520-4.

Baselga, J. (2001). "The EGFR as a target for anticancer therapy--focus on cetuximab." Eur J Cancer **37 Suppl 4**: S16-22.

Bates, S.E., Valverius, E.M., Ennis, B.W., Bronzert, D.A., Sheridan, J.P., Stampfer, M.R., Mendelsohn, J., Lippman, M.E. and Dickson, R.B. (1990). "Expression of the transforming growth factor-alpha/epidermal growth factor receptor pathway in normal human breast epithelial cells." Endocrinology **126**(1): 596-607.

Baulida, J., Kraus, M.H., Alimandi, M., Di Fiore, P.P. and Carpenter, G. (1996). "All ErbB receptors other than the epidermal growth factor receptor are endocytosis impaired." J Biol Chem **271**(9): 5251-7.

Bell, A.W., Ward, M.A., Blackstock, W.P., Freeman, H.N., Choudhary, J.S., Lewis, A.P., Chotai, D., Fazel, A., Gushue, J.N., Paiement, J., Palcy, S., Chevet, E., Lafreniere-Roula, M., Solari, R., Thomas, D.Y., Rowley, A. and Bergeron, J.J. (2001). "Proteomics characterization of abundant Golgi membrane proteins." J Biol Chem **276**(7): 5152-65.

Benvenuti, S., Cramer, R., Bruce, J., Waterfield, M.D. and Jat, P.S. (2002). "Identification of novel candidates for replicative senescence by functional proteomics." Oncogene **21**(28): 4403-13.

Berggren, K.N., Schulenberg, B., Lopez, M.F., Steinberg, T.H., Bogdanova, A., Smejkal, G., Wang, A. and Patton, W.F. (2002). "An improved formulation of SYPRO Ruby protein gel stain: Comparison with the original formulation and with a ruthenium II tris (bathophenanthroline disulfonate) formulation." Proteomics **2**(5): 486-98.

Biemann, K. and Scoble, H.A. (1987). "Characterization by tandem mass spectrometry of structural modifications in proteins." Science **237**(4818): 992-8.

Biemann, K. (1990). "Appendix 5. Nomenclature for peptide fragment ions (positive ions)." Methods Enzymol **193**: 886-7.

Bjellqvist, B., Ek, K., Righetti, P.G., Gianazza, E., Gorg, A., Westermeier, R. and Postel, W. (1982). "Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications." J Biochem Biophys Methods **6**(4): 317-39.

Blow, J.J. and Hodgson, B. (2002). "Replication licensing-defining the proliferative state?" Trends Cell Biol **12**(2): 72-8.

Blow, J.J. and Laskey, R.A. (1988). "A role for the nuclear envelope in controlling DNA replication within the cell cycle." Nature **332**(6164): 546-8.

Borden, E.C., Esserman, L., Linder, D.J., Campbell, M.J. and Fulton, A.M. (1999). "Biological therapies for breast carcinoma: concepts for improvement in survival." Semin Oncol **26**(4 Suppl 12): 28-40.

Brichory, F., Beer, D., Le Naour, F., Giordano, T. and Hanash, S. (2001). "Proteomics-based identification of protein gene product 9.5 as a tumor antigen that induces a humoral immune response in lung cancer." Cancer Res **61**(21): 7908-12.

Burke, P.M. and Wiley, H.S. (1999). "Human mammary epithelial cells rapidly exchange empty EGFR between surface and intracellular pools." J Cell Physiol **180**(3): 448-60.

Caballero, O.L., Resto, V., Patturajan, M., Meerzaman, D., Guo, M.Z., Engles, J., Yochem, R., Ratovitski, E., Sidransky, D. and Jen, J. (2002). "Interaction and colocalization of PGP9.5 with JAB1 and p27(Kip1)." Oncogene **21**(19): 3003-10.

Cantley, L.C., Auger, K.R., Carpenter, C., Duckworth, B., Graziani, A., Kapeller, R. and Soltoff, S. (1991). "Oncogenes and signal transduction." Cell **64**(2): 281-302.

Carraway, K.L., 3rd, Sliwkowski, M.X., Akita, R., Platko, J.V., Guy, P.M., Nuijens, A., Diamonti, A.J., Vandlen, R.L., Cantley, L.C. and Cerione, R.A. (1994). "The *erbB3* gene product is a receptor for heregulin." J Biol Chem **269**(19): 14303-6.

Carter, P., Presta, L., Gorman, C.M., Ridgway, J.B., Henner, D., Wong, W.L., Rowland, A.M., Kotts, C., Carver, M.E. and Shepard, H.M. (1992). "Humanization of an

anti-p185HER2 antibody for human cancer therapy." Proc Natl Acad Sci U S A **89**(10): 4285-9.

Castellanos-Serra L, Huerta V, Moritz RL, Simpson RJ (1999). "Proteome analysis of polyacrylamide gel-separated proteins visualized by reversible negative staining using imidazole-zinc salts." Electrophoresis **20**(4-5): 732-7.

Caudell, E.G., Caudell, J.J., Tang, C.H., Yu, T.K., Frederick, M.J. and Grimm, E.A. (2000). "Characterization of human copine III as a phosphoprotein with associated kinase activity." Biochemistry **39**(42): 13034-43.

Celis, J.E., Ostergaard, M., Rasmussen, H.H., Gromov, P., Gromova, I., Varmark, H., Palsdottir, H., Magnusson, N., Andersen, I., Basse, B., Lauridsen, J.B., Ratz, G., Wolf, H., Orntoft, T.F., Celis, P. and Celis, A. (1999). "A comprehensive protein resource for the study of bladder cancer: http://biobase.dk/cgi-bin/celis." Electrophoresis **20**(2): 300-9.

Chan, R., Hardy, W.R., Laing, M.A., Hardy, S.E. and Muller, W.J. (2002). "The catalytic activity of the ErbB-2 receptor tyrosine kinase is essential for embryonic development." Mol Cell Biol **22**(4): 1073-8.

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. and Beer, D.G. (2002). "Discordant protein and mRNA expression in lung adenocarcinomas." Mol Cell Proteomics **1**(4): 304-13.

Chen, G., Gharib, T.G., Huang, C.C., Thomas, D.G., Shedden, K.A., Taylor, J.M., Kardia, S.L., Misek, D.E., Giordano, T.J., Iannettoni, M.D., Orringer, M.B., Hanash, S.M. and Beer, D.G. (2002). "Proteomic analysis of lung adenocarcinoma: identification of a highly expressed set of proteins in tumors." Clin Cancer Res **8**(7): 2298-305.

Church, S.L., Grant, J.W., Ridnour, L.A., Oberley, L.W., Swanson, P.E., Meltzer, P.S. and Trent, J.M. (1993). "Increased manganese superoxide dismutase expression suppresses the malignant phenotype of human melanoma cells." Proc Natl Acad Sci U S A **90**(7): 3113-7.

Ciardiello, F. (2000). "Epidermal growth factor receptor tyrosine kinase inhibitors as

anticancer agents." Drugs **60**(Suppl 1): 25-32; discussion 41-2.

Claret, F.X., Hibi, M., Dhut, S., Toda, T. and Karin, M. (1996). "A new group of conserved coactivators that increase the specificity of AP-1 transcription factors." Nature **383**(6599): 453-7.

Corthals, G.L., Wasinger, V.C., Hochstrasser, D.F. and Sanchez, J.C. (2000). "The dynamic range of protein expression: a challenge for proteomic research." Electrophoresis **21**(6): 1104-15.

Cox, D.M., Du, M., Guo, X., Siu, K.W. and McDermott, J.C. (2002). "Tandem affinity purification of protein complexes from mammalian cells." Biotechniques **33**(2): 267-8, 270.

Creutz, C.E., Tomsig, J.L., Snyder, S.L., Gautier, M.C., Skouri, F., Beisson, J. and Cohen, J. (1998). "The copines, a novel class of C2 domain-containing, calcium-dependent, phospholipid-binding proteins conserved from Paramecium to humans." J Biol Chem **273**(3): 1393-402.

Cullen, B.R. (2002). "RNA interference: antiviral defense and genetic tool." Nat Immunol **3**(7): 597-9.

Daban, J.R. (2001). "Fluorescent labeling of proteins with nile red and 2-methoxy-2,4-diphenyl-3(2H)-furanone: physicochemical basis and application to the rapid staining of sodium dodecyl sulfate polyacrylamide gels and Western blots." Electrophoresis **22**(5): 874-80.

Daban, J.R., Bartolome, S. and Samso, M. (1991). "Use of the hydrophobic probe Nile red for the fluorescent staining of protein bands in sodium dodecyl sulfate-polyacrylamide gels." Anal Biochem **199**(2): 169-74.

Dahmann, C., Diffley, J.F. and Nasmyth, K.A. (1995). "S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state." Curr Biol **5**(11): 1257-69.

Day, I.N. and Thompson, R.J. (1987). "Molecular cloning of cDNA coding for human PGP 9.5 protein. A novel cytoplasmic marker for neurones and neuroendocrine cells." FEBS Lett **210**(2): 157-60.

Egeblad, M., Mortensen, O.H., van Kempen, L.C. and Jaattela, M. (2001).

"BIBX1382BS, but not AG1478 or PD153035, inhibits the ErbB kinases at different concentrations in intact cells." Biochem Biophys Res Commun **281**(1): 25-31.

Famulok, M. and Verma, S. (2002). "In vivo-applied functional RNAs as tools in proteomics and genomics research." Trends Biotechnol **20**(11): 462-6.

Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M. (1989). "Electrospray ionization for mass spectrometry of large biomolecules." Science **246**(4926): 64-71.

Ferguson, A.T., Evron, E., Umbricht, C.B., Pandita, T.K., Chan, T.A., Hermeking, H., Marks, J.R., Lambers, A.R., Futreal, P.A., Stampfer, M.R. and Sukumar, S. (2000). "High frequency of hypermethylation at the 14-3-3 sigma locus leads to gene silencing in breast cancer." Proc Natl Acad Sci U S A **97**(11): 6049-54.

Ficarro, S., Chertihin, O., Westbrook, V.A., White, F., Jayes, F., Kalab, P., Marto, J.A., Shabanowitz, J., Herr, J.C., Hunt, D.F. and Visconti, P.E. (2003). "Phosphoproteome analysis of capacitated human sperm. Evidence of tyrosine phosphorylation of a kinase-anchoring protein 3 and valosin- containing protein/p97 during capacitation." J Biol Chem **278**(13): 11579-89.

Fields, S. and Song, O. (1989). "A novel genetic system to detect protein-protein interactions." Nature **340**(6230): 245-6.

Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D., Moch, J.K., Muster, N., Sacci, J.B., Tabb, D.L., Witney, A.A., Wolters, D., Wu, Y., Gardner, M.J., Holder, A.A., Sinden, R.E., Yates, J.R. and Carucci, D.J. (2002). "A proteomic view of the *Plasmodium falciparum* life cycle." Nature **419**(6906): 520-6.

Forler, D., Kocher, T., Rode, M., Gentzel, M., Izaurralde, E. and Wilm, M. (2003). "An efficient protein complex purification method for functional proteomics in higher eukaryotes." Nat Biotechnol **21**(1): 89-92.

Gavin, A.C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.M., Cruciat, C.M., Remor, M., Hofert, C., Schelder, M., Brajenovic, M., Ruffner, H., Merino, A., Klein, K., Hudak, M., Dickson, D.,

Rudi, T., Gnau, V., Bauch, A., Bastuck, S., Huhse, B., Leutwein, C., Heurtier, M.A., Copley, R.R., Edelmann, A., Querfurth, E., Rybin, V., Drewes, G., Raida, M., Bouwmeester, T., Bork, P., Seraphin, B., Kuster, B., Neubauer, G. and Superti-Furga, G. (2002). "Functional organization of the yeast proteome by systematic analysis of protein complexes." Nature **415**(6868): 141-7.

Gorg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R. and Weiss, W. (2000). "The current state of two-dimensional electrophoresis with immobilized pH gradients." Electrophoresis **21**(6): 1037-53.

Gorg, A., Obermaier, C., Boguth, G. and Weiss, W. (1999). "Recent developments in two-dimensional gel electrophoresis with immobilized pH gradients: wide pH gradients up to pH 12, longer separation distances and simplified procedures." Electrophoresis **20**(4-5): 712-7.

Goshe, M.B., Conrads, T.P., Panisko, E.A., Angell, N.H., Veenstra, T.D. and Smith, R.D. (2001). "Phosphoprotein isotope-coded affinity tag approach for isolating and quantitating phosphopeptides in proteome-wide analyses." Anal Chem **73**(11): 2578-86.

Gottlieb, T.M. and Jackson, S.P. (1993). "The DNA-dependent protein kinase: requirement for DNA ends and association with Ku antigen." Cell **72**(1): 131-42.

Graus-Porta, D., Beerli, R.R., Daly, J.M. and Hynes, N.E. (1997). "ErbB-2, the preferred heterodimerization partner of all ErbB receptors, is a mediator of lateral signaling." Embo J **16**(7): 1647-55.

Guy, P.M., Platko, J.V., Cantley, L.C., Cerione, R.A. and Carraway, K.L., 3rd (1994). "Insect cell-expressed p180erbB3 possesses an impaired tyrosine kinase activity." Proc Natl Acad Sci U S A **91**(17): 8132-6.

Gygi, S.P., Corthals, G.L., Zhang, Y., Rochon, Y. and Aebersold, R. (2000). "Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology." Proc Natl Acad Sci U S A **97**(17): 9390-5.

Gygi, S.P., Rist, B., Gerber, S.A., Turecek, F., Gelb, M.H. and Aebersold, R. (1999). "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags." Nat Biotechnol **17**(10): 994-9.

Gygi, S.P., Rochon, Y., Franza, B.R. and Aebersold, R. (1999). "Correlation between protein and mRNA abundance in yeast." Mol Cell Biol **19**(3): 1720-30.

Hackel, P.O., Zwick, E., Prenzel, N. and Ullrich, A. (1999). "Epidermal growth factor receptors: critical mediators of multiple receptor pathways." Curr Opin Cell Biol **11**(2): 184-9.

Hahn, W.C., Counter, C.M., Lundberg, A.S., Beijersbergen, R.L., Brooks, M.W. and Weinberg, R.A. (1999). "Creation of human tumour cells with defined genetic elements." Nature **400**(6743): 464-8.

Haj, F.G., Markova, B., Klaman, L.D., Bohmer, F.D. and Neel, B.G. (2003). "Regulation of receptor tyrosine kinase signaling by protein tyrosine phosphatase-1B." J Biol Chem **278**(2): 739-44.

Haller, O., Frese, M. and Kochs, G. (1998). "Mx proteins: mediators of innate resistance to RNA viruses." Rev Sci Tech **17**(1): 220-30.

Hanahan, D. and Weinberg, R.A. (2000). "The hallmarks of cancer." Cell **100**(1): 57-70.

Hanash, S. (2003). "Disease proteomics." Nature **422**(6928): 226-32.

Hanash, S.M. (2000). "Biomedical applications of two-dimensional electrophoresis using immobilized pH gradients: current status." Electrophoresis **21**(6): 1202-9.

Hanash, S.M., Strahler, J.R., Neel, J.V., Hailat, N., Melhem, R., Keim, D., Zhu, X.X., Wagner, D., Gage, D.A. and Watson, J.T. (1991). "Highly resolving two-dimensional gels for protein sequencing." Proc Natl Acad Sci U S A **88**(13): 5709-13.

Harari, D., Tzahar, E., Romano, J., Shelly, M., Pierce, J.H., Andrews, G.C. and Yarden, Y. (1999). "Neuregulin-4: a novel growth factor that acts through the ErbB-4 receptor tyrosine kinase." Oncogene **18**(17): 2681-9.

Harris, R.A., Eichholtz, T.J., Hiles, I.D., Page, M.J. and O'Hare, M.J. (1999). "New model of ErbB-2 over-expression in human mammary luminal epithelial cells." Int J Cancer **80**(3): 477-84.

Harris, R.A., Yang, A., Stein, R.C., Lucy, K., Brusten, L., Herath, A., Parekh, R., Waterfield, M.D., O'Hare, M.J., Neville, M.A., Page, M.J. and Zvelebil, M.J. (2002). "Cluster analysis of an extensive human breast cancer cell line protein

expression map database." Proteomics **2**(2): 212-23.

Hartman, B.K. and Udenfriend, S. (1969). "A method for immediate visualization of proteins in acrylamide gels and its use for preparation of antibodies to enzymes." Anal Biochem **30**(3): 391-4.

Harvat, B.L. and Jetten, A.M. (1996). "Gamma-interferon induces an irreversible growth arrest in mid-G1 in mammary epithelial cells which correlates with a block in hyperphosphorylation of retinoblastoma." Cell Growth Differ **7**(3): 289-300.

Henzel, W.J., Billeci, T.M., Stults, J.T., Wong, S.C., Grimley, C. and Watanabe, C. (1993). "Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases." Proc Natl Acad Sci U S A **90**(11): 5011-5.

Herbert, B. (1999). "Advances in protein solubilisation for two-dimensional electrophoresis." Electrophoresis **20**(4-5): 660-3.

Hibi, K., Westra, W.H., Borges, M., Goodman, S., Sidransky, D. and Jen, J. (1999). "PGP9.5 as a candidate tumor marker for non-small-cell lung cancer." Am J Pathol **155**(3): 711-5.

Hicke, L. (1997). "Ubiquitin-dependent internalization and down-regulation of plasma membrane proteins." Faseb J **11**(14): 1215-26.

Hochstrasser DF, P.A., Merril CR (1988). "Development of polyacrylamide gels that improve the separation of proteins and their detection by silver staining." Anal Biochem **173**(2): 412-23.

Hondermarck, H., Vercoutter-Edouart, A.S., Revillion, F., Lemoine, J., el-Yazidi-Belkoura, I., Nurcombe, V. and Peyrat, J.P. (2001). "Proteomics of breast cancer for marker discovery and signal pathway profiling." Proteomics **1**(10): 1216-32.

Horowitz, P.M. and Bowman, S. (1987). "Ion-enhanced fluorescence staining of sodium dodecyl sulfate- polyacrylamide gels using bis(8-p-toluidino-1-naphthalenesulfonate)." Anal Biochem **165**(2): 430-4.

Horton, J. (2002). "Trastuzumab use in breast cancer: clinical issues." Cancer Control **9**(6): 499-507.

Hua, X.H., Yan, H. and Newport, J. (1997). "A role for Cdk2 kinase in negatively

regulating DNA replication during S phase of the cell cycle." J Cell Biol **137**(1): 183-92.

Hunt, D.F., Shabanowitz, J., Yates, J.R., 3rd, Zhu, N.Z., Russell, D.H. and Castro, M.E. (1987). "Tandem quadrupole Fourier-transform mass spectrometry of oligopeptides and small proteins." Proc Natl Acad Sci U S A **84**(3): 620-3.

Hunter, T. (2000). "Signaling-2000 and beyond." Cell **100**(1): 113-27.

Iacopino, F., Ferrandina, G., Scambia, G., Benedetti-Panici, P., Mancuso, S. and Sica, G. (1996). "Interferons inhibit EGF-stimulated cell growth and reduce EGF binding in human breast cancer cells." Anticancer Res **16**(4A): 1919-24.

Issaq, H.J., Veenstra, T.D., Conrads, T.P. and Felschow, D. (2002). "The SELDI-TOF MS approach to proteomics: protein profiling and biomarker identification." Biochem Biophys Res Commun **292**(3): 587-92.

Ito, A., Shinkai, M., Honda, H., Yoshikawa, K., Saga, S., Wakabayashi, T., Yoshida, J. and Kobayashi, T. (2003). "Heat shock protein 70 expression induces antitumor immunity during intracellular hyperthermia using magnetite nanoparticles." Cancer Immunol Immunother **52**(2): 80-8.

James, P., Quadroni, M., Carafoli, E. and Gonnet, G. (1993). "Protein identification by mass profile fingerprinting." Biochem Biophys Res Commun **195**(1): 58-64.

Jat, P.S. and Sharp, P.A. (1989). "Cell lines established by a temperature-sensitive simian virus 40 large- T-antigen gene are growth restricted at the nonpermissive temperature." Mol Cell Biol **9**(4): 1672-81.

Jenkins, R.E. and Pennington, S.R. (2001). "Arrays for protein expression profiling: towards a viable alternative to two-dimensional gel electrophoresis?" Proteomics **1**(1): 13-29.

Johnson, N.L., Gardner, A.M., Diener, K.M., Lange-Carter, C.A., Gleavy, J., Jarpe, M.B., Minden, A., Karin, M., Zon, L.I. and Johnson, G.L. (1996). "Signal transduction pathways regulated by mitogen- activated/extracellular response kinase kinase kinase induce cell death." J Biol Chem **271**(6): 3229-37.

Jones, S.L., Wang, J., Turck, C.W. and Brown, E.J. (1998). "A role for the actin-bundling protein L-plastin in the regulation of leukocyte integrin function." Proc Natl Acad

Sci U S A **95**(16): 9331-6.

Karas, M. and Hillenkamp, F. (1988). "Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons." Anal Chem **60**(20): 2299-301.

Kenrick, K.G. and Margolis, J. (1970). "Isoelectric focusing and gradient gel electrophoresis: a two- dimensional technique." Anal Biochem **33**(1): 204-7.

Kernec, F., Unlu, M., Labeikovsky, W., Minden, J.S. and Koretsky, A.P. (2001). "Changes in the mitochondrial proteome from mouse hearts deficient in creatine kinase." Physiol Genomics **6**(2): 117-28.

Kim, H.H., Vijapurkar, U., Hellyer, N.J., Bravo, D. and Koland, J.G. (1998). "Signal transduction by epidermal growth factor and heregulin via the kinase-deficient ErbB3 protein." Biochem J **334**(Pt 1): 189-95.

Klawitz, I., Preuss, U. and Scheidtmann, K.H. (2001). "Interaction of SV40 large T antigen with components of the nucleo/cytoskeleton." Int J Oncol **19**(6): 1325-32.

Klose, J. (1999). "Large-gel 2-D electrophoresis." Methods Mol Biol **112**: 147-72.

Klose, J. and Spielmann, H. (1975). "Gel isoelectric focusing of mouse lactate dehydrogenase: heterogeneity of the isoenzymes A4 and X4." Biochem Genet **13**(9-10): 707-20.

Koller, A., Washburn, M.P., Lange, B.M., Andon, N.L., Deciu, C., Haynes, P.A., Hays, L., Schieltz, D., Ulaszek, R., Wei, J., Wolters, D. and Yates, J.R., 3rd (2002). "Proteomic survey of metabolic pathways in rice." Proc Natl Acad Sci U S A **99**(18): 11969-74.

Land, H., Parada, L.F. and Weinberg, R.A. (1983). "Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes." Nature **304**(5927): 596-602.

Lane, H.A., Beuvink, I., Motoyama, A.B., Daly, J.M., Neve, R.M. and Hynes, N.E. (2000). "ErbB2 potentiates breast tumor proliferation through modulation of p27(Kip1)-Cdk2 complex formation: receptor overexpression does not determine growth dependency." Mol Cell Biol **20**(9): 3210-23.

Lapillonne, A., Coue, O., Friederich, E., Nicolas, A., Del Maestro, L., Louvard, D., Robine, S. and Sastre-Garau, X. (2000). "Expression patterns of L-plastin isoform

in normal and carcinomatous breast tissues." Anticancer Res **20**(5A): 3177-82.

Larsen, C.N., Krantz, B.A. and Wilkinson, K.D. (1998). "Substrate specificity of deubiquitinating enzymes: ubiquitin C-terminal hydrolases." Biochemistry **37**(10): 3358-68.

Le Bivic, A., Real, F.X. and Rodriguez-Boulan, E. (1989). "Vectorial targeting of apical and basolateral plasma membrane proteins in a human adenocarcinoma epithelial cell line." Proc Natl Acad Sci U S A **86**(23): 9313-7.

Lee, J.O., Rieu, P., Arnaout, M.A. and Liddington, R. (1995). "Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18)." Cell **80**(4): 631-8.

Lee, K.F., Simon, H., Chen, H., Bates, B., Hung, M.C. and Hauser, C. (1995). "Requirement for neuregulin receptor erbB2 in neural and cardiac development." Nature **378**(6555): 394-8.

Lenferink, A.E., Busse, D., Flanagan, W.M., Yakes, F.M. and Arteaga, C.L. (2001). "ErbB2/neu kinase modulates cellular p27(Kip1) and cyclin D1 through multiple signaling pathways." Cancer Res **61**(17): 6583-91.

Lenferink, A.E., Pinkas-Kramarski, R., van de Poll, M.L., van Vugt, M.J., Klapper, L.N., Tzahar, E., Waterman, H., Sela, M., van Zoelen, E.J. and Yarden, Y. (1998). "Differential endocytic routing of homo- and hetero-dimeric ErbB tyrosine kinases confers signaling superiority to receptor heterodimers." Embo J **17**(12): 3385-97.

Levkowitz, G., Klapper, L.N., Tzahar, E., Freywald, A., Sela, M. and Yarden, Y. (1996). "Coupling of the c-Cbl protooncogene product to ErbB-1/EGF-receptor but not to other ErbB proteins." Oncogene **12**(5): 1117-25.

Levkowitz, G., Waterman, H., Ettenberg, S.A., Katz, M., Tsygankov, A.Y., Alroy, I., Lavi, S., Iwai, K., Reiss, Y., Ciechanover, A., Lipkowitz, S. and Yarden, Y. (1999). "Ubiquitin ligase activity and tyrosine phosphorylation underlie suppression of growth factor signaling by c-Cbl/Sli-1." Mol Cell **4**(6): 1029-40.

Levkowitz, G., Waterman, H., Zamir, E., Kam, Z., Oved, S., Langdon, W.Y., Beguinot, L., Geiger, B. and Yarden, Y. (1998). "c-Cbl/Sli-1 regulates endocytic sorting and

ubiquitination of the epidermal growth factor receptor." Genes Dev **12**(23): 3663-74.

Lewis, T.S., Hunt, J.B., Aveline, L.D., Jonscher, K.R., Louie, D.F., Yeh, J.M., Nahreini, T.S., Resing, K.A. and Ahn, N.G. (2000). "Identification of novel MAP kinase pathway signaling targets by functional proteomics and mass spectrometry." Mol Cell **6**(6): 1343-54.

Li, J.J., Oberley, L.W., St Clair, D.K., Ridnour, L.A. and Oberley, T.D. (1995). "Phenotypic changes induced in human breast cancer cells by overexpression of manganese-containing superoxide dismutase." Oncogene **10**(10): 1989-2000.

Lin, C.S., Aebersold, R.H., Kent, S.B., Varma, M. and Leavitt, J. (1988). "Molecular cloning and characterization of plastin, a human leukocyte protein expressed in transformed human fibroblasts." Mol Cell Biol **8**(11): 4659-68.

Lin, C.S., Chen, Z.P., Park, T., Ghosh, K. and Leavitt, J. (1993). "Characterization of the human L-plastin gene promoter in normal and neoplastic cells." J Biol Chem **268**(4): 2793-801.

Lindberg, L.E., Hedjazifar, S. and Baeckstrom, D. (2002). "c-erbB2-induced Disruption of Matrix Adhesion and Morphogenesis Reveals a Novel Role for Protein Kinase B as a Negative Regulator of alpha(2)beta(1) Integrin Function." Mol Biol Cell **13**(8): 2894-908.

Lobo, C., Ruiz-Bellido, M.A., Aledo, J.C., Marquez, J., Nunez De Castro, I. and Alonso, F.J. (2000). "Inhibition of glutaminase expression by antisense mRNA decreases growth and tumourigenicity of tumour cells." Biochem J **348 Pt 2**: 257-61.

Longva, K.E., Blystad, F.D., Stang, E., Larsen, A.M., Johannessen, L.E. and Madshus, I.H. (2002). "Ubiquitination and proteasomal activity is required for transport of the EGF receptor to inner membranes of multivesicular bodies." J Cell Biol **156**(5): 843-54.

Lopez, M.F., Berggren, K., Chernokalskaya, E., Lazarev, A., Robinson, M. and Patton, W.F. (2000). "A comparison of silver stain and SYPRO Ruby Protein Gel Stain with respect to protein detection in two-dimensional gels and identification by peptide mass profiling." Electrophoresis **21**(17): 3673-83.

Lopez, M.F., Kristal, B.S., Chernokalskaya, E., Lazarev, A., Shestopalov, A.I., Bogdanova, A. and Robinson, M. (2000). "High-throughput profiling of the mitochondrial proteome using affinity fractionation and automation." Electrophoresis **21**(16): 3427-40.

Ma, J. and Ptashne, M. (1988). "Converting a eukaryotic transcriptional inhibitor into an activator." Cell **55**(3): 443-6.

Mann, M., Hojrup, P. and Roepstorff, P. (1993). "Use of mass spectrometric molecular weight information to identify proteins in sequence databases." Biol Mass Spectrom **22**(6): 338-45.

Maemura, M., Iino, Y., Horiguchi, J., Takei, H., Horii, Y., Koibuchi, Y., Yokoe, T., Takeyoshi, I., Ohwada, S. and Morishita, Y. (1999). "Effects of interferon-alpha on cellular proliferation and adhesion of breast carcinoma cells." Oncol Rep **6**(3): 557-61.

Melen, K., Keskinen, P., Lehtonen, A. and Julkunen, I. (2000). "Interferon-induced gene expression and signaling in human hepatoma cell lines." J Hepatol **33**(5): 764-72.

Mirgorodskaya, O.A., Kozmin, Y.P., Titov, M.I., Korner, R., Sonksen, C.P. and Roepstorff, P. (2000). "Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using ^{18}O-labeled internal standards." Rapid Commun Mass Spectrom **14**(14): 1226-32.

Mujumdar, R.B., Ernst, L.A., Mujumdar, S.R., Lewis, C.J. and Waggoner, A.S. (1993). "Cyanine dye labeling reagents: sulfoindocyanine succinimidyl esters." Bioconjug Chem **4**(2): 105-11.

Mujumdar, R.B., Ernst, L.A., Mujumdar, S.R. and Waggoner, A.S. (1989). "Cyanine dye labeling reagents containing isothiocyanate groups." Cytometry **10**(1): 11-9.

Muthuswamy, S.K., Gilman, M. and Brugge, J.S. (1999). "Controlled dimerization of ErbB receptors provides evidence for differential signaling by homo- and heterodimers." Mol Cell Biol **19**(10): 6845-57.

Naaby-Hansen, S., Waterfield, M.D. and Cramer, R. (2001). "Proteomics-post-genomic cartography to understand gene function." Trends Pharmacol Sci **22**(7): 376-84.

Neuhoff, V., Arold, N., Taube, D. and Ehrhardt, W. (1988). "Improved staining of

proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250." Electrophoresis **9**(6): 255-62.

Neve, R.M., Holbro, T. and Hynes, N.E. (2002). "Distinct roles for phosphoinositide 3-kinase, mitogen-activated protein kinase and p38 MAPK in mediating cell cycle progression of breast cancer cells." Oncogene **21**(29): 4567-76.

Neville, D.C., Rozanas, C.R., Price, E.M., Gruis, D.B., Verkman, A.S. and Townsend, R.R. (1997). "Evidence for phosphorylation of serine 753 in CFTR using a novel metal- ion affinity resin and matrix-assisted laser desorption mass spectrometry." Protein Sci **6**(11): 2436-45.

Nicholson, R.I., Gee, J.M. and Harper, M.E. (2001). "EGFR and cancer prognosis." Eur J Cancer **37 Suppl 4**: S9-15.

Oberley, L.W. and Buettner, G.R. (1979). "Role of superoxide dismutase in cancer: a review." Cancer Res **39**(4): 1141-9.

O'Farrell, P.H. (1975). "High resolution two-dimensional electrophoresis of proteins." J Biol Chem **250**(10): 4007-21.

Olayioye, M.A., Badache, A., Daly, J.M. and Hynes, N.E. (2001). "An essential role for Src kinase in ErbB receptor signaling through the MAPK pathway." Exp Cell Res **267**(1): 81-7.

Olayioye, M.A., Graus-Porta, D., Beerli, R.R., Rohrer, J., Gay, B. and Hynes, N.E. (1998). "ErbB-1 and ErbB-2 acquire distinct signaling properties dependent upon their dimerization partner." Mol Cell Biol **18**(9): 5042-51.

Olayioye, M.A., Neve, R.M., Lane, H.A. and Hynes, N.E. (2000). "The ErbB signaling network: receptor heterodimerization in development and cancer." Embo J **19**(13): 3159-67.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A. and Mann, M. (2002). "Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics." Mol Cell Proteomics **1**(5): 376-86.

Ortiz, M.L., Calero, M., Fernandez Patron, C., Patron, C.F., Castellanos, L. and Mendez,

E. (1992). "Imidazole-SDS-Zn reverse staining of proteins in gels containing or not SDS and microsequence of individual unmodified electroblotted proteins." FEBS Lett **296**(3): 300-4.

Osada, H., Tatematsu, Y., Yatabe, Y., Nakagawa, T., Konishi, H., Harano, T., Tezel, E., Takada, M. and Takahashi, T. (2002). "Frequent and histological type-specific inactivation of 14-3-3sigma in human lung cancers." Oncogene **21**(15): 2418-24.

Ostergaard, M., Rasmussen, H.H., Nielsen, H.V., Vorum, H., Orntoft, T.F., Wolf, H. and Celis, J.E. (1997). "Proteome profiling of bladder squamous cell carcinomas: identification of markers that define their degree of differentiation." Cancer Res **57**(18): 4111-7.

Ostman, A. and Bohmer, F.D. (2001). "Regulation of receptor tyrosine kinase signaling by protein tyrosine phosphatases." Trends Cell Biol **11**(6): 258-66.

Otsuka, M., Kato, M., Yoshikawa, T., Chen, H., Brown, E.J., Masuho, Y., Omata, M. and Seki, N. (2001). "Differential expression of the L-plastin gene in human colorectal cancer progression and metastasis." Biochem Biophys Res Commun **289**(4): 876-81.

Overholser, J.P., Prewett, M.C., Hooper, A.T., Waksal, H.W. and Hicklin, D.J. (2000). "Epidermal growth factor receptor blockade by antibody IMC-C225 inhibits growth of a human pancreatic carcinoma xenograft in nude mice." Cancer **89**(1): 74-82.

Page, M.J., Amess, B., Townsend, R.R., Parekh, R., Herath, A., Brusten, L., Zvelebil, M.J., Stein, R.C., Waterfield, M.D., Davies, S.C. and O'Hare, M.J. (1999). "Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties." Proc Natl Acad Sci U S A **96**(22): 12589-94.

Park, T., Chen, Z.P. and Leavitt, J. (1994). "Activation of the leukocyte plastin gene occurs in most human cancer cells." Cancer Res **54**(7): 1775-81.

Patel, K., Stein, R., Benvenuti, S. and Zvelebil, M.J. (2002). "Combinatorial use of mRNA and two-dimensional electrophoresis expression data to choose relevant features for mass spectrometric identification." Proteomics **2**(10): 1464-73.

Patton, W.F. (2000). "A thousand points of light: the application of fluorescence detection technologies to two-dimensional gel electrophoresis and proteomics." Electrophoresis **21**(6): 1123-44.

Patton, W.F. (2002). "Detection technologies in proteome analysis." J Chromatogr B Analyt Technol Biomed Life Sci **771**(1-2): 3-31.

Pawson, T. (1995). "Protein modules and signalling networks." Nature **373**(6515): 573-80.

Peles, E. and Yarden, Y. (1993). "Neu and its ligands: from an oncogene to neural factors." Bioessays **15**(12): 815-24.

Perrot, M., Sagliocco, F., Mini, T., Monribot, C., Schneider, U., Shevchenko, A., Mann, M., Jeno, P. and Boucherie, H. (1999). "Two-dimensional gel protein database of *Saccharomyces cerevisiae* (update 1999)." Electrophoresis **20**(11): 2280-98.

Pina, B., Aragay, A.M., Suau, P. and Daban, J.R. (1985). "Fluorescent properties of histone-1-anilinonaphthalene 8-sulfonate complexes in the presence of denaturant agents: application to the rapid staining of histones in urea and Triton-urea-polyacrylamide gels." Anal Biochem **146**(2): 431-3.

Poltoratsky, V.P., Shi, X., York, J.D., Lieber, M.R. and Carter, T.H. (1995). "Human DNA-activated protein kinase (DNA-PK) is homologous to phosphatidylinositol kinases." J Immunol **155**(10): 4529-33.

Rabilloud, T. (1994). "Two-dimensional electrophoresis of basic proteins with equilibrium isoelectric focusing in carrier ampholyte-pH gradients." Electrophoresis **15**(2): 278-82.

Rabilloud, T. (1998). "Use of thiourea to increase the solubility of membrane proteins in two-dimensional electrophoresis." Electrophoresis **19**(5): 758-60.

Rabilloud, T. (2000). "Detecting proteins separated by 2-D gel electrophoresis." Anal Chem **72**(1): 48A-55A.

Rabilloud, T., Blisnick, T., Heller, M., Luche, S., Aebersold, R., Lunardi, J. and Braun-Breton, C. (1999). "Analysis of membrane proteins by two-dimensional electrophoresis: comparison of the proteins extracted from normal or *Plasmodium falciparum*-infected erythrocyte ghosts." Electrophoresis **20**(18): 3603-10.

Rabilloud, T., Valette, C. and Lawrence, J.J. (1994). "Sample application by in-gel rehydration improves the resolution of two-dimensional electrophoresis with immobilized pH gradients in the first dimension." Electrophoresis **15**(12): 1552-8.

Rabilloud, T., Vuillard, L., Gilly, C. and Lawrence, J.J. (1994). "Silver-staining of proteins in polyacrylamide gels: a general overview." Cell Mol Biol (Noisy-le-grand) **40**(1): 57-75.

Raineri, I., Huang, T.T., Epstein, C.J. and Epstein, L.B. (1996). "Antisense manganese superoxide dismutase mRNA inhibits the antiviral action of interferon-gamma and interferon-alpha." J Interferon Cytokine Res **16**(1): 61-8.

Riese, D.J., 2nd and Stern, D.F. (1998). "Specificity within the EGF family/ErbB receptor family signaling network." Bioessays **20**(1): 41-8.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Seraphin, B. (1999). "A generic protein purification method for protein complex characterization and proteome exploration." Nat Biotechnol **17**(10): 1030-2.

Roy, A.L., Du, H., Gregor, P.D., Novina, C.D., Martinez, E. and Roeder, R.G. (1997). "Cloning of an inr- and E-box-binding protein, TFII-I, that interacts physically and functionally with USF1." Embo J **16**(23): 7091-104.

Ruepp, S.U., Tonge, R.P., Shaw, J., Wallis, N. and Pognan, F. (2002). "Genomics and proteomics analysis of acetaminophen toxicity in mouse liver." Toxicol Sci **65**(1): 135-50.

Salazar, G. and Gonzalez, A. (2002). "Novel mechanism for regulation of epidermal growth factor receptor endocytosis revealed by protein kinase A inhibition." Mol Biol Cell **13**(5): 1677-93.

Sarto, C., Binz, P.A. and Mocarelli, P. (2000). "Heat shock proteins in human cancer." Electrophoresis **21**(6): 1218-26.

Scheele, G.A. (1975). "Two-dimensional gel analysis of soluble proteins. Charaterization of guinea pig exocrine pancreatic proteins." J Biol Chem **250**(14): 5375-85.

Scheler, C., Lamer, S., Pan, Z., Li, X.P., Salnikow, J. and Jungblut, P. (1998). "Peptide mass fingerprint sequence coverage from differently stained proteins on two-dimensional electrophoresis patterns by matrix assisted laser

desorption/ionization-mass spectrometry (MALDI-MS)." Electrophoresis **19**(6): 918-27.

Schirmer, E.C. and Gerace, L. (2002). "Organellar proteomics: the prizes and pitfalls of opening the nuclear envelope." Genome Biol **3**(4).

Schlessinger, J. (2000). "Cell signaling by receptor tyrosine kinases." Cell **103**(2): 211-25.

Seeger, M., Kraft, R., Ferrell, K., Bech-Otschir, D., Dumdey, R., Schade, R., Gordon, C., Naumann, M. and Dubiel, W. (1998). "A novel protein complex involved in signal transduction possessing similarities to 26S proteasome subunits." Faseb J **12**(6): 469-78.

Shevchenko, A., Wilm, M., Vorm, O. and Mann, M. (1996). "Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels." Anal Chem **68**(5): 850-8.

Simon, A., Fah, J., Haller, O. and Staeheli, P. (1991). "Interferon-regulated Mx genes are not responsive to interleukin-1, tumor necrosis factor, and other cytokines." J Virol **65**(2): 968-71.

Slamon, D.J., Clark, G.M., Wong, S.G., Levin, W.J., Ullrich, A. and McGuire, W.L. (1987). "Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene." Science **235**(4785): 177-82.

Songyang, Z., Shoelson, S.E., Chaudhuri, M., Gish, G., Pawson, T., Haser, W.G., King, F., Roberts, T., Ratnofsky, S., Lechleider, R.J. *et al.* (1993). "SH2 domains recognize specific phosphopeptide sequences." Cell **72**(5): 767-78.

Sorkin, A., Di Fiore, P.P. and Carpenter, G. (1993). "The carboxyl terminus of epidermal growth factor receptor/erbB-2 chimerae is internalization impaired." Oncogene **8**(11): 3021-8.

Sorkin, A. and Waters, C.M. (1993). "Endocytosis of growth factor receptors." Bioessays **15**(6): 375-82.

Sorkina, T., Huang, F., Beguinot, L. and Sorkin, A. (2002). "Effect of Tyrosine Kinase Inhibitors on Clathrin-coated Pit Recruitment and Internalization of Epidermal Growth Factor Receptor." J Biol Chem **277**(30): 27433-41.

Stamps, A.C., Davies, S.C., Burman, J. and O'Hare, M.J. (1994). "Analysis of proviral integration in human mammary epithelial cell lines immortalized by retroviral infection with a temperature-sensitive SV40 T-antigen construct." Int J Cancer **57**(6): 865-74.

Stark, G.R., Kerr, I.M., Williams, B.R., Silverman, R.H. and Schreiber, R.D. (1998). "How cells respond to interferons." Annu Rev Biochem **67**: 227-64.

Steinberg, T.H., Chernokalskaya, E., Berggren, K., Lopez, M.F., Diwu, Z., Haugland, R.P. and Patton, W.F. (2000). "Ultrasensitive fluorescence protein detection in isoelectric focusing gels using a ruthenium metal chelate stain." Electrophoresis **21**(3): 486-96.

Steinberg, T.H., Jones, L.J., Haugland, R.P. and Singer, V.L. (1996). "SYPRO orange and SYPRO red protein gel stains: one-step fluorescent staining of denaturing gels for detection of nanogram levels of protein." Anal Biochem **239**(2): 223-37.

Steinberg, T.H., Lauber, W.M., Berggren, K., Kemper, C., Yue, S. and Patton, W.F. (2000). "Fluorescence detection of proteins in sodium dodecyl sulfate-polyacrylamide gels using environmentally benign, nonfixative, saline solution." Electrophoresis **21**(3): 497-508.

Swaminathan, S., Amerik, A.Y. and Hochstrasser, M. (1999). "The Doa4 deubiquitinating enzyme is required for ubiquitin homeostasis in yeast." Mol Biol Cell **10**(8): 2583-94.

Takihara, Y., Matsuda, Y. and Hara, J. (2000). "Role of the beta isoform of 14-3-3 proteins in cellular proliferation and oncogenic transformation." Carcinogenesis **21**(11): 2073-7.

Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y. and Yoshida, T. (1988). "Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry." Rapid Commun Mass Spectrom **2**(8): 151-153.

Tezel, E., Hibi, K., Nagasaka, T. and Nakao, A. (2000). "PGP9.5 as a prognostic factor in pancreatic cancer." Clin Cancer Res **6**(12): 4764-7.

Timms, J.F., White, S.L., O'Hare, M.J. and Waterfield, M.D. (2002). "Effects of ErbB-2 overexpression on mitogenic signalling and cell cycle progression in human

breast luminal epithelial cells." Oncogene **21**(43): 6573-86.

Tomsig, J.L., Snyder, S.L. and Creutz, C.E. (2003). "Identification of targets for calcium signalling through the copine family of proteins. Characterization of a coiled-coil copine-binding motif." J Biol Chem **8**: 8.

Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I. and Davison, M. (2001). "Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology." Proteomics **1**(3): 377-96.

Tossing, G. (2001). "New developments in interferon therapy." Eur J Med Res **6**(2): 47-65.

Tzahar, E., Pinkas-Kramarski, R., Moyer, J.D., Klapper, L.N., Alroy, I., Levkowitz, G., Shelly, M., Henis, S., Eisenstein, M., Ratzkin, B.J., Sela, M., Andrews, G.C. and Yarden, Y. (1997). "Bivalence of EGF-like ligands drives the ErbB signaling network." Embo J **16**(16): 4938-50.

Tzahar, E., Waterman, H., Chen, X., Levkowitz, G., Karunagaran, D., Lavi, S., Ratzkin, B.J. and Yarden, Y. (1996). "A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor." Mol Cell Biol **16**(10): 5276-87.

Ullrich, A., Coussens, L., Hayflick, J.S., Dull, T.J., Gray, A., Tam, A.W., Lee, J., Yarden, Y., Libermann, T.A., Schlessinger, J. and et al. (1984). "Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells." Nature **309**(5967): 418-25.

Unlu, M., Morgan, M.E. and Minden, J.S. (1997). "Difference gel electrophoresis: a single gel method for detecting changes in protein extracts." Electrophoresis **18**(11): 2071-7.

Urwin, V.E. and Jackson, P. (1993). "Two-dimensional polyacrylamide gel electrophoresis of proteins labeled with the fluorophore monobromobimane prior to first-dimensional isoelectric focusing: imaging of the fluorescent protein spot patterns using a cooled charge-coupled device." Anal Biochem **209**(1): 57-62.

Valverius, E.M., Bates, S.E., Stampfer, M.R., Clark, R., McCormick, F., Salomon, D.S.,

Lippman, M.E. and Dickson, R.B. (1989). "Transforming growth factor alpha production and epidermal growth factor receptor expression in normal and oncogene transformed human mammary epithelial cells." Mol Endocrinol **3**(1): 203-14.

Van Criekinge, W. and Beyaert, R. (1999). "Yeast-Two-Hybrid: State of the Art." Biological Procedures Online **2**(1): 38.

Vanhaesebroeck, B., Leevers, S.J., Panayotou, G. and Waterfield, M.D. (1997). "Phosphoinositide 3-kinases: a conserved family of signal transducers." Trends Biochem Sci **22**(7): 267-72.

VerBerkmoes, N.C., Bundy, J.L., Hauser, L., Asano, K.G., Razumovskaya, J., Larimer, F., Hettich, R.L. and Stephenson, J.L., Jr. (2002). "Integrating "top-down" and "bottom-up" mass spectrometric approaches for proteomic analysis of *Shewanella oneidensis*." J Proteome Res **1**(3): 239-52.

Vercoutter-Edouart, A.S., Lemoine, J., Le Bourhis, X., Louis, H., Boilly, B., Nurcombe, V., Revillion, F., Peyrat, J.P. and Hondermarck, H. (2001). "Proteomic analysis reveals that 14-3-3sigma is down-regulated in human breast cancer cells." Cancer Res **61**(1): 76-80.

von Wussow, P., Jakschies, D., Hochkeppel, H.K., Fibich, C., Penner, L. and Deicher, H. (1990). "The human intracellular Mx-homologous protein is specifically induced by type I interferons." Eur J Immunol **20**(9): 2015-9.

Wang, Z., Zhang, L., Yeung, T.K. and Chen, X. (1999). "Endocytosis deficiency of epidermal growth factor (EGF) receptor-ErbB2 heterodimers in response to EGF stimulation." Mol Biol Cell **10**(5): 1621-36.

Washburn, M.P., Ulaszek, R., Deciu, C., Schieltz, D.M. and Yates, J.R., 3rd (2002). "Analysis of quantitative proteomic data generated via multidimensional protein identification technology." Anal Chem **74**(7): 1650-7.

Washburn, M.P., Wolters, D. and Yates, J.R., 3rd (2001). "Large-scale analysis of the yeast proteome by multidimensional protein identification technology." Nat Biotechnol **19**(3): 242-7.

Waterman, H. and Yarden, Y. (2001). "Molecular mechanisms underlying endocytosis

and sorting of ErbB receptor tyrosine kinases." FEBS Lett **490**(3): 142-52.

Weggen, S., Preuss, U., Pietsch, T., Hilger, N., Klawitz, I., Scheidtmann, K.H., Wiestler, O.D. and Bayer, T.A. (2001). "Identification of amplified genes from SV40 large T antigen-induced rat PNET cell lines by subtractive cDNA analysis and radiation hybrid mapping." Oncogene **20**(16): 2023-31.

Weisiger, R.A. and Fridovich, I. (1973). "Superoxide dismutase. Organelle specificity." J Biol Chem **248**(10): 3582-92.

Weiss, A. and Schlessinger, J. (1998). "Switching signals on or off by receptor dimerization." Cell **94**(3): 277-80.

Westbrook, J.A., Yan, J.X., Wait, R. and Dunn, M.J. (2001). "A combined radiolabelling and silver staining technique for improved visualisation, localisation, and identification of proteins separated by two-dimensional gel electrophoresis." Proteomics **1**(3): 370-6.

Westbrook, J.A., Yan, J.X., Wait, R., Welson, S.Y. and Dunn, M.J. (2001). "Zooming-in on the proteome: very narrow-range immobilised pH gradients reveal more protein species and isoforms." Electrophoresis **22**(14): 2865-71.

Whitehouse, C.M., Dreyer, R.N., Yamashita, M. and Fenn, J.B. (1985). "Electrospray interface for liquid chromatographs and mass spectrometers." Anal Chem **57**(3): 675-9.

Wiley, H.S., Herbst, J.J., Walsh, B.J., Lauffenburger, D.A., Rosenfeld, M.G. and Gill, G.N. (1991). "The role of tyrosine kinase activity in endocytosis, compartmentation, and down-regulation of the epidermal growth factor receptor." J Biol Chem **266**(17): 11083-94.

Wilkins, M.R., Sanchez, J.C., Gooley, A.A., Appel, R.D., Humphery-Smith, I., Hochstrasser, D.F. and Williams, K.L. (1996). "Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it." Biotechnol Genet Eng Rev **13**: 19-50.

Wilkinson, K.D., Lee, K.M., Deshpande, S., Duerksen-Hughes, P., Boss, J.M. and Pohl, J. (1989). "The neuron-specific protein PGP 9.5 is a ubiquitin carboxyl-terminal hydrolase." Science **246**(4930): 670-3.

Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T. and Mann, M. (1996). "Femtomole sequencing of proteins from polyacrylamide gels by nano- electrospray mass spectrometry." Nature **379**(6564): 466-9.

Wong, G.H., Kaspar, R.L. and Vehar, G. (1996). "Tumor necrosis factor and lymphotoxin: protection against oxidative stress through induction of MnSOD." Exs **77**: 321-33.

Worthylake, R., Opresko, L.K. and Wiley, H.S. (1999). "ErbB-2 amplification inhibits down-regulation and induces constitutive activation of both ErbB-2 and epidermal growth factor receptors." J Biol Chem **274**(13): 8865-74.

Wulfkuhle, J.D., Sgroi, D.C., Krutzsch, H., McLean, K., McGarvey, K., Knowlton, M., Chen, S., Shu, H., Sahin, A., Kurek, R., Wallwiener, D., Merino, M.J., Petricoin, E.F., 3rd, Zhao, Y. and Steeg, P.S. (2002). "Proteomics of human breast ductal carcinoma in situ." Cancer Res **62**(22): 6740-9.

Yamazaki, T., Hibi, K., Takase, T., Tezel, E., Nakayama, H., Kasai, Y., Ito, K., Akiyama, S., Nagasaka, T. and Nakao, A. (2002). "PGP9.5 as a marker for invasive colorectal cancer." Clin Cancer Res **8**(1): 192-5.

Yan, J.X., Devenish, A.T., Wait, R., Stone, T., Lewis, S. and Fowler, S. (2002). "Fluorescence two-dimensional difference gel electrophoresis and mass spectrometry based proteomic analysis of Escherichia coli." Proteomics **2**(12): 1682-98.

Yan, J.X., Wait, R., Berkelman, T., Harry, R.A., Westbrook, J.A., Wheeler, C.H. and Dunn, M.J. (2000). "A modified silver staining protocol for visualization of proteins compatible with matrix-assisted laser desorption/ionization and electrospray ionization-mass spectrometry." Electrophoresis **21**(17): 3666-72.

Yanagida, M., Shimamoto, A., Nishikawa, K., Furuichi, Y., Isobe, T. and Takahashi, N. (2001). "Isolation and proteomic characterization of the major proteins of the nucleolin-binding ribonucleoprotein complexes." Proteomics **1**(11): 1390-404.

Yang, W. and Desiderio, S. (1997). "BAP-135, a target for Bruton's tyrosine kinase in response to B cell receptor engagement." Proc Natl Acad Sci U S A **94**(2): 604-9.

Yao, X., Freas, A., Ramirez, J., Demirev, P.A. and Fenselau, C. (2001). "Proteolytic 18O

labeling for comparative proteomics: model studies with two serotypes of adenovirus." Anal Chem **73**(13): 2836-42.

Yarden, Y. (2001). "The EGFR family and its ligands in human cancer. signalling mechanisms and therapeutic opportunities." Eur J Cancer **37 Suppl 4**: S3-8.

Yarden, Y. and Sliwkowski, M.X. (2001). "Untangling the ErbB signalling network." Nat Rev Mol Cell Biol **2**(2): 127-37.

Yates, J.R., 3rd, Speicher, S., Griffin, P.R. and Hunkapiller, T. (1993). "Peptide mass maps: a highly informative approach to protein identification." Anal Biochem **214**(2): 397-408.

You, Z., Ishimi, Y., Masai, H. and Hanaoka, F. (2002). "Roles of Mcm7 and Mcm4 subunits in the DNA helicase activity of the mouse Mcm4/6/7 complex." J Biol Chem **277**(45): 42471-9.

Zheng, J., Rudra-Ganguly, N., Powell, W.C. and Roy-Burman, P. (1999). "Suppression of prostate carcinoma cell invasion by expression of antisense L-plastin gene." Am J Pathol **155**(1): 115-22.

Zhou, G., Li, H., DeCamp, D., Chen, S., Shu, H., Gong, Y., Flaig, M., Gillespie, J.W., Hu, N., Taylor, P.R., Emmert-Buck, M.R., Liotta, L.A., Petricoin, E.F., 3rd and Zhao, Y. (2002). "2D differential in-gel electrophoresis for the identification of esophageal scans cell cancer-specific protein markers." Mol Cell Proteomics **1**(2): 117-24.

Zhou, P., Fernandes, N., Dodge, I.L., Reddi, A.L., Rao, N., Safran, H., DiPetrillo, T.A., Wazer, D.E., Band, V. and Band, H. (2003). "ErbB2 degradation mediated by the co-chaperone protein CHIP." J Biol Chem **6**: 6.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R.A., Gerstein, M. and Snyder, M. (2001). "Global analysis of protein activities using proteome chips." Science **293**(5537): 2101-5.

Zwick, E., Bange, J. and Ullrich, A. (2002). "Receptor tyrosine kinases as targets for anticancer drugs." Trends Mol Med **8**(1): 17-23.

APPENDIX

Evaluation of Two-dimensional Differential Gel Electrophoresis for Proteomic Expression Analysis of a Model Breast Cancer Cell System*

Severine Gharbi, Piers Gaffney, Alice Yang, Marketa J. Zvelebil, Rainer Cramer, Michael D. Waterfield, and John F. Timms‡

The technique of fluorescent two-dimensional (2D) difference gel electrophoresis for differential protein expression analysis has been evaluated using a model breast cancer cell system of ErbB-2 overexpression. Labeling of paired cell lysate samples with *N*-hydroxy succinimidyl ester-derivatives of fluorescent Cy3 and Cy5 dyes for separation on the same 2D gel enabled quantitative, sensitive, and reproducible differential expression analysis of the cell lines. SyproRuby staining was shown to be a highly sensitive and 2D difference gel electrophoresis-compatible method for post-electrophoretic visualization of proteins, which could then be picked and identified by matrix-assisted laser-desorption ionization mass spectroscopy. Indeed, from these experiments, we have identified multiple proteins that are likely to be involved in ErbB-2-mediated transformation. A triple dye labeling methodology was used to identify proteins differentially expressed in the cell system over a time course of growth factor stimulation. A Cy2-labeled pool of samples was used as a standard with all Cy3- and Cy5-labeled sample pairs to facilitate cross-gel quantitative analysis. DeCyder (Amersham Biosciences, Inc.) software was used to distinguish clear statistical differences in protein expression over time and between the cell lines. ***Molecular & Cellular Proteomics 1:91–98, 2002.***

The ability to determine statistically significant alterations in protein expression that correlate with disease or occur consequent to experimentally induced changes in cells is fundamental to the exploitation of proteomics. Previous studies from our laboratory have used two-dimensional (2D)[1] gel analysis of immunomagnetic affinity-sorted primary human breast cells to establish the protein expression profiles of luminal and myoepithelial cells (1). These differential expression studies were carried out using replicate 2D gels of each sample and post-staining with a fluorescent protein dye. The subsequent detailed curation and correlation of images was used to derive data sets for statistical analysis. This approach was attractive, because the fluorescent protein stain has a wide linear range of detection, thus giving improved quantification of both high and low abundance proteins. However, in these experiments, many pairs of gels were required to establish statistically significant differences in protein expression as each gel contains inherent experimental variations that limit image superimposition.

The introduction of fluorescent 2D differential gel electrophoresis (DIGE) by Unlu *et al.* (2) has now made it possible to detect and quantitate differences between experimental pairs of samples resolved on the same 2D gel. The basis of the technique is the use of two mass- and charge-matched *N*-hydroxy succinimidyl ester derivatives of the fluorescent cyanine dyes Cy3 and Cy5, which possess distinct excitation and emission spectra. These are used to differentially label lysine residues of two protein samples for comparative analysis of the mixed sample on one gel. The ability to directly compare two samples on the same gel not only avoids the complications of gel-to-gel variation but also enables a more accurate and rapid analysis of differences and reduces the number of gels that need to be run. This procedure has been further developed by Amersham Biosciences, Inc. and has been evaluated recently in an *in vivo* mouse toxicology study (3).

Labeling reactions are carried out under conditions where proteins are "minimally" labeled, such that only 20% of molecules of a particular protein are covalently modified with one Cy dye molecule. Detection of proteins for excision and mass spectroscopy requires post-staining of gels with a general protein stain, because the unlabeled majority of a protein will not exactly co-migrate with the labeled protein, particularly in the low molecular weight range. A third fluorescent Cy dye (Cy2) has also been introduced, making it possible to com-

From the Ludwig Institute for Cancer Research, 91 Riding House Street, London W1W 7BS, United Kingdom

Received, September 20, 2001, and in revised form, November 12, 2001

Published, MCP Papers in Press, November 19, 2001, DOI 10.1074/mcp.T100007-MCP200

[1] The abbreviations used are: 2D, two-dimensional; DIGE, differential gel electrophoresis; MALDI, matrix-assisted laser desorption ionization; MS, mass spectroscopy; HRGβ1, heregulin β1; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonic acid; IPG, immobilized pH gradient; AmBic, ammonium bicarbonate; Hsp, heat shock protein.

pare three samples on one gel. An experimental design that should allow a much more accurate statistical analysis of expression differences across multiple gels can now be developed, because different Cy3- and Cy5-labeled samples can be compared with a Cy2-labeled "standard" run on every gel.

In this paper we describe further evaluation of the 2D-DIGE technique and seek to extend our protein expression studies in breast cancer. Specifically we investigated ErbB-2-mediated transformation in a model cell line system comprised of an immortalized luminal epithelial cell line and a derivative stably overexpressing ErbB-2 at a similar level to that seen in breast carcinomas (4). The ErbB-2 receptor tyrosine kinase (also known as neu/HER2) is overexpressed in 25–30% of breast cancers and is often associated with poor prognosis. Differentially expressed proteins detected in this cell system are likely to be involved in the processes of ErbB-2-mediated transformation. In the first experiments we evaluated the feasibility of the technique for monitoring protein expression changes in a model cell line system and tested its utility for high sensitivity, high throughput differential expression proteomics. We have examined the sensitivity and reproducibility of Cy3/Cy5 dye labeling and employed SyproRuby (Molecular Probes) gel staining to visualize proteins for automated spot picking. We have identified a number of differentially expressed proteins resulting from ErbB-2 overexpression using matrix-assisted laser-desorption ionization (MALDI) mass spectroscopy (MS). In a second set of experiments, we introduced Cy2 dye labeling of a standard pooled sample for linking gel images of pairs of samples from differentially treated cells. In this case we have employed proprietary software from Amersham Biosciences, Inc. (DeCyder) to collect, process, and derive statistical data in an experiment that examines the effect of a growth factor on the expression profile of a normal, mammary luminal epithelial cell line *versus* its ErbB-2-overexpressing derivative.

EXPERIMENTAL PROCEDURES

Cell Culture and Mitogen Stimulation—The parental HB4a cell line was established from flow-sorted normal human breast luminal epithelial cells by immortalization with a non-DNA binding, temperature-sensitive mutant of SV40 large T-antigen (U19tsA58) (5). The ErbB-2-overexpressing variant HBc3.6 clone was derived from HB4a by stable co-transfection with a full-length normal human ErbB-2 cDNA under control of the murine mammary tumor virus long terminal repeat promoter and an SV40 polyadenylation signal. The HBc3.6 clone was selected, because it overexpresses a level of ErbB-2 similar to that seen in many breast carcinomas of luminal cell origin (4).

Cells were routinely maintained in RPMI 1640 with 10% (v/v) fetal calf serum, 2 mM glutamine, 100 IU/ml penicillin, and 100 μg/ml streptomycin (all from Invitrogen) and 5 μg/ml hydrocortisone and 5 μg/ml insulin (both from Sigma) at 37 °C in a 10% CO_2 humidified incubator. Cells were starved for 48 h in insulin-free medium containing 0.1% fetal calf serum prior to harvesting. Starved cells were also stimulated with 1 nM (8 ng/ml) HRGβ1 (R & D Systems) for 4, 8, and 24 h. The concentration of HRGβ1 used was determined to be the minimum required to induce maximal extracellular signal-regulated kinase 1/2 and Akt phosphorylation after 10 min of stimulation as determined by immunoblotting (data not shown).

Sample Preparation and Protein Labeling—Cells at ~80% confluence were washed twice in 0.5× phosphate-buffered saline, lysed in lysis buffer (4% (w/v) CHAPS, 2 M thiourea, 8 M urea, 10 mM Tris-HCl, pH 8.5), and then homogenized by passing through a 25-gauge needle six times. Insoluble material was removed by centrifugation at 14,000 rpm for 20 min at 10 °C. Protein concentration was determined using the Coomassie protein assay reagent (Pierce).

Cell lysates were labeled with *N*-hydroxy succinimidyl ester-derivatives of the cyanine dyes Cy2, Cy3, and Cy5 (Amersham Biosciences, Inc.) following the protocol described previously (3). Typically, 100 μg of lysate was minimally labeled with 400 pmol of either Cy3 or Cy5 for comparison on the same 2D gel. Labeling reactions were performed on ice in the dark for 30 min and then quenched with a 50-fold molar excess of free lysine to dye for 10 min on ice. Differentially labeled samples were mixed and reduced with 65 mM dithiothreitol for 15 min. Ampholines/pharmalytes, pH 3–10 (1% (v/v) each; Amersham Biosciences, Inc.), and bromphenol blue were added, and the final volume was adjusted to 350 μl with lysis buffer. For the HRGβ1 stimulation experiments, lysates from three separate time course experiments were run in parallel. The triplicate sets of cells were serum-starved for 48 h and then stimulated for 4, 8, and 24 h with 1 nM HRGβ1 or left unstimulated (0 h). The 24 lysates generated were labeled with Cy3 (for HB4a) and Cy5 (for HBc3.6). A pool of all samples was also prepared and labeled with Cy2 to be used as a standard on all gels to aid image matching and cross-gel statistical analysis. The Cy3 and Cy5 labeling reactions (100 μg of each) from each time point were mixed and run on the same gels with an equal amount (100 μg) of Cy2-labeled standard. Thus, the triplicate samples and the standard were run on 12 gels (*i.e.* three gels with two cell lines from each of the four time points), to generate 36 images.

Protein Separation by 2D Gel Electrophoresis and Gel Imaging—Immobilized non-linear pH gradient (IPG) strips, pH 3–10 (Amersham Biosciences, Inc.), were rehydrated with Cy-labeled samples in the dark at room temperature overnight, according to the manufacturers guidelines. Isoelectric focusing was performed using a Multiphor II apparatus (Amersham Biosciences, Inc.) for a total of 80 kV-h at 20 °C, 10 mA. Strips were equilibrated for 15 min in 50 mM Tris-HCl, pH 8.8, 6 M urea, 30% (v/v) glycerol, 1% (w/v) SDS containing 65 mM dithiothreitol and then for 15 min in the same buffer containing 240 mM iodoacetamide. Equilibrated IPG strips were transferred onto 18 × 20-cm 9–16% gradient or 12% uniform polyacrylamide gels poured between low fluorescence glass plates. Gels were bonded to the inner plate using bind-saline solution (PlusOne) according to the manufacturer's protocol. Strips were overlaid with 0.5% (w/v) low melting point agarose in running buffer containing bromphenol blue. Gels were run in Protean II gel tanks (Bio-Rad) at 30 mA per gel at 10 °C until the dye front had run off the bottom of the gels.

2D gels were scanned directly between glass plates using a 2920 2DMaster Imager (Amersham Biosciences, Inc.). This charge-coupled device-based instrument possesses two six-position filter wheels (excitation and emission) enabling scanning at the different wavelengths specific for each of the Cy dyes and for SyproRuby fluorescent protein stain. An image is built up and converted to gray scale pixel values. Gel images were normalized by adjusting the exposure times according to the average pixel values observed. The images generated were exported as tagged image format (.tif) files for further protein profile analysis.

Post-staining, Image Analysis, and Spot Picking—Gels were fixed in 30% (v/v) methanol, 7.5% (v/v) acetic acid overnight and washed in water, and total protein was detected by post-staining with SyproRuby dye (Molecular Probes) for 3 h at room temperature. Excess dye

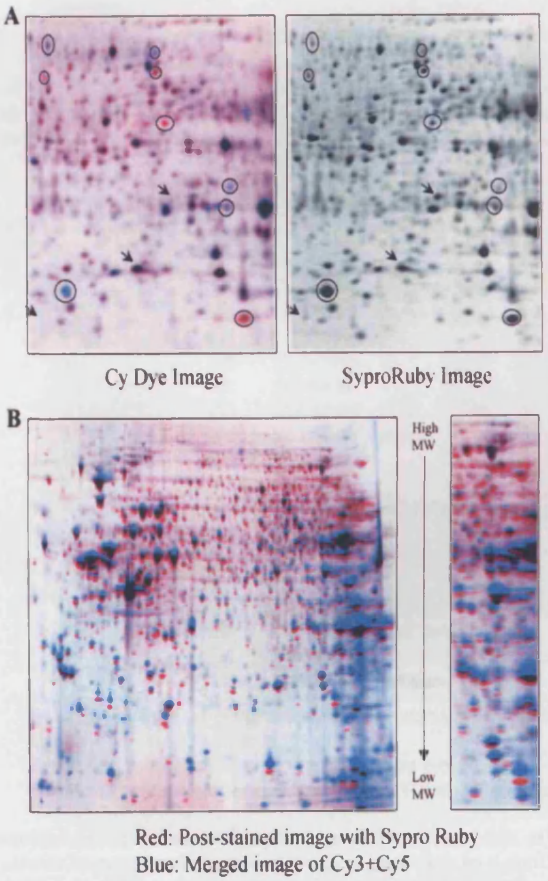


FIG. 1. **Sensitivity of 2D-DIGE and compatibility with SYPRO gel staining.** *A*, comparison of 2D-DIGE imaging and SyproRuby post-staining. Merged Cy dye image of HB4a lysate labeled with Cy3 (*red*) and HBc3.6 lysate labeled with Cy5 (*blue*) (*left panel*). The same gel was post-stained with SyproRuby dye (*right panel*). 100 μg of each lysate from serum-starved cells were analyzed on a 9–16% gradient gel. *Circles* represent differentially expressed proteins detectable by both methods. *Arrows* represent spots detected by SyproRuby but not Cy dye labeling. *B*, the shift in molecular weight between the modified and unmodified proteins was visualized by image overlaying. The DIGE image (*Blue*) was overlaid with the SYPRO image (*Red*).

was removed by washing twice in water, and gels were imaged using the 2920 2D Imager at the appropriate excitation and emission wavelengths for the stain. Gels were also post-stained superficially with silver according to the protocol of Shevchenko *et al.* (6). Images were curated and analyzed using Melanie III (Swiss Institute of Bioinformatics, Geneva, Switzerland) or with DeCyder software (on trial from Amersham Biosciences, Inc.). Differences were also detected visually by direct overlay of images using Adobe PhotoShop (Adobe Systems Incorporated).

For DeCyder image analysis, the differential in-gel analysis mode of DeCyder was first used to merge the Cy2, Cy3, and Cy5 images for each gel and to detect spot boundaries for the calculation of normalized spot volumes/protein abundance. At this stage, features resulting from non-protein sources (*e.g.* dust particles and scratches) were filtered out. The analysis was used to rapidly calculate abundance differences between samples run on the same gel. The biological

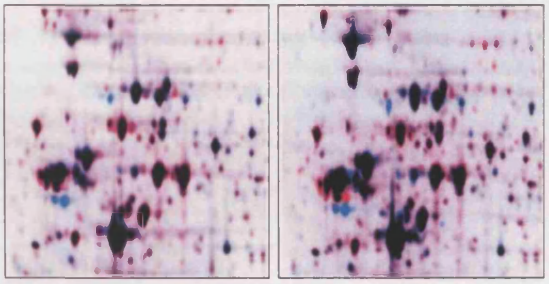


FIG. 2. **2D-DIGE is a reproducible detection method.** Duplicate samples of HB4a and HBc3.6 were labeled separately with Cy3 (*red*) and Cy5 (*blue*), respectively, and run on two separate gels. Images were generated using Adobe PhotoShop (Adobe Systems Incorporated).

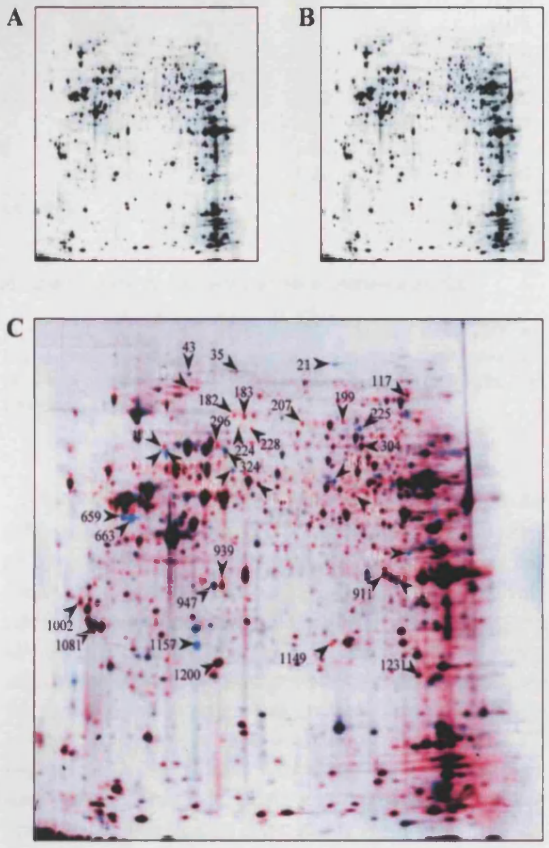


FIG. 3. **Cy dye images of protein expression differences between serum-starved cells.** *A*, HB4a image. *B*, HBc3.6 image. *C*, merged image of HB4a (*red*) and HBc3.6 (*blue*) lysates showing differentially expressed protein spots picked for identification by MALDI-MS and peptide mass fingerprinting. Not all indicated proteins have been identified (see Table I). The image shown is of a 200-μg sample (100 μg each of Cy3- and Cy5-labeled lysates) run on a pH 3–10 non-linear gradient IPG strip and 9–16% gradient polyacrylamide gel.

TABLE I

Summary of differentially expressed proteins in ErdB-2-overexpressing cells identified by MALDI-MS

The table shows identified proteins that were significantly up- or down-regulated in response to ErbB-2 overexpression. -Fold differences were calculated within the BVA mode of DeCyder using the average standardized abundance values from triplicate spots that had normalized spot volumes of >100,000. Zero time point values (*i.e.* no HRGβ1 treatment) were used for calculations except † and ‡, which were calculated from the 4- and 8-h time points, respectively. All differences are statistically relevant having *p* values of <0.05.

Average -fold difference (HBc3.6 *vs.* HB4a)[a]	Spot no. (Fig. 2C)[b]	Protein name	NCBI gene identifier	Predicted M_r	Predicted pI	% Coverage[c]	Match[d]
+18.6	1157	Hsp27	123571	22327	7.83	48	11/36
+13.9	21	Carbamoyl phosphate synthetase	13636759	131192	5.79	38	42/47
+8.5	663	Cytokeratin 17	4557701	48105	4.97	63	31/51
+3.9	332	L-plastin	11434009	70288	5.29	47	25/38
+3.3	466	Glutaminase C	5690372	65474	8.09	45	23/40
+2.0	407	Copine III	4503015	60131	5.60	40	22/63
+1.9	933	Aldose reductase	493797	35706	6.55	32	17/30
+1.4†	35	Leucine-rich protein	1730078	145202	5.50	37	42/72
+1.4	324	T-plastin	2506254	70436	5.52	48	27/68
+1.2	1081	14-3-3b	4507949	28082	4.76	85	32/64
−1.4	542	DNA helicase p50 (RuvB-like 1)	4506753	50228	6.02	45	17/32
−1.4	296	Hsp70	4204880	69995	5.56	49	36/72
−1.4	911	CLP-36	13994151	36071	6.56	65	18/56
−1.5	300	Lamin B	5031877	66408	5.11	58	36/54
−1.5	224	MxA (interferon-induced viral resistance	127566	75534	5.60	66	37/45
−1.6	182	Ku p70/80 subunit (acidic)	10863945	82705	5.55	54	43/63
−1.6‡	1231	Mitochondrial superoxide dismutase 2	11418405	24750	8.34	68	13/71
−1.7	183	Ku p70/80 subunit	10863945	82705	5.55	28	40/93

[a] Average -fold increase (+) or decrease (−) in expression in HBc3.6 *vs.* HB4a.

[b] Spot numbers refer to those in Fig. 2C.

[c] Amino acid sequence coverage for the identified protein.

[d] Number of peptide masses matching the top hit from MS-Fit peptide mass mapping *vs.* the total number of masses submitted.

variation analysis mode of DeCyder was then used to match all pairwise image comparisons from difference in-gel analysis for a comparative cross-gel statistical analysis. Operator intervention was required at this point to set landmarks on gels for more accurate cross-gel image superimposition. Comparison of normalized Cy3 and Cy5 spot volumes with the corresponding Cy2 standard spot volumes within each gel gave a standardized abundance. This value was compared across all gels for each matched spot, and a statistical analysis was performed using the triplicate values from each experimental condition.

Changes observed by 2D-DIGE analyses were aligned with Sypro-Ruby protein patterns, and spots were selected for picking according to this post-stained image. Spots of interest were excised from 2D gels using a Syprot automated spot picker (Amersham Biosciences, Inc.) following the manufacturer's instructions. Spots were collected in 200 μl of water in 96-well plates and kept frozen at −20 °C for protein identification by MALDI-MS.

Protein Identification by MALDI-MS—Gel pieces were washed twice in 25 mM ammonium bicarbonate (AmBic) in 50% acetonitrile and dried in a SpeedVac for 10 min. Samples were reduced in 10 mM dithiothreitol, 25 mM AmBic for 45 min at 50 °C and then alkylated in 50 mM iodoacetic acid, 25 mM AmBic for 1 h at room temperature in the dark. Gel pieces were then washed twice in 25 mM AmBic, 50% acetonitrile and vacuum-dried. Proteins were proteolysed with 30 ng of modified trypsin (Promega, Southampton, United Kingdom) in 25 mM AmBic for 16 h or overnight at 37 °C. Supernatant was collected, and peptides were further extracted in 5% trifluoroacetic acid, 50% acetonitrile. Peptide extracts were vacuum-dried and resuspended in 3 μl of water. Digests (0.5 μl) were spotted onto a MALDI target in 1 μl of matrix (2,5-dihydroxybenzoic acid). MALDI-MS was performed using a Reflex III reflector time-of-flight mass spectrometer (Bruker Daltonik, Bremen, Germany) in the reflector mode with delayed extraction. All mass spectra were internally calibrated with trypsin autolysis peaks. Peptide mass mapping was carried out using the MS-Fit program (Protein Prospector; University of California, San Francisco, CA).

RESULTS AND DISCUSSION

2D-DIGE Is a Sensitive and Reproducible Technique for Differential Expression Analysis—Initially, we sought to compare the sensitivity of protein detection using the *N*-hydroxy succinimidyl ester derivatives of the Cy3 and Cy5 dyes with post-staining methods that are compatible with subsequent MALDI-MS identification. In preliminary studies, varying amounts of a standard protein were labeled with Cy dyes, and the sensitivity of detection was compared with that of MS-compatible, superficial silver staining (6). This analysis revealed that Cy dye labeling was in fact more sensitive than silver staining; we could detect less than 1 ng of standard protein by Cy dye labeling compared with 5 ng with the silver staining protocol (data not shown).

Our planned high-throughput approach required bonding of gels to glass plates for automated spot picking. This was to prevent gel shrinkage upon fixation and movement of the gel during picking but also because fluorescent markers used for coordinating the picking process are mounted on the plate. We found that superficial silver staining was even less sensitive on the bonded gels and staining varied considerably from

gel to gel, thus making it unsuitable for 2D-DIGE post-staining. The fluorescent protein stain SyproRuby was much more sensitive than silver, consistently gave uniform staining from gel to gel, and its ability to detect proteins was unaffected by the process of bonding gels to glass plates. SyproRuby was also slightly more sensitive than Cy3 and Cy5 labeling, because a number of proteins detectable by SyproRuby were not easily visualized in the Cy3/5 images (Fig. 1*A*). Although this effect may be because of the enhanced sensitivity of SyproRuby staining, we cannot rule out the possibility that some proteins may not be modified by the Cy dyes as efficiently as others. Using the Melanie III software (7), we detected an average of 1.4 ± 0.1 times more gel features using SyproRuby *versus* Cy dye labeling. Thus, in practice, SyproRuby appears to be an ideal post-stain that is compatible with 2D-DIGE labeling.

Comparison of SyproRuby-stained images with Cy3 or Cy5 images from the same gel revealed that many spots did not exactly align. This was particularly noticeable in the lower molecular weight range of the gels (Fig. 1*B*). This is because proteins are minimally labeled, and the labeled protein has the additional mass of a covalently attached dye molecule (~580 Da). However, this shift was not uniform for all proteins, suggesting that dye labeling could alter the migration or that more than one lysine residue is labeled on that particular protein. For this reason, post-electrophoretic staining was required, and spots of interest were visually aligned prior to picking.

The preliminary Cy dye labeling experiments revealed obvious differences in protein expression between the parental and ErbB-2-overexpressing cell lines, with most of these proteins detectable by SyproRuby post-staining (Fig. 1*A*). For most of the spots of interest, we found that the volume of the spots from Cy dye images correlated with the amount of protein and hence the ability to identify them by MALDI-MS. Several low volume spots that were barely detectable by SyproRuby staining could not be identified by MALDI-MS, and it will be necessary to pick these spots from preparative gels on which more protein is loaded.

We next tested the reproducibility of Cy dye labeling. When the same sample (either an HB4a or an HBc3.6 urea-solubilized lysate) was labeled with Cy3 and Cy5, and the mixed labeled samples were run on the same gels, we detected only very minor differences in the abundance of some proteins (data not shown). These differences were more apparent for the very low abundance proteins. In a further experiment, replicate samples of HB4a and HBc3.6 lysates were labeled with Cy3 and Cy5, respectively, and run on different gels. We were able to detect the same differences in the expression of particular proteins from gel to gel (Fig. 2). These differences were also detectable when the sample-dye combinations were reversed (data not shown). Taken together, these results are in agreement with the conclusions drawn from the more rigorous validation of the technique carried out by Tonge *et al.* (3) and show that this technique is both sensitive and reproducible and can be used for the rapid identification of differences in the protein content of two separate cell lysate samples.

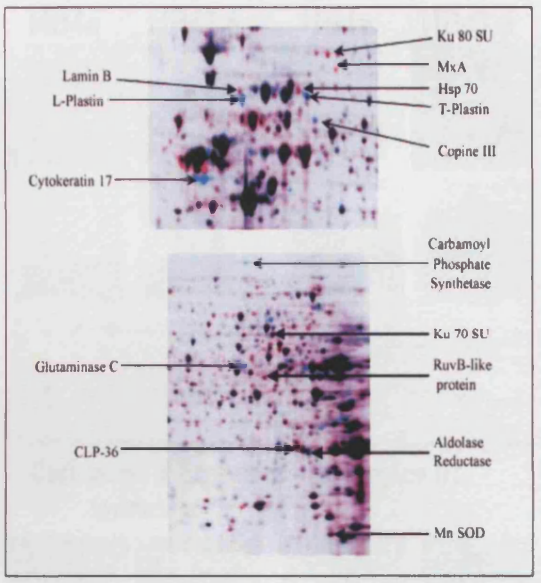


FIG. 4. **2D gel migration patterns of differentially expressed protein identified by MALDI-MS and peptide mass fingerprinting.** Enlarged regions of merged Cy dye images from samples run on pH 3–10 non-linear gradient IPG strips and 9–16% gradient polyacrylamide gels are shown. Identified proteins up-regulated in the HBc3.6 cell line are shown in *blue*, whereas those down-regulated in the HBc3.6 cells are shown in *red*. Refer to Table I for -fold changes in expression calculated using DeCyder software.

Identification of ErbB-2-mediated Protein Expression Changes in Serum-starved Mammary Luminal Epithelial Cells—To identify differences in protein expression resulting from ErbB-2 overexpression, we ran four sets of duplicate gels of HB4a and HBc3.6 lysates labeled with Cy3 and Cy5 in both combinations. We initially chose to compare the expression profiles of only the serum-starved cell pair to simplify experiments and to avoid the problem of sample contamination with bovine serum proteins. Under these conditions, the two cell lines show differences in morphology and proliferative capacity (4), suggesting differential protein expression occurs in the absence of serum. The protein image sets were curated and analyzed using Melanie III. Many of the differences were also detected visually by directly overlaying the Cy3 and Cy5 images using Adobe PhotoShop.

We detected 35 distinct protein spots that showed consistent differences in expression levels between the two cell lines, which were present in all gel images and were detectable by SyproRuby post-staining (Fig. 3). These spots were selected for automated spot picking, and 18 of them were identified with confidence by MALDI-MS and peptide mass fingerprinting (Table I). The migration patterns and differential labeling of the identified proteins are shown on the Cy dye images in Fig. 4. DeCyder software was used to calculate the average -fold

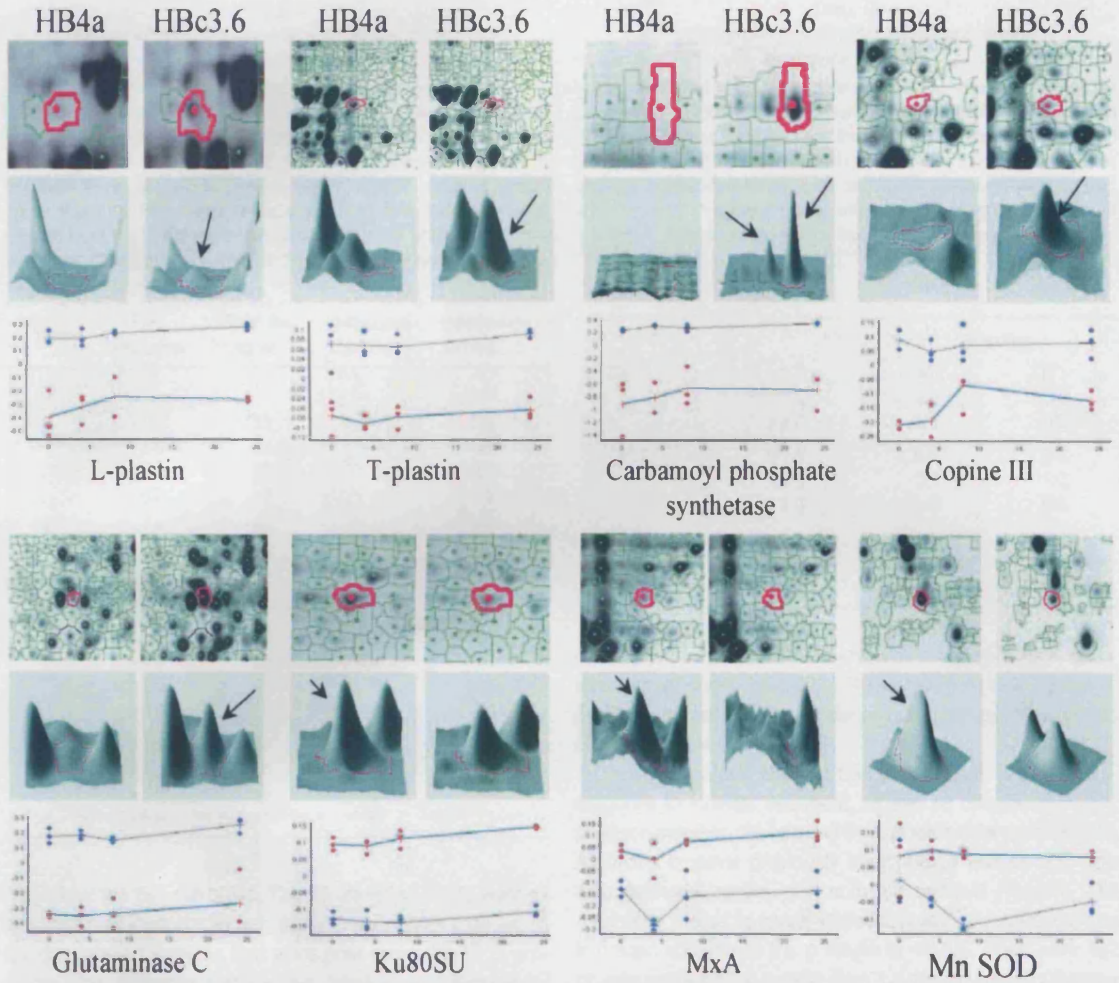


FIG. 5. **DeCyder output showing several of the identified, differentially expressed proteins.** Enlarged regions of images of Cy3-labeled HB4a (*left panels*) and Cy5-labeled HBc3.6 (*right panels*) lysates are shown, with the DeCyder-matched spots of interest highlighted. The three-dimensional fluorescence intensity profiles of the individual spots are shown in the *middle panels*. Graphical representations of all matched spots for a particular protein are shown in the *lower panels*, with changes in expression over time. Values are the standardized log of abundance (*i.e.* log abundance of Cy3- or Cy5-labeled spot over log abundance of Cy2-labeled standard spot). The three points at each time point on the graph represent single values from one gel, and lines are plotted using the averaged values.

difference in expression under serum-starved conditions between the two cell lines (see below). Six of these proteins showed a 2-fold or greater increase in expression in the ErbB-2-overexpressing cell line. Hsp27 showed the greatest increase in expression (Table I). This molecular chaperone protein is involved in various cellular stress responses, apoptosis and actin reorganization, but more significantly, it has been shown to be overexpressed in numerous human cancers, including invasive ductal carcinoma of the breast (reviewed in Ref. 8). Interestingly, Hsp70, another member of the heat shock protein family dysregulated in cancer (8), was down-regulated in the ErbB-2-overexpressing cell line.

Several of the other identified proteins have also been implicated in cellular transformation and metastasis. These are the actin-bundling protein L-plastin (9, 10), the glutamine-metabolizing enzyme glutaminase C (11), and the cell signaling modulator protein 14–3-3β (12). Thus, the up-regulation of these proteins appears to be involved in ErbB-2-mediated hyperproliferation and transformation of this cell type and perhaps confirms that this cell system is a valid cell culture model of breast cancer. We also speculate that increased expression of the metabolic enzymes carbamoyl-phosphate synthetase, glutaminase, and aldose reductase in the HBc3.6 cells is a direct consequence of their enhanced proliferation caused by ErbB-2 overexpression. Moreover, changes in the expression of the identified structural proteins (cytokeratin 17, L-plastin, T-plastin, leucine-rich protein, and lamin B) could account, at least in part, for the different morphologies dis-

TABLE II
Number of protein spots with a significant difference in abundance between cell lines

The total number of matched protein spots with a greater than 1.3-, 1.5-, 2.0-, and 5.0-fold difference in abundance between the cell lines is shown for each time point of HRGβ1 stimulation. The number of spots increased in the HB4a and HBc3.6 cell lines is also shown. Numbers were obtained by filtering data from 2607 matched protein spots within the BVA mode of DeCyder. -Fold differences were calculated from the average standardized abundance of triplicate spots with normalized spot volumes of >100,000. Only averages with *p* values of <0.05 were included.

Time	-Fold difference	Total No. spots	Increased in HB4a	Increased in HBc3.6
h				
T0	>1.3	135	34	101
	>1.5	81	14	67
	>2.0	23	3	20
	>5.0	2	0	2
T4	>1.3	160	53	107
	>1.5	74	22	52
	>2.0	16	2	14
	>5.0	2	0	2
T8	>1.3	159	56	103
	>1.5	68	24	44
	>2.0	20	3	17
	>5.0	3	0	3
T24	>1.3	165	44	121
	>1.5	75	18	57
	>2.0	16	3	13
	>5.0	1	0	1

played by the two cell types. The 18 identified spots were all cut from gels loaded with only analytical amounts (200 μg) of Cy dye-labeled samples that were post-stained with SyproRuby. This suggests that Cy dye labeling and SyproRuby post-staining *per se* do not affect the ability to identify proteins by MALDI-MS. Because mostly unlabeled protein is picked for MS identification using this minimal labeling methodology, it is unlikely that covalent Cy dye modification of lysine residues affects the generation of tryptic peptides. The remaining 17 unidentified proteins are either currently awaiting identification by MALDI-MS peptide mass fingerprinting, or their identities (six of them) cannot be given with statistical confidence because of the low abundance of recovered tryptic peptides. Identification of the low abundance proteins will require resolution of preparative quantities of samples. Ultimately, tandem MS will be applied for identification of differentially expressed protein spots that cannot be identified by MALDI-MS peptide mass fingerprinting.

Effects of HRGβ1 Treatment and ErbB-2 Overexpression on Protein Expression in Mammary Luminal Epithelial Cells: Statistical Analysis of Differential Expression—We further extended our studies to examine the effects of HRGβ1 treatment over time on the expression profiles of the parental and ErbB-2-overexpressing cells (see "Experimental Procedures"). We have recently observed differences in the responsiveness of these cells to HRGβ1,[2] which is a ligand of ErbB-3, the preferred heterodimerization partner of activated ErbB-2 (reviewed in Ref. 13).

TABLE III
Number of protein spots with a significant change in abundance over time

The number of matched protein spots down- and up-regulated in the HB4a and HBc3.6 cell lines over time following HRGβ1 stimulation are shown. Values were obtained by filtering data from 2607 matched protein spots. -Fold differences between each time point and time zero were calculated for each cell line from the average standardized abundance of triplicate spots with normalized spot volumes of >100,000. Only averages with *p* values of <0.05 were included.

Time	-Fold difference	HB4a	HBc3.6
h		*down/up*	
T0/T4	>1.3	5/0	6/1
	>1.5	2/0	0/0
	>2.0	1/0	0/0
T0/T8	>1.3	0/1	2/3
	>1.5	0/0	1/0
	>2.0	0/0	0/0
T0/T24	>1.3	3/6	2/11
	>1.5	3/4	2/4
	>2.0	0/0	0/1

DeCyder analysis showed trends of protein expression in response to HRGβ1 treatment, as well as differences in expression between the two cell lines at each time point (Fig. 5). All of the proteins previously identified by MALDI-MS were also detected by DeCyder software analysis (Table I). The data were filtered to reveal statistically relevant -fold changes in protein abundance (*i.e. p* values of <0.05). There were 135 protein spots with a greater than 1.3-fold average difference in abundance between the two cell lines at the zero time point (Table II). Of these 135 proteins, 34 were higher in the parental cell line (HB4a), whereas 101 were higher in the ErbB-2-overexpressing cell line (HBc3.6). At a 2-fold difference cutoff, there were 23 differentially regulated proteins. The number of significant differences between the cell lines was increased following HRGβ1 stimulation. Overall, the HBc3.6 cell line had a greater number of proteins with increased expression, perhaps correlating with its increased proliferation in response to HRGβ1 treatment. A comparison of the change in expression over time indicated that there were more proteins down-regulated than up-regulated at 4 h, whereas a higher number of proteins were up-regulated than down-regulated at 24 h in both cell lines (Table III).

The abundance of most of the identified proteins (Table I) did not change substantially with HRGβ1 treatment; however notable exceptions were copine III, interferon-induced viral

[2] John F. Timms, Sarah L. White, Michael J. O'Hare, and Michael D. Waterfield, submitted for publication.

resistance protein MxA, and mitochondrial superoxide dismutase 2 (Fig. 5). copine III expression increased in HB4a cells in response to growth factor, whereas its elevated expression remained unchanged in HBc3.6 cells. copine III is a phosphoprotein that possesses an associated kinase activity and exhibits Ca^{2+}-dependent phospholipid binding (14, 15), although its cellular function is unknown. MxA and mitochondrial superoxide dismutase 2 were repressed in the HBc3.6 cells, and their expression was further reduced by HRGβ1 treatment (Fig. 5), suggesting that ErbB-2 signaling negatively regulates their expression. It is imperative to follow up on our findings, using both biochemical and biological approaches to further assess the role of the identified proteins in ErbB-2-mediated transformation. We are also currently trying to identify the additional proteins whose expression is significantly altered by HRGβ1 treatment and ErbB-2 overexpression, as revealed by DeCyder analysis.

In conclusion, we show that 2D-DIGE and DeCyder image analysis is a sensitive, MS-compatible, and reproducible technique for identifying statistically significant differences in the protein expression profiles of multiple samples. This approach was more rapid than conventional analyses that compare multiple, post-electrophoretic stained gels and thus require more runs for statistical certainty. Using this methodology we have identified numerous proteins that are now implicated in ErbB-2-mediated transformation and may represent future targets for breast cancer therapies.

Acknowledgments—We gratefully acknowledge Drs. Sarah L. White and Carol Box for critical reading of the manuscript, Dr. Soren Naaby-Hansen and Steve Corless for technical advice, and Ian Currie and Alison Hillier of Amersham Biosciences, Inc. for DeCyder training and support.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

This manuscript is dedicated to the memory of Craig A. Brooks, who is greatly missed.

‡ To whom correspondence should be addressed. Tel.: 44-207-8784126; Fax: 44-207-8784040; E-mail: jtimms@ludwig.ucl.uk.

REFERENCES

1. Page, M. J., Amess, B., Townsend, R. R., Parekh, R., Herath, A., Brusten, L., Zvelebil, M. J., Stein, R. C., Waterfield, M. D., Davies, S. C., and O'Hare, M. J. (1999) Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 12589–12594
2. Unlu, M., Morgan, M. E., and Minden, J. S. (1997) Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* **18,** 2071–2077
3. Tonge, R., Shaw, J., Middleton, B., Rowlinson, R., Rayner, S., Young, J., Pognan, F., Hawkins, E., Currie, I., and Davison, M. (2001) Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* **1,** 377–396
4. Harris, R. A., Eichholtz, T. J., Hiles, I. D., Page, M. J., and O'Hare, M. J. (1999) New model of ErbB-2 over-expression in human mammary luminal epithelial cells. *Int. J. Cancer* **80,** 477–484
5. Stamps, A. C., Davies, S. C., Burman, J., and O'Hare, M. J. (1994) Analysis of proviral integration in human mammary epithelial cell lines immortalized by retroviral infection with a temperature-sensitive SV40 T-antigen construct. *Int. J. Cancer* **57,** p. 865–74
6. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68,** 850–858
7. Wilkins, M. R., Hochstrasser, D. F., Sanchez, J. C., Bairoch, A., and Appel R. D. (1996) Integrating two-dimensional gel databases using the Melanie II software. *Trends Biochem. Sci.* **21,** 496–497
8. Sarto, C., Binz, P. A., and Mocarelli, P. (2000) Heat shock proteins in human cancer. *Electrophoresis* **21,** 1218–1226
9. Lin, C. S., Aebersold, R. H., Kent, S. B., Varma, M., and Leavitt, J. (1988) Molecular cloning and characterization of plastin, a human leukocyte protein expressed in transformed human fibroblasts. *Mol. Cell. Biol.* **8,** 4659–4668
10. Lapillonne, A., Coue, O., Friederich, E., Nicolas, A., Del Maestro, L., Louvard, D., Robine, S., and Sastre-Garau, X. (2000) Expression patterns of L-plastin isoform in normal and carcinomatous breast tissues. *Anticancer Res.* **20,** 3177–3182
11. Lobo, C., Ruiz-Bellido, M. A., Aledo, J. C., Marquez, J., Nunez De Castro, I., and Alonso, F. J. (2000) Inhibition of glutaminase expression by antisense mRNA decreases growth and tumorigenicity of tumor cells. *Biochem. J.* **348,** 257–261
12. Takihara, Y., Matusa, Y., and Hara, J. (2000) Role of the beta isoform of 14-3-3 proteins in cellular proliferation and oncogenic transformation. *Carcinogenesis* **21,** 2073–2077
13. Yarden, Y., and Sliwkowski, M. X. (2000) Untangling the ErbB signaling network. *Nat Rev. Mol. Cell. Biol.* **2,** 127–137
14. Caudell, E. G., Caudell, J. J., Tang, C. H., Yu, T. K., Frederick, M. J., and Grimm, E. A. (2000) Characterization of human copine III as a phosphoprotein with associated kinase activity. *Biochemistry* **39,** 13034–13043
15. Creutz, C. E., Tomsig, J. L., Snyder, S. L., Gautier, M. C., Skouri, F., Beisson, J., and Cohen, J. (1998) The copines, a novel class of C2 domain-containing, calcium-dependent, phospholipid-binding proteins conserved from Paramecium to humans. *J. Biol. Chem.* **273,** 1393–1402