TGF-beta signaling in carcinogenic transformation of breast epithelial cells, using proteomics

Nimesh Bhaskaran

Stockholm 2008
To

Pappa, Mummy and Smikutty
Fate knows where you are going, but it is up to you to drive there.

Michelle Keesling

The dictionary is the only place where success comes before work.

Mark Twain
ABSTRACT

Transforming growth factor-β (TGFβ) is a key regulator of cell proliferation, death, migration and differentiation. One of the features of TGFβ signaling is to inhibit tumor growth at the early stage of cancer but promotes tumorigenesis at the advanced stage. Such a shift may be explained by differences in expression and activities of components of intracellular signaling.

Studies in this thesis have been focused on the mechanism of TGFβ signaling in carcinogenic transformation of human breast epithelial cells. Using comprehensive and unbiased proteomics technologies, we identified a number of novel targets of TGFβ in human breast epithelial cells (MCF10A). The main aim was to identify proteins involved in tumourigenic transformation of these cell lines. In a systemic and network analysis, we identified a number of novel TGFβ-dependent pathways, whose regulation is pivotal for TGFβ signaling in breast carcinogenesis. We focused on one such pathway (PLAG1-CNK1-RASSF1A) involved in TGFβ-dependent ERK1/2 activation. We observed that TGFβ1-induced regulation of these proteins was important for TGFβ1 dependent ERK1/2 activation and TGFβ1-dependent induction of Erk1/2 phosphorylation was detected predominantly at the endogenous levels of Plag1, CNK1 and RASSF1A. Interestingly, we found the scaffold role of CNK1 provided duality in its effect on Erk1/2 activation i.e. pro-zone effect, not previously known in TGFβ signaling. Further we were able to show that TGFβ1 inhibited cell growth under conditions when TGFβ1 induced Erk1/2 activation and any alteration in expression of these proteins could contribute to the switch of TGFβ from a tumor suppressor into a tumor promoter.

We were also interested to unravel the mechanism of cell transformation by exploring and comparing proteomes of MCF10A and 184A1 breast epithelial cells. MCF-10A and 184A1 lines are examples of transformed HMEC with different proliferation potential. Generation of protein expression maps, identification of functional domains and building proliferation related network lead us to a conclusion that components of mitogenic signaling such as PDGF, PI3K, Rac, Sos, MAP kinases are more represented in MCF10A cells as compared to 184A1. We further validated correlation of expression of endogenous cdk4, cyclinD3, cdc25B, p38γ with cell proliferation. Moreover, down-regulation of cdk4 and cyclin D3 with specific siRNA inhibited cell proliferation in human breast epithelial cells emphasizing their role in enhancement of cell proliferation.

Many important cellular decisions occur through post translational modifications (PTMs) of proteins. We also identified novel PTMs of Smad2 proteins in human breast carcinoma cell lines using mass spectrometry. Data from peptide mass finger printing (PMF) indicated that Smad2 can be acetylated, methylated, citrullinated, nitrated and palmitoylated. Further sequencing of selected peptides confirmed methylation at Gly122 and hydroxylation at Trp18 of Smad2. We were also able to observe novel modification previously not reported on Try128 and Try151. We therefore were able to show novel modifications on Smad2.
From our proteomics profiling study, we identified XRCC3 a DNA repair protein, whose expression was inhibited by TGFβ1. We found that XRCC3 is a negative regulator of TGFβ1-dependent transcriptional activation as observed by luciferase assay. Subsequently, we demonstrated that XRCC3 was found in a complex together with Smad3 on CAGA probe. Moreover the binding of XRCC3 on CAGA was specific, as no binding was observed on scrambled CAGA probe. Furthermore, XRCC3 was shown to induce/modulate endogenous expression of TGFβ1-responsive genes such as c-myc, ccnd1, ccne1, c-fos and cdkn1a which have an important role in cell proliferation and cell cycle. Overexpression of XRCC3 was also observed to increase cell proliferation considerably, as compared to control cells. We therefore establish a novel regulatory mechanism of TGFβ signaling. Thus, our investigations provided insights into molecular mechanisms of TGFβ signaling in human breast epithelial cells and in mechanisms of human breast tumorigenesis.
LIST OF PUBLICATIONS

I. **Nimesh Bhaskaran**, Serhiy Souchelnytskyi.
   Systemic analysis of TGFβ proteomics revealed involvement of Plag1/CNK1/RASSF1A/Src network in TGFβ 1-dependent activation of Erk1/2 and cell proliferation.
   *Proteomics, Vol.8, 2008*

II. **Nimesh Bhaskaran***, Kah Wai Lin***, Aude Gautier***, Hanna Woksepp, Ulf Hellman, Serhiy Souchelnytskyi.
   Comparative proteome profiling of MCF10A and 184A1 human breast epithelial cells emphasized involvement of CDK4 and Cyclin D3 in cell proliferation.
   *Proteomics Clinical Applications, Vol.1, Jan. 2009*

III. **Nimesh Bhaskaran**, Hiroyuki Iwahana, Jonas Bergquist, Ulf Hellman, Serhiy Souchelnytskyi.
   Novel post-translational modifications of Smad2 identified by mass spectrometry.
   *Central European Journal of Biology, 2008: (In press)*

IV. **Nimesh Bhaskaran**, Serhiy Souchelnytskyi.
   XRCC3 inhibits TGFβ signaling.
   *Submitted.*

* These authors contributed equally
Additional papers

# CONTENTS

## Introduction

1. Cancer 1
   1.1. Breast Cancer 1

2. TGFβ signaling 4
   2.1. Smad Proteins 7
   2.2. ERK1/2 MAP Kinase in TGFβ signaling 9
   2.3. Role of TGFβ in cancer 11
       2.3.1 TGFβ role in tumor suppression 12
       2.3.2 TGFβ role in tumor progression 14

3. Concept of Proteomics 16
   3.1. Proteomics technology /Methodology 17

4. Systems biology 21
   4.1. What is Systems biology? Why biologist needs to study what biology is? 22
   4.2. Systems biology and TGFβ signaling 23

## Aims of thesis 25

## Results and discussion 26

## General conclusions 35

## Acknowledgements 36

## References 39
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFT2</td>
<td>Iron-regulated transcriptional activator 2</td>
</tr>
<tr>
<td>ALK</td>
<td>Activin-receptors-like kinases</td>
</tr>
<tr>
<td>BEMAD</td>
<td>Beta-elimination followed by Micheals addition of DTT</td>
</tr>
<tr>
<td>BMP</td>
<td>Bone morphogenetic proteins</td>
</tr>
<tr>
<td>BRCA1</td>
<td>Breast Cancer 1 gene</td>
</tr>
<tr>
<td>CAF</td>
<td>Chemically assisted fragmentation</td>
</tr>
<tr>
<td>CellML</td>
<td>Cell Markup Language</td>
</tr>
<tr>
<td>CKI</td>
<td>Cyclin-dependent kinase inhibitor</td>
</tr>
<tr>
<td>CRM1</td>
<td>Exportin1, Chromosome region maintenance protein1</td>
</tr>
<tr>
<td>C-JUN</td>
<td>Avian sarcoma virus 17 oncogene homolog, Activator protein 1</td>
</tr>
<tr>
<td>CK2</td>
<td>Casein kinase 2, alpha-1</td>
</tr>
<tr>
<td>C-MYC</td>
<td>Avian Myelocytomatosis viral oncogene homolog</td>
</tr>
<tr>
<td>CSG</td>
<td>Coomasie stained gel</td>
</tr>
<tr>
<td>DAF7</td>
<td>Caenorhabditis elegans abnormal Dauer Formation family member</td>
</tr>
<tr>
<td>DAXX</td>
<td>Death-associated protein 6</td>
</tr>
<tr>
<td>DIGE</td>
<td>Differential gel electrophoresis</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiotrietol</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal growth factor</td>
</tr>
<tr>
<td>E2F4</td>
<td>E2F transcription factor 4</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial to mesenchymal transition</td>
</tr>
<tr>
<td>ER</td>
<td>Estrogen receptor</td>
</tr>
<tr>
<td>ERBB2</td>
<td>Avian erythroblast leukemia viral oncogene homolog 2</td>
</tr>
<tr>
<td>ESI</td>
<td>Electrospray ionization</td>
</tr>
<tr>
<td>FKBP12</td>
<td>FK506 binding protein 1A, 12kDa</td>
</tr>
<tr>
<td>FOXO</td>
<td>Forkhead box O</td>
</tr>
<tr>
<td>FTα</td>
<td>Farnesyl transferase alpha</td>
</tr>
<tr>
<td>FUSEBP</td>
<td>Far upstream binding protein</td>
</tr>
<tr>
<td>GDFs</td>
<td>Growth and differentiation factors</td>
</tr>
<tr>
<td>GO</td>
<td>Gene ontology</td>
</tr>
<tr>
<td>HDAC1</td>
<td>Histone deacetylase 1</td>
</tr>
<tr>
<td>IEF</td>
<td>Isoelectric focusing</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilized metal-affinity chromatography</td>
</tr>
<tr>
<td>IPG</td>
<td>Immobilized pH gradient gel</td>
</tr>
<tr>
<td>HMEC</td>
<td>Human mammary epithelial cells</td>
</tr>
<tr>
<td>LAP</td>
<td>Latency-associated protein</td>
</tr>
<tr>
<td>LOH</td>
<td>Loss of heterozygosity</td>
</tr>
<tr>
<td>LTBP</td>
<td>Latent TGF-β-binding protein</td>
</tr>
<tr>
<td>LUMIER</td>
<td>Luminescence-based mammalian interactome mapping</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>Interferon gamma</td>
</tr>
<tr>
<td>IPG</td>
<td>Immobilized pH gradient gel</td>
</tr>
<tr>
<td>MALDI</td>
<td>Matrix-assisted laser desorption/ionization</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinase</td>
</tr>
<tr>
<td>MAPKKK</td>
<td>Mitogen-activated protein kinase kinase kinase</td>
</tr>
<tr>
<td>MH1</td>
<td>Mad Homology domain 1</td>
</tr>
<tr>
<td>MH2</td>
<td>Mad Homology domain 2</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MIS</td>
<td>Muellerian inhibiting substance</td>
</tr>
<tr>
<td>MIZ-1</td>
<td>Myc-interacting zinc finger 1</td>
</tr>
<tr>
<td>MMP-2</td>
<td>Metalloproteases 2</td>
</tr>
<tr>
<td>MMP-9</td>
<td>Metalloproteases 9</td>
</tr>
<tr>
<td>NES</td>
<td>Nuclear export signal</td>
</tr>
<tr>
<td>NK</td>
<td>Natural killer cells</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localisation signal</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Name</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>OCLN</td>
<td>Occludin</td>
</tr>
<tr>
<td>PAR6</td>
<td>Partitioning defective 6 homolog</td>
</tr>
<tr>
<td>PEM</td>
<td>Polymorphic epithelial mucins</td>
</tr>
<tr>
<td>P38 MAPK</td>
<td>Mitogen-activated protein kinase 1</td>
</tr>
<tr>
<td>p53</td>
<td>Tumor protein P53, Transformation related protein 53</td>
</tr>
<tr>
<td>p15(INK4B)</td>
<td>Cyclin-dependent kinase inhibitor 2B</td>
</tr>
<tr>
<td>p16</td>
<td>Cyclin dependent kinase inhibitor 2A, Multiple tumor suppressor 1</td>
</tr>
<tr>
<td>p27(KIP1)</td>
<td>Cyclin-dependent kinase inhibitor 1B</td>
</tr>
<tr>
<td>p21(CIP1)</td>
<td>Cyclin-dependent kinase inhibitor 1A</td>
</tr>
<tr>
<td>p57</td>
<td>Cyclin dependent kinase inhibitor 1C</td>
</tr>
<tr>
<td>PI3K</td>
<td>Phosphoinositide 3-Kinase</td>
</tr>
<tr>
<td>PMBD</td>
<td>Pre-malignant breast disease</td>
</tr>
<tr>
<td>pRB</td>
<td>Retinoblastoma susceptibility gene product</td>
</tr>
<tr>
<td>PTHrP</td>
<td>Parathyroid hormone-releasing peptide</td>
</tr>
<tr>
<td>PTM</td>
<td>Post translational modification</td>
</tr>
<tr>
<td>PY</td>
<td>Proline-tyrosine motif</td>
</tr>
<tr>
<td>R-SMAD</td>
<td>Regulated SMAD</td>
</tr>
<tr>
<td>RTK</td>
<td>Receptor tyrosine kinase</td>
</tr>
<tr>
<td>RUNX3</td>
<td>Runt-related transcription factor 3</td>
</tr>
<tr>
<td>SA-β-gal</td>
<td>Senescence-associated- β-galactosidase</td>
</tr>
<tr>
<td>SAD</td>
<td>Smad activation domain</td>
</tr>
<tr>
<td>SARA</td>
<td>Smad anchor for receptor activation</td>
</tr>
<tr>
<td>SBML</td>
<td>XML-based Systems Biology Markup Language</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulphate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SILAC</td>
<td>Stable-isotope labeling of cells</td>
</tr>
<tr>
<td>SSG</td>
<td>Silver stained gel</td>
</tr>
<tr>
<td>SMAD</td>
<td>SMA- and MAD related protein 1, Mothers against decapentaplegic drosophila homolog</td>
</tr>
<tr>
<td>SMURF</td>
<td>Smad ubiquitination regulatory factors</td>
</tr>
<tr>
<td>SNX6</td>
<td>Sortin nexin 6</td>
</tr>
<tr>
<td>SP1</td>
<td>Specificity protein 1</td>
</tr>
<tr>
<td>TAB1</td>
<td>TGF-β-activated kinase-binding protein</td>
</tr>
<tr>
<td>TAK1</td>
<td>TGF-β-activated kinase</td>
</tr>
<tr>
<td>TβRI/ALK</td>
<td>TGF-β binding proteins/receptors type I</td>
</tr>
<tr>
<td>TβRII</td>
<td>TGF-β binding proteins/receptors type II</td>
</tr>
<tr>
<td>TFE3</td>
<td>Transcription factor binding to immunoglobulin heavy chain enhancer 3</td>
</tr>
<tr>
<td>TGF-β</td>
<td>Transformation growth factor- beta</td>
</tr>
<tr>
<td>TIE</td>
<td>TGF-β inhibitory element</td>
</tr>
<tr>
<td>TNF-α</td>
<td>Tumor necrosis factor alpha</td>
</tr>
<tr>
<td>TRF</td>
<td>Terminal restriction fragment</td>
</tr>
<tr>
<td>TRIP</td>
<td>TGF-β receptor-interacting protein</td>
</tr>
<tr>
<td>TSP</td>
<td>Thrombospondin</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular endothelial cell growth factor</td>
</tr>
<tr>
<td>VHL</td>
<td>Von Hippel-Lindau</td>
</tr>
<tr>
<td>WNT</td>
<td>Wingless-type MMTV integration site family</td>
</tr>
<tr>
<td>XIAP</td>
<td>X-linked inhibitor of apoptosis</td>
</tr>
<tr>
<td>Y2H</td>
<td>Yeast two-hybrid screening</td>
</tr>
<tr>
<td>ZO-1</td>
<td>Zona occludens 1</td>
</tr>
<tr>
<td>1D-PAGE</td>
<td>One-dimensional gel electrophoresis</td>
</tr>
<tr>
<td>2DE</td>
<td>Two-dimensional electrophoresis</td>
</tr>
<tr>
<td>2D-DIGE</td>
<td>Two-dimensional differential gel electrophoresis</td>
</tr>
</tbody>
</table>
1 CANCER

1.1 Breast cancer

Cancer is a generic term for a group of diseases that is characterized by uncontrolled growth and spread of abnormal cells that may affect almost any tissue of the body. Tumorigenic cells first form a primary tumor which may further acquire abilities to spread in the body which is known as metastases. Metastases are the major cause of death from cancer.

Despite a great success in the fight against cancer, from a total of 58 million deaths worldwide, cancer accounts for 7.9 million (or 13%) of all deaths, in 2007 and is estimated to increase in coming years\(^1\). Among different kinds of cancer, breast cancer is the most common malignancy affecting women. Comparing cancers, breast cancer alone accounts for about 20% of death in women. This emphasizes the importance of improvement in its diagnostics and treatment. Clinical and histological data in combination with biomarkers suggest that breast neoplasms can be differentially staged, from benign fibroadenomas to aggressive infiltrating carcinomas. The studies on the reported and elevated breast cancer biomarkers suggest that most of the tumors originate from epithelial cell. This underscores the need of a better understanding of the signaling events in the epithelial cells of human breast tissue and establishment of a standard model system to investigate the transformation of these cells.

Experimental models that explore human breast cancer progression will provide valuable insights for potential avenues of clinical intervention. In vitro and in vivo studies of rodent mammary carcinogenesis and human mammary cell cultures have provided support to build such model systems. Although these models have limitation when compared to the whole animal studies, they generate valuable data. Potential rate-limiting steps involved in early stages of breast cancer may be present in long lived animals such as human, that exhibit stringent replicative senescence and control of telomerase expression. Short lived animals such as rodents may lack these rate limiting steps; therefore rodents may be inappropriate model systems for studies of human breast disease. Thus, for the best modelling of human tumorigenesis, a use of human cells is preferential.

The first tumorigenetic changes of cells include immortalization, corruption of the cell cycle and evasion of cell death regulation (Fig.1). Various in vitro models of breast cancer progression have been utilized to examine the process immortal
transformation of normal human cells. The lack of spontaneous immortalization of normal human cells together with rarity of carcinogen-induced \textit{in vitro} transformation, has led to the predominant use of viral oncogene exposure for efficient immortal transformation. However, it is important to know that these viral oncogenes, capable of inactivating most of the cellular check point proteins simultaneously, are not associated with most of the human cancers, including breast cancer\textsuperscript{2}.

\begin{figure}[h!]
\centering
\includegraphics[width=\textwidth]{image.png}
\end{figure}

\textit{In vitro} transformed human mammary epithelial cells (HMEC) such as MCF10A and 184A1 have been used as model systems to investigate early breast carcinogenesis\textsuperscript{2}. The conversion model based on the immortal transformation of CDK inhibitor (p16) deficient finite lifespan HMEC suggests that in a normal lifespan of a cell there are two stages of senescence growth arrest. The first senescence occurs after \textasciitilde15-20 population doublings and is associated generally with high p16 levels, expression of senescence-associated- \( \beta \)-galatosidase (SA-\( \beta \)-gal) and a mean terminal restriction fragment (TRF) \textasciitilde6-8 kb. Post selection of these cells (with carcinogens) for next 20-70 population doublings, show no expression of p16 or SA-\( \beta \)-gal. This population then enters the second senescence growth arrest, as the mean TRF approaches \textasciitilde5 kb. The initial cell population that overcome the second senescence
block are called the “conditionally immortal”. These cell are characterised by p16 (-), SA-ß-gal (+), p57 (+), increase in telomerase activity and their mean TRF length decline to <3 kb. They exhibit a slow heterogeneous growth. As the mean TRF gets stabilized at <3 kb, the cells become fully immortal, characterised by high levels of telomerase activity, p57 (-) and uniform growth in absence and presence of TGFβ (Fig. 2).  

MCF-10A and 184A1 lines are the examples of such transformed HMEC. MCF-10A is fully immortalized cell lines whereas 184A1 lines are conditionally immortal. MCF-10A originated from “spontaneous” transformation of HMEC when grown in a medium with low calcium. Unlike the parental cell, the immortal lines were capable of growing in high calcium medium. MCF-10A express polymorphic epithelial mucins (PEM), keratin 14/8/18, and are ER negative 3, 4 They differ from normal HMEC in that they express telomerase and show loss of the CKI p15 (as well as p16) 2 (Fig. 2). Overexpression of c-erbB2 via retroviral transduction could transform MCF10A cells to strongly reduced mitogenic response (such as EGF and TGF-alpha ) and higher cloning efficiency in soft agar5.

184A1 lines were obtained after several extended life cultures of HMEC exposed to the chemical carcinogen benzo(a)pyrene 6. Extended life culture showed loss of p16 expression and stable form of p53 and lost proliferative capacity 7, 8. These cells express PEM, keratin 5/14 and 8/18 and are ER negative9. No defects were found in these immortal lines in expression or phosphorylation of pRB 7. Introduction of v-Ha-ras and SV40T could render 184A1 increased growth in methocel and increased tumors in nude mouse10. Both MCF10A and 184A1 cell lines retain growth factor and anchorage dependence and are nontumorigenic. MCF-10A and 184A1 cell lines were used as a model system for my investigations, as they represent early stages of carcinogenic transformation of human breast epithelial cells (Fig.2).
2 TGFβ signaling

Transforming growth factor beta is a secreted cytokine that regulates diverse biological responses including proliferation, differentiation, migration and apoptosis. The superfamily of TGFβ consists of more than 30 members which is divided based on sequences similarity and the specific signaling pathways they activate into two subfamilies; the TGFβ/Activin/Nodal subfamily and the bone morphogenetic proteins (BMP)/growth and differentiation factors (GDFs)/Muellerian inhibiting substance (MIS) subfamily (Fig. 3) \(^\text{11}\).
Figure 3. TGFβ-superfamily ligands, receptors and ligand traps. (reprinted from Cell, Vol 113, Yigong Shi and Joan Massague, Mechanism of TGFβ Signaling from Cell Membrane to the Nucleus, Pages No 685-700, Copyright (2003), with permission from Elsevier).

TGFβ is secreted by cells in a noncovalent complex with the latency-associated protein 12, which renders it inactive. In this complex both LAP and mature TGFβ are disulphide bonded dimers 13. The latent TGFβ complex often contains another protein, latent TGFβ-binding protein (LTBP) which are required for their extracellular localization 14, 15. TGFβ is activated in vivo by proteolytic cleavage of the LAP protein, exposure to low pH or upon interaction of the latent TGFβ with other proteins (e.g., thrombospondin (TSP) and αvβ6 integrin) 15-18 (Fig.4). Upon activation TGFβ is released from the latent complex as a dimer, which is stabilized by hydrophobic interactions and disulphide bonds. Each monomer comprises of several β strands interlocked by three conserved disulphide bonds forming a tight structure known as “cysteine knot” 19.

There are at least 5 types of TGFβ binding proteins/receptors 20. Type I (TβRI/ALK5) and type II (TβRII) TGFβ receptors were found to be essential for initiation of intracellular signaling. Type I and type II receptors are single transmembrane proteins which have serine/threonine kinases in their intracellular parts. In humans there are five type II and seven type I receptors 21 (Fig.4).
Ligand binding to TβRII receptor initiates TβRI/ TβRII complex formation and TβRII phosphorylates TβRI/ALK5 in the juxtamembrane region (GS domain) \(^{22}\). The activated TβRI/ALK5 canalizes the signal inside the cell through the phosphorylation of receptor regulated Smads. TGFβ/activin/nodal transduce signal transduction through TGFβ R-Smads such as Smad2 and Smad3, whereas BMP type I receptors mediates signaling through BMP R-Smads such as Smad1, Smad5 and Smad8 \(^{11,22,23}\) (Fig.5).

**Figure 4. Overview of TGFβ signaling.** The arrows indicate signal flow and are color coded: orange for ligand and receptor activation, gray for Smad and receptor inactivation, green for Smad activation and formation of a transcriptional complex, and blue for Smad nucleocytoplasmic shuttling. Phosphate groups and ubiquitin are represented by green and red circles, respectively (reprinted from *Cell*, Vol 113, Yigong Shi and Joan Massague, Mechanism of TGFβ Signaling from Cell Membrane to the Nucleus, Pages No 685-700, Copyright (2003), with permission from Elsevier).
2.1. Smad Proteins

Smad proteins can be divided into three subfamilies: the first comprises receptor activated Smads or R-Smads (Smad2 and Smad3 for TGFβ signaling), the second is common mediator Smad4 or Co-Smad, and the third includes the inhibitory Smads or I-Smad (Smad6 and Smad7). R-Smads and Co-Smads comprises of highly conserved amino terminal Mad homology 1 (MH1) and a carboxy terminal Mad homology 2 (MH2) domains which form globular structures and are linked by a divergent proline-rich region of variable length, which is usually referred to as linker region. With an exception of Smad2, the MH1 domain of both Smad3 and Smad4 account for their ability to bind directly to DNA. Crystal structure of Smad3 MH1 domain revealed that a conserved 11-residue β-hairpin loop is responsible for its DNA binding. The MH1 regions in R- and Co-Smads contain a nuclear localization signal-like (NLS-like) sequence which allows their recognition by the nuclear import machinery. Smad4 contains a nuclear export signal (NES) in the linker region which interacts with the nuclear exporter CRM1. MH1 domain also mediates protein-protein interaction with transcription factors, such as AFT-2, c-Jun, SP1 or TFE3.

Linker region of Smads contains several important regulatory motifs. There are multiple consensus phosphorylation sites for mitogen-activated protein kinases (MAPK) which upon phosphorylation causes sequestration of R-Smads in the cytoplasm by impeding nuclear translocation. Additionally, a proline-tyrosine (PY) motif present in R-Smads enables recognition by the WW domains of the E3 ubiquitin ligases, Smurf1 and Smurf2. Smad4 linker region also contains Smad activation domain (SAD) that is essential for transcriptional transactivation activity (Fig.4, 5).

The MH2 domain is required for the formation of homomeric and heteromeric complexes among R- and Co-Smads. This domain accounts also for Smad-receptor association. The specificity for distinct Smads is thereby determined by the L3 loop of the R-Smads and the L45 loop of the type I receptor as well as the interaction of the phosphorylated GS motif with the Smad sequence downstream from the L3 loop. The interaction of the L45 loop of the type I receptor and L3 loop of the R-Smads plays an important role in initial receptor-Smad selection but the interaction between phosphorylated GS motif with the Smad sequence downstream from the L3 loop stabilizes the receptor-Smad interaction. As with the L45 loop, only a few amino acids in the L3 loop define receptor-binding specificity. The L3 sequences are invariant between Smad2 and Smad3 proteins. MH2 domain of R- and Co-Smads can also
recruit transcriptional co-activators (such as CBP/p300, PCAF, ARC105) and co-repressors (such as c-Myc, c-Ski/SnoN, Ev1-1) \(^{22,35}\).


TβRI/ALK5 phosphorylates R-Smad at the C-terminal SSXS-motif which leads to their activation. The activated R-Smad forms a heteromeric complex with Co-Smad4. These Smad complexes then translocate to the nucleus where they control gene expression in a cell-type-specific and ligand dose-dependent manner through interactions with transcription factors, coactivators and corepressors \(^{11,36}\). Smad4 is the only Smad protein which mediates signaling by all TGFβ superfamily receptors.

I-Smads have a conserved MH2 domain but their N-terminal region show weak similarity to that of R- and Co-Smads \(^{37}\). The inhibitory Smads (Smad7 and Smad6) lack the SSXS-motif and hence don’t serve as a substrate for activated TβRI. But Smad7 stably interacts with the receptor, thereby interfering with the R-Smads for receptor association. As a result, phosphorylation of R-Smads and subsequent downstream signaling is blocked by Smad7. Thus, Smad7 acts as an efficient negative feedback loop \(^{38,39}\). I-Smads also associates through their PY motif with WW domains of the E3 ubiquitin ligases (Smurfs) \(^{32}\). Bridging of this complex to the activated receptor; leads to the induction of receptor degradation through proteosomal pathways.
I-Smads also interacts with phosphatases such as GADD34-PP1c which in turn leads to the type I receptor dephosphorylation and inactivation. Although Smad6 and Smad7 show functional similarities, it is observed that Smad6 is functionally conserved to BMP signaling and Smad7 shows broader action against both TGFβ and BMP signaling. This specificity is mainly determined by specific lysine residues in MH2 domain of I-Smads. Since I-Smads are known to block TGFβ signaling, they also affect TGFβ growth inhibitory response in epithelial cells. Thus they might act as tumor promoters, at least when overexpressed. This has been demonstrated for Smad7 in a mouse epidermal tumor model study and in human pancreatic cancers. On the other hand Smad7 also inhibits TGFβ dependent tumor progression. Moreover TGFβ induced apoptosis in PC-3U cells is prevented by inhibition of Smad7 expression by antisense mRNA, suggesting that Smad7 is required for TGFβ dependent apoptosis.

2.2. ERK1/2 MAP Kinase in TGFβ signaling

MAP kinases constitute a family of serine-threonine kinases that allow numerous extracellular signals reaching the cell surface to converge into the nucleus by activating transcription factors. MAP kinase pathways are evolutionarily conserved signaling modules and the prototypical MAP kinase pathway is the ERK1/2 pathway, which is activated as a result of various extracellular stimuli including cytokines, ultraviolet irradiations, cell-cell or cell-matrix contacts etc.

Basic principles that govern the activation of ERK1/2 are well established and understood. Ligand binding to the cell surface receptors lead to the activation of small GTPase Ras, which further recruits the MAP kinase kinase kinase Raf to the membrane for subsequent activation by phosphorylation. Activated Raf proteins then phosphorylate and activate the MAP kinase kinases MEK1/MEK2, which in turn activate the effector MAP kinases ERK1 and ERK2 by phosphorylation of their Thr and Tyr residues within their catalytic domain.

Erk1/2 in turn are able to modulate gene expression by phosphorylation/activation of transcription factors directly or by activating other protein kinases, which then activate proteins involved in gene expression. The most common among these factors is activating protein-1 (AP-1), a family of pleiotropic transcription factors comprised of homo and heterodimers of Fos, Jun and ATF family members. Erk-mediated pathways are mostly involved in proliferation and differentiation and are considered to be anti-apoptotic.
Even though the pathway governing the activation of ERK1/2 is well established, there has been some controversy in the TGFβ field with regard to the role of Ras/ERK1/2 pathways in mediating TGFβ responses.

Earlier studies have focused on the ability of TGFβ to block the regulation of Ras/ERK cascades which were induced by tyrosine kinase receptor signaling. Mulder et al. suggested that ERK1/2 activation by TGFβ is not only cell-type specific, but is also dependent upon the cell cycle phase. They claim that TGFβ treatment of cells results in a rapid activation of ERK1 (beginning as early as 5 mins) after addition of growth factors to proliferating cultures of epithelial cells, in contrast to previous report which indicated that TGFβ inhibited the activation of ERK1 in quiescent mink lung epithelial cells, released from TGFβ-induced G1 growth arrest by addition of serum and EGF. ERK1/2 activation by TGFβ in epithelial cells may be mediated by Ras signaling. Expression of dominant negative RasN17 completely abrogated the TGFβ-mediated activation of ERK1. Finlay and colleagues demonstrated a delayed activation of ERKs in response to TGFβ in primary human lung fibroblasts. In these cells, TGFβ stimulated ERK activity beginning at 2 h, with the peak activity at 16 h. Simeone et al. described a delayed activation of ERKs by TGFβ in pancreatic acinar cells, suggesting an indirect response requiring protein translation. This activation ERKs by TGFβ is dependent on Smad4, MEK1, and Ras and is also important in the regulation of DNA binding of the transcription factor AP-1.

Although TGFβ stimulates low level of ERK induction which is Ras dependent, the mechanism underlying this induction is still unclear. Despite its low level, ERK activation is required for TGFβ signaling. It was shown in HaCaT cells, that induction in p21Cip1 expression by TGFβ required constitutively active components of the ERK pathway, and inhibitors or dominant negative constructs for the ERK pathway significantly decrease p21Cip1 induction by TGFβ while having no effect on Smad activity. These findings suggest that ERK pathway may be an independent pathway from Smad signaling in TGFβ signaling.

Abnormal activation of Ras/ERK pathway has been shown to divert TGFβ signaling towards pro-oncogenic response. In kidney epithelial cells, activation of Raf/ERK pathway blocks TGFβ induced apoptosis while increasing invasiveness characteristic of malignancy and induction of EMT. Activation of ERK pathway by TGFβ is required for the induction of EMT in breast tumor cells, and the observation that MEK1/2 inhibitors blocked TGFβ-induced ERK activation suggests that MEK/ERK pathway is involved in the induction of EMT.
Recently it has been shown that activated TβRI/ALK5 initiate ERK pathway by directly phosphorylating and activating ShcA protein. TβRI/ALK5 directly phosphorylates ShcA on serine and tyrosine to induce low levels of ShcA/Grb2/Sos complex formation that correlate with low levels of Erk activation. ERK MAP kinase pathway has also been reported to modify signaling by phosphorylating receptor activated R-Smads. Erk has been shown to phosphorylate the linker regions of Smad1, Smad2 and Smad3, which results in blocking of nuclear accumulation of these activated Smads. Four sites for Smad1 have been identified (Ser187, Ser195, Ser206, Ser214). Smad2 has only one Erk site in this region (Thr220) and Smad3 has two (Thr178, Ser212) sites. It was proposed that such mechanism might explain why some cells which have hyperactive Ras mutations do not respond to TGFβ dependent growth inhibition. However, other reports have demonstrated efficient nuclear Smad translocation in Ras transformed cells. Smad2 was shown to be phosphorylated by Erk1 and it was demonstrated that this phosphorylation increases Smad2 half-life, Smad4 complex formation and nuclear translocation resulting in increase of the transcriptional activity. It was suggested that Erk phosphorylated Smad2 in response to HGF and EGF in its C-terminal SSXS motif, which are specifically targeted by TβRI. It was also reported that Erk phosphorylates Smad4 on Thr276, which leads to increase TGFβ dependent R-Smad/Smad4 complex formation, nuclear accumulation and transcriptional activity. Erks were also reported to phosphorylate Smad2 on Thr8, a site that overlaps with the calmodulin binding region. This phosphorylation was shown to increase half-life of Smad2 and it was also reported that calmodulin inhibits Smad2 phosphorylation by ERK1 and vice versa.

ERK MAP kinase has thus been reported to have tight integration in TGFβ signaling. Understanding the influence of ERK pathway in TGFβ signaling is of prime importance for the complete conception of TGFβ-induced response and also understanding the dual nature of TGFβ signaling in cancer.

2.3. Role of TGFβ in cancer

TGFβ has a dual role in human tumorigenesis. In healthy tissues of developing and adult organisms TGFβ is thought to acts as tumor suppressor but whereas in more advanced stages of malignant disease, TGFβ promotes tumor growth. During the early stages of breast cancer development, the transformed epithelial cells appear to still be responsive to TGFβ mediated growth arrest, and TGFβ acts as a tumor suppressor.
contrast, advanced breast cancers are mostly nonsensitive to TGFβ-mediated growth inhibition and produce large amounts of TGFβ (Fig. 6).

**Figure 6. Dual role of TGF- β in cancer. For explanations, see the text.**

### 2.3.1 TGFβ role in tumor suppression

The growth inhibitory effects of TGFβ primarily involve the induction of cell cycle arrest at the G1 phase. This is accomplished by multiple mechanisms, mostly through transcriptional regulation by Smads. One of the main targets of TGFβ is c-Myc which promotes cell proliferation and growth. Downregulation of c-Myc is one of the key events in TGFβ mediated cytostatic effect, which in most cases is defective in breast cancer cells.\(^70\) Both mRNA and protein levels of c-Myc are rapidly reduced in response to TGFβ.

Activated TGFβ signaling stimulates a repressor complex containing Smad3, Smad4, E2F4 (or E2F5) and the Rb family member p107 to bind to the TGFβ inhibitory element (TIE) on the c-Myc promoter and suppress its transcription. The MH2 domain of Smad3 facilitates the formation of this multimeric complex required for transcriptional repression of c-Myc promoter. E2F4 and p107 play a key role and are functionally required for TGFβ mediated repression of c-Myc.\(^71,72\).

TGFβ also leads to the downregulation of c-Myc through the degradation of far upstream binding protein (FUSEBP), a transcriptional activator of c-Myc through a tRNA synthetase cofactor p38. TGFβ induces the expression of p38 and promotes its translocation to nucleus. Binding of p38 to FUSEBP stimulates the ubiquitination and degradation of FUSEBP, leading to the downregulation of c-Myc.\(^73\).
TGFβ exerts its cytostatic effect also by the inhibition of G1 CDKs activity. CDKs and cyclins interact to form an active cell cycle promoting complex without which the cell cannot complete the transition from G1 to S phase. TGFβ has been associated with up-regulation of the cyclin dependent kinase inhibitors (CKI) p15\textsuperscript{Ink4B}, p21\textsuperscript{Cip1} and p2\textsuperscript{7Kip1} while c-Myc represses the expression of p15\textsuperscript{Ink4B} and p21\textsuperscript{Kip1}\textsuperscript{74, 75}. TGFβ treatment of HaCaT keratinocytes and CCL64 mink lung epithelial cells leads to the induction of p15\textsuperscript{Ink4B} which then binds to CDK4/6 and prevents their interaction with cyclin D\textsuperscript{76, 77}. Consequently, the CDK inhibitor p2\textsuperscript{7Kip1} is released from the cyclin D-CDK4 complex, and redistributes to bind to cyclin E-CDK2 complex to inhibit its activity. p21\textsuperscript{Ink4B} also inhibits cyclin E-CDK2 activity. The different subcellular localization of p21\textsuperscript{Ink4B} (cytoplasm) and p2\textsuperscript{7Kip1} (nucleus) have been suggested to facilitate p21\textsuperscript{Ink4B} and p2\textsuperscript{7Kip1} to coordinately inhibit CDK2 and CDK4/6 kinase activities. It has also been shown that TGFβ up-regulation of p21 is Smad3 dependent\textsuperscript{77}.

Myc-interacting zinc finger 1 (Miz-1) is a transcription activator for p15\textsuperscript{Ink4B} and p21\textsuperscript{Kip1}. By binding to Miz-1, c-Myc inhibits Miz-1 mediated transcription activation of p15\textsuperscript{Ink4B} and p21\textsuperscript{Kip1}\textsuperscript{78}. c-Myc also inhibits TGFβ dependent induction of p15\textsuperscript{Ink4B} and p21\textsuperscript{Kip1} by binding to Smad2 and Smad3 and repressing their growth inhibitory function\textsuperscript{74}. In addition p21\textsuperscript{Kip1} expression is also induced in response to TGFβ by transcription factors like FoxO and Runx3. FoxO interacts with Smad3 and Smad4 in response to TGFβ at p21\textsuperscript{Kip1} promoter and activates the expression of p21\textsuperscript{Kip1}\textsuperscript{79}. Similarly Runx3 interacts with Smad2, Smad3 and Smad4 and synergizes with Smads for TGFβ-dependent induction of p21\textsuperscript{Kip1}\textsuperscript{80}.

In epithelial cells and keratinocytes, induction of p15\textsuperscript{Ink4B} and p21\textsuperscript{Kip1} is crucial for cytostatic effect of TGFβ but this does not occur in hematopoietic cells. Studies have shown that up-regulation of CDK inhibitor p57 is essential for TGFβ-dependent inhibition of the cell cycle in hematopoietic cells. p57 mRNA and protein levels are both induced in response to TGFβ\textsuperscript{81}. Dephosphorylation and activation of CDK2 and CDK4/6 by CDC25A promotes cell cycle progression from G1 to S phase. Downregulation of CDC25A is mediated through various discrete mechanisms by TGFβ. Smad3 appears to aid CDC25A and Skp1-cullinβ-TrCp interaction and its subsequent ubiquitination and degradation\textsuperscript{82}. TGFβ signaling also activates p160ROCK which then translocates to nucleus and inhibits the phosphatase activity of CDC25A\textsuperscript{83}. Moreover, in keratinocytes there are indications that TGFβ may induce an E2F4-p130
to complex with histone deacetylase 1 (HDAC1) in order to bind to the repressor site in the CDC25A promoter, leading to the decrease in the expression of CDC25A.

TGFβ downregulates the expression of cyclin A, which is necessary for the S-phase entry. Smad3 plays important role in this downregulation. TGFβ has been shown to suppress Id proteins (Id1, Id2 and Id3). Id proteins inhibit cell cycle exit and differentiation. TGFβ activates AFT3, a transcriptional repressor, which then together with Smad3 and Smad4 represses the transcription of Id1 promoter. TGFβ also induces the expression of Mad which in turn binds to and inhibits Id2 promoter. TGFβ has been shown to inhibit CDK2/cyclin E complex by inhibiting CDK2 CAK activity in human HepG2 hepatocellular carcinoma cells. It was also observed that TGFβ can specifically downregulate CDK2 CAK activity without altering CDK4 CAK activity.

These data indicates that TGFβ may acts as a tumor suppressor by inhibiting proliferation of epithelial, neuronal and hematopoietic cells.

2.3.2 TGFβ role in tumor progression

In spite of the tumor suppressive activity of TGFβ, tumor cells often show increased production of TGFβ. Studies proving TGFβs role in enhancing cell migration, EMT (epithelial to mesenchymal transition), angiogenesis and immunosuppression provide strong evidence in positive role of TGFβ in cancer progression.

TGFβ was shown to be a major effector of breast tumor metastasis in vivo. MDA-MB-231 breast cancer cells normally metastasize to bone. Blocking of TGFβ signaling in these cells by overexpressing dominant negative TβRII mutant and subsequent intracardiac injection of these cells into immunodeficient mice decreased tumors with fewer osteoclast recruitment, produced less bone destruction at metastatic sites and prolonged host survival. In this model it was observed that TGFβ induced the expression and secretion of parathyroid hormone-releasing peptide (PTHrP), resulting in bone-reabsorbing osteoclasts.

TGFβ can induce EMT in cultures of normal and transformed breast epithelial cells, squamous carcinoma cells, ovarian adenosarcoma cells and melanoma cells. TGFβ induces downregulation of epithelial markers such as E-cadherin, ZO-1, vinculin and keratin, and induces the expression of mesenchymal markers such as vimentin and N-cadherin. The mechanism of the TGFβ-induced EMT is intriguingly complex. Ras, PI-3-K signaling and TGFβ-induced activation of Smad and RhoA signaling seem to have distinct roles in this process.
induced phenotypical transition in some cell lines is similar to that of transitions stimulated by the growth factors that act through tyrosine kinase receptors and Ras signaling in other cell systems. This synergy explains why the invasive growth of Ras- or Raf-transformed epithelial cells in vitro requires intact autocrine TGFβ signaling 92, 96.

Tumor angiogenesis is a crucial phenomenon for tumor growth and invasion. In human breast tumors, high levels of TGFβ mRNA associate with increased microvessel density, and elevated circulating levels of plasma TGFβ associate with high tumor burden and poor patient prognosis 97. Enhanced tumor angiogenesis in immunodeficient mice caused by overexpression of TGFβ in Chinese hamster ovary 67 cells where strongly reduced by local administration of neutralizing antibodies to TGFβ 98. Intraperitoneal injection of TGFβ antibodies reduced angiogenesis and tumorigenicity of a renal carcinoma cell line that lacked TβRII and was therefore not responsive to TGFβ in T-, NK- and B-cell deficient mice. These cells also lacked the von Hippel-Lindau (VHL) tumor suppressor gene, and reintroduction of this gene decreased TGFβ synthesis, suggesting that VHL regulates TGFβ levels and has a tumor suppressor effects through an anti-angiogenic mechanism involving TGFβ 99.

TGFβ-induced changes in the microenvironment provide favorable conditions for endothelial cell migration and capillary formation. TGFβ induces the expression of metalloproteases MMP-2 and MMP-9 and downregulation of the protease inhibitor TIMP in tumor cells which are required for angiogenesis 100. Moreover, TGFβ was also shown to induce the expression of angiogenesis inducing vascular endothelial cell growth factor (VEGF) 101.

Immunosuppressive activities of TGFβ also stimulate tumor progression. Transfection of TGFβ cDNA into tumor cell line enhanced tumorigenicity by allowing evasion of host T cell-mediated tumor cytotoxicity. Injection of neutralizing antibodies against TGFβ suppressed MDA-MB-231 breast tumor formation and lung metastasis in nude mice that lack T-cells, but not in nude-beige mice that lack both T- and NK-cells 102. This suggests that the effect of TGFβ in tumorigenesis in vivo involves NK cell activity. TGFβ also reduces the expression of Major histocompatibility complex (MHC) class II molecules in tumor cells, thus acting indirectly in immunomodulation 103. Thus these roles of TGFβ in tumor progression offer an attractive target for therapy.

Clinical investigation using TGFβ antagonists have already started as a strategic approach for controlling tumor progression and metastasis. Neutralizing antibodies to TGFβ like 2G7IgG2b, 16-3G1, 5-2G6, CAT-152/Trabio, CAT-192 have
shown to increase the potency of immune response in animals which were reduced by tumor cells. Antisense oligonucleotide based approach represented by AP-12009 and AP-11014 designed to inhibit TGFβ expression and therefore inhibit TGFβ mediated immunosuppression is already been implied for the successful treatment of gliosarcomas. Various small-molecule ATP-competitive compounds LY550410, LY580276 and SB-505124 are in clinical trials for cancer therapy. It is also shown that it is possible to inhibit TGFβ function by the use of extracellular (soluble) domain of TβRII (sTβRII). sTβRII significantly reduces TGFβ metastases.

These findings are a good indicator that TGFβ has a dual role as a tumor suppressor and tumor promoter, and is a strong candidate for therapeutic application for cancer treatment.

3 Concept of Proteomics

The term ‘proteome’ was first introduced in 1995, means the protein complement of the genome. According to the central dogma of molecular biology, genomic DNA transcribes into mRNA and the mRNA is translated into proteins, which are functional units in cells. While the genome is comparatively static, proteome is much more dynamic, as the protein content of a given cell will vary with respect to changes in the external environment, physiological state of the cell and diseases etc. While genomic and mRNA profiling technologies provide important information, there are limitations to the type of the data obtained at the DNA and mRNA levels, as these data lack to explain information about splice variants, co- and post-translational modification of the proteins. There are more than 200 PTMs which combined with alternative splicing can generate more than 1,000,000 protein entities. Not every protein and its isoforms are present in a cell at one time. Moreover, protein functions are also dependent upon their localization. Proteomics encompasses the global analysis of all the proteins expressed by cell, tissue or organism. In this field, it is critical to analyze qualitative and quantitative changes that occur as a function of disease, treatment or environment. Proteomics plays an important role in the rapid identification of the molecular events associated with appearance and development of various cancers. Recent progress of proteomic technologies allowed approaching these challenges, although with various degree of success.
3.1 Proteomics technology /Methodology

Components of proteomics technology comprise 1) separation of complex protein mixtures, 2) quantification of the separated components, 3) identification of proteins and their posttranslational modifications and 4) functional analysis of the identified proteins. There are various ways of achieving these aims, but two-dimensional gel electrophoresis is still the benchmark for large-scale separations of complex protein mixtures. 2DE involves the separation of proteins in the first dimension according to their charge, isoelectric focusing (IEF) and in the second dimension according to their relative molecular mass by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). This technique can provide high efficiency of protein separation. 2D electrophoresis in combination with mass spectrometric identification is the most consistent and successful technology to date. Even though one-dimensional gel electrophoresis (1D-PAGE) can separate proteins, degree of protein resolution obtained is relatively low and a single band from a 1D-PAGE will contain several proteins. Moreover many proteins may also have overlapping physiochemical properties which may further reduce the resolution of separation by 1D-PAGE. 1D-PAGE may be less appropriate for proteomics of biological biopsies or identification of biomarkers because of limited separation capacities. Since the introduction of immobilized pH gradients (IPGs), IEF had a significant impact on separation of complex mixtures over a wide range of pH. The carrier ampholytes are attached to the acrylamide gels and cast into the gels to form fixed pH gradients. Fixing the pH gradient prevents drifting in the gel and also improves reproducibility. The IPGs enable the formation of a stable pH gradient and is capable of focusing protein on the gels in both acidic and basic region. Using narrow-range IPG strips allowed proteins to be separated in narrow pI-range.

2-D PAGE not only resolves the proteins but staining these proteins also enables relative abundance of the proteins to be quantified. In 2D electrophoresis, the second dimension is most commonly the SDS-PAGE separation. Consequently, detection on the 2-D gels uses the method developed for SDS PAGE. But often some issues are important for 2-D gels detection which is not so important for 1-D electrophoresis. Apart from dynamic range of detection, the linearity, reproducibility between different staining experiments and variability between proteins are important parameters to be considered. Based on theses there are five different commonly applied general protein visualizing approaches: 1) organic dyes, 2) precipitation of salts, 3) metal ion reduction, 4) fluorescence dyes and 5) radioactive isotopes.
Silver staining was first in use back in 80’s 118. Silver staining techniques are based on saturating gels with silver ions, washing the excess of ions out of gel matrix and reducing the protein-bound silver ions to form metallic silver. Driving force for staining is the affinity of the proteins for cations, silver. However, much substance, like SDS, amino acid, chloride etc show high affinity for silver so they must be removed prior contacting proteins to silver ion. Silver staining process is a multi step process which can be divided into 4 main steps in common 1) fixation, to insolubilize the proteins and remove the interfering compounds present in the gels, 2) sensitization to increase subsequent image formation, 3) silver impregnation, 4) image development. Considering various fixation and sensitization process there are more than 100 different variants of silver staining. However, in general there are two main categories: alkaline and acidic silver stains depending on the condition used for silver impregnation. Alkaline methods work with a diamine complex of silver nitrate in a highly alkaline environment (ammonia or sodium hydroxide). The developer is formaldehyde in a dilute acidic solution of citric acid 119, 120. Acidic methods use silver nitrate in water (weakly acidic solution) for gel impregnation and developer is formaldehyde in an alkaline carbonate solution 120, 121. Sensitivity of silver staining is consistently in low nanogram range. In general, silver staining detects the proteins on the gel surface. This is also one reason why silver stained gels can be used for further analysis of silver stained proteins.

Protein expression from biological samples can also be compared by using differential gel electrophoresis (DIGE). Two-dimensional differential gel electrophoresis (2D-DIGE) has significantly improved the sensitivity and reproducibility of 2DE, permitting its use for high throughput research 122, 123. Quantitative and statistical analysis of visualized proteins is achieved through computerized analysis of 2-D gel digitized image. Imaging softwares are required to interpret complex data generated by 2-DE into valid biological information. A number of commercially available sophisticated imaging and computer programs such as Melanie-3, Z3, Image Master 2D Platinum, Nonlinear Dynamics SameSpots and Decyder are available for this purpose. The introduction of these programs allows fast data mining spot detection, normalization background correction and reporting and exporting of data. These programs have given a possibility of statistical and comparative quantification of protein features and thereby feasible for comparison of biological samples. Even though with significant progress in non-gel based techniques,
2-D PAGE will remain an essential technique for the characterization of proteomes for many more years to come.

Even though 2DE has been the most widely used techniques for protein complexes, it has certain limitations. Identification of 1000 to 3000 proteins in 2D gel electrophoresis is considered standard but the estimated number of the proteins in a given cell is roughly 50,000 to 30,000. Thus 2DE separation allows detection of only a part of cellular proteins.

Mass spectrometric identification of proteins by peptide mass fingerprinting (PMF) and de novo sequencing has shown to have high efficiency. Advance mass spectrometric instruments along with enhanced data processing search engines provide a possibility of protein profiling and PTM identification at a rate which is unattainable by traditional biochemical technique. Mass spectrometer determines the mass-to-charge ratio (m/z) of the peptide or protein ions generated by electrospray ionization or matrix-assisted laser desorption/ionization (MALDI) source. The sensitivity of the mass spectrometer allows the identification of proteins or peptides from nanogram quantities of samples. Development of powerful search engines, such as MASCOT, ProFound and Find Mode, enables peptide to be sequenced and PTM to be identified and characterized in a relatively fast way. Determination of PTMs by western-blots is a widespread and easy method for analysis of protein modification. However western-blots analysis relies mostly on prior knowledge of the kind of the modification and good quality antibodies. In contrast, MS can be used to detect known PTMs as well as novel modification.

Analysis of a chemical moiety on an amino acid modification using mass spectrometry is well suited since covalent addition leads to increase of the molecular mass of that residue. For example, the phosphorylation of a tyrosine residue increases its molecular mass by addition of an HPO$_3$ group. This allows the identification of the phospho-Tyr residue through the observation of a discrete mass increment (Fig.7). This approach can be applied for further identification of different PTMs like glycosylation, acetylation and methylation etc.
Mass spectrometric methods can be complemented by various chemical methods for the immobilization and alterations of the post translationally modified peptides to improve overall sensitivity of analysis. Phosphoprotein and phosphopeptide enrichment can be achieved by using immobilized metal-affinity chromatography (IMAC) where phosphate moiety are chelated on Fe(III) or Ga(III) metal ions. IMAC takes advantage of the affinity of chelated Fe(III) or Ga(III) ions towards the phosphate group of phosphopeptides. Recently TiO$_2$ columns were introduced for phosphopeptide enrichment prior to MS analysis. Introduction of affinity tags by $\beta$-elimination of phosphoric acid from pSer and pThr followed by Michael addition of affinity-groups, such as biotin, enable enrichment of phosphopeptides and phosphoproteins. $O$-glycosylated amino acid residues also undergo $\beta$-elimination
Phosphotyrosine peptides can also be captured on solid support using phosphoramidate chemistry where these peptides are coupled to specific resins. Several mass spectrometry (MS)-based quantitative proteomic methods have been developed such as isotope-coded affinity tag (ICAT), Stable-isotope labeling with amino acid in cell culture (SILAC) and isobaric tag for relative and absolute quantitation (iTRAQ) which allow accurate comparative quantification of protein expression levels and post-translation modification.

Chemical modification of peptides also enabled enhanced fragmentation and detection of ions. This is important for identification of modified amino acid residue by fragmentation of selected peptides. Chemical addition of a sulphonyl-ester (Chemically assisted fragmentation or CAF) on the N-terminal of the peptides allows the identification of predominantly ‘y’ ion series which permits improved sequencing of proteins. This approach consists of a blocking step, where lysine residues are converted to homoarginine using O-methylisourea-hydrogen sulphate (giving rise to a 42.012 Da mass addition) or 2-methoxy-4,5-dihydro-1H-imidazole compound (Lys Tag 4H, Agilent Technologies) (giving rise to a 68.037 Da mass addition). This blocking step protects the lysine side chains from subsequent sulfonation labeling. This reaction is specific for lysine ε–amino group and therefore does not react with N-terminal amino group. In the second step the CAF reagent (3-sulfopropionic acid NHS ester) reacts with the amino group of the peptide giving rise to 136 Da increase in mass. Peptides obtained from trypsin digestion show high efficiency of fragmentation with CAF reaction due to the presence of a strong basic amino acid (i.e. lysine or arginine) at the C-terminus.

Using these technologies it has become evident that proteins are much more complex and dynamic than was predicted. There are tens of thousands of proteins and potentially hundreds of thousands of relations between them as a result a global view on the proteome is hampered. To increase our knowledge in global organization of the proteomes, there is always a need for the integration of computational tools to reveal the molecular features of complex networks.

4. **Systems biology**

Growing rate of wide range of “omics” and development in high throughput technologies are generating a huge volume of complex data which has to be analyzed using various tools and interpreted to understand a biological function. These
approaches will allow better understanding in the outcomes of diseases and should help to improve diagnostics, design better treatments.

4.1. What is Systems biology? Why biologist needs to study what biology is?

Systems biology is an integration of high through-put biological studies to understand how biological systems function. To understand living cell we must study them as systems rather than individual molecules. Generation of networks from intracellular process has proved fruitful in understanding complex systems. 

Network models of the whole system can be developed by studying the relationship and interaction between various parts of biological system (interactions of gene, protein, and metabolic pathways, physiological systems etc). Theses models provide a clear representation of complicated relationship between large numbers of biological elements. A network is defined simply as a system of components/proteins (called as nodes) that interact with each other (represented as links or edges). Systems biology aims to understand biological behavior at the systems level in terms of mathematical and computational formalisms. This is a challenging field due to the complexity in the biological system.

As proteins are the key components of any living organism, they become the essential molecules for biological modeling. In-depth analyses of selected proteins generate valuable information about these proteins, but such analyses seldom provide quantitative information which can be suitable for modeling. For example, modeling requires detailed information about multiple species in the same experiment, which may not be available in studies of few selected proteins.

Proteomics is a high throughput technology which enables the measurement of multiple features of complex system at various levels from a single cell to the whole organism. These technologies can generate massive amount of information about protein expression, rate of synthesis and degradation, PTMs, enzymatic activities, localization and interacting partners etc. Some level of standardization is needed for the acquisition and management of such large- scale data in a format that can be used for modeling. Systems biology based standards such as Gene Ontology (GO) for describing gene function, XML-based Systems Biology Markup Language (SBML), Cell Markup Language (CellML) for describing biomolecular simulation provide such initiatives. As our biological knowledge is incomplete and is developing fast, there is always a need to evaluate how modeling tools will manage the data. Standards for describing biological functions should also be flexible.
There are several databases for protein interactions which are curated from the literature and continuously updated. These include DIP \(^{137-139}\), BIND \(^{140}\), MIPS \(^{141}\), REACTOME \(^{142}\), and Ingeuniy database (\url{http://www.ingenuity.com}).

**Figure 8.** Flowchart of the steps of biological modeling using proteomics data is shown here. For details, see the text. (Souchelnnytskyi, Bridging proteomics and systems biology: what are the roads to be traveled?, Proteomics 2005, Vol 5, Pages 4123-37. Copyright Wiley-VCH Verlag GmbH & Co.KGaA. Reproduced with permission.)

**1.4.2. Systems biology and TGFβ signaling**

In recent years there has been dramatic progress in the field of TGFβ which allows building and analysis of accurate model or “signaling network”. Systemic analysis of receptor tyrosine kinase (RTK) such as EGF, PDGF and Toll-like receptor has already provided successful insights into their signaling pathways which can be further used for designing drugs and predict disease outcomes \(^{143-145}\). Studies based on transcriptional networks and large scale identifications of protein-protein interactions have made attempts to delineate the TGFβ/Smad pathway.

The first study to use such techniques was based on yeast two-hybrid screening (Y2H) in *C.elegans, daf7* pathway. *daf-7* encodes a TGFβ homolog that signals through type I and type II receptor \(^{146}\). From these studies the authors were able to define a network containing 59 interacting proteins of DAF-7/ TGFβ signaling components. Further functional analysis by using RNAi technology identified 9 novel components in TGFβ/Smad pathway which includes ADP ribosylation factor.
Second study was also based on a similar Y2H approach on mammalian Smad signaling network\textsuperscript{147}. The investigators used 11 human Smad pathway-related proteins as baits to screen the human placenta cDNA library for interacting partners. 12 novel interacting proteins identified from these screens were again used as baits to screen human cDNA library for interacting clones. After the completion of the screen 755 protein-protein interactions were identified and a network containing 591 proteins was generated. 14 newly identified interactors were further functionally analysed by RNAi technique. Seven of the 14 proteins were found to be modulated by TGF\(\beta\) response. One of these partners (LAPTm5) was identified as lysosomal protein, and based on these and other data authors suggested a link between Smad signaling and endocytosis.

Recently protein pairs were assayed for interaction in human 293T cells based on a high-throughput approach called Luminescence-based mammalian interactome mapping (LUMIER)\textsuperscript{148}. A dominant attribute about this technique was that the interactions were examined in the presence and absence of exogenous TGF\(\beta\). Examining the screens guided a network containing 947 protein interactions was generated. Further cluster and functional analysis revealed strong links between TGF\(\beta\) signaling and two novel proteins Partitioning defective 6 homolog (PAR6) and Occludin (OCLN) in the induction of EMT.

Computational modelling approaches have already been applied to analyse Smad pathways signaling\textsuperscript{149}. Investigators used an Ordinary Differential Equation (ODEs) to build a computational model of the canonical Smad signaling pathway. This study represents an excellent example of systems modelling in biology. All these evidences from published data suggest a wide range of dynamic complexity in TGF\(\beta\) signaling. More detailed investigations have to be performed to understand the dual nature of TGF\(\beta\) in cancer progression.

Focus of my investigation is to understand the mechanism of TGF\(\beta\) mediated growth inhibitory response, and also to examine the potential oncogenic mechanism that relapse tumor suppression activity of TGF\(\beta\) in mammary epithelial cells.
AIMS OF THE THESIS

I have been pursuing following studies:

1) Investigate network mechanisms of TGFβ dependent regulation of various cellular processes by using proteomics and systems biology approach. To examine molecular connections between different TGFβ targets and unravel cross-talks with other oncogenic pathways.

2) To explore the mechanism of cell transformation by describing and comparing proteomes of MCF10A and 184A1 breast epithelial cells.

3) To identify novel PTMs of Smad proteins in breast carcinoma cell lines and to unravel the functional importance of these modifications using mass spectrometry.

4) To identify and explore in depth novel TGFβ targets and understand their role in carcinogenesis.

These findings should guide improved understanding of the molecular events involved in breast carcinogenesis and should help in better drug development.
RESULTS AND DISCUSSION

PAPER I

Systemic analysis of TGFβ proteomics revealed involvement of Plag1/CNK1/ RASSF1A/Src network in TGFβ 1-dependent activation of Erk1/2 and cell proliferation

Transforming growth factor-β (TGFβ) has been described as a potent anti-mitogenic factor for epithelial cell regulating cell cycle, cell death, migration and differentiation. The role TGFβ in breast cancer is ambiguous, as it is known to both inhibit and promote breast tumor development in later stages of cancer. In this study, we investigated the mechanism which defines diversity of TGFβ signaling.

To identify novel targets of TGFβ on the protein level in a comprehensive and unbiased way, we performed proteome profiling of MCF10A cells. Protein expression maps which were treated or not TGFβ1 were generated using two-dimensional (2D) gel electrophoresis. To assess the short and long term effects of TGFβ1, MCF10A cells were treated with TGFβ1 for 2, 8 and 24hrs. To evaluate the protein presence in the cells, 2-D gels were stained with silver and to explore the rate of protein turnover, cells were metabolically labeled with [35S] methionine and [35S] cysteine. Statistical studies were performed by computer-assisted gel image analysis and the proteins from the spots which changed their level of expression (i.e. volume of the protein spot) upon treatment with TGFβ1 were identified using MALDI TOF/TOF mass spectrometry. We identified 144 proteins in silver-stained gels and 73 35S-labeled proteins. Most of the proteins were novel targets of TGF-β signaling. Proteomics data was also validated by performing western blots for some of the identified proteins.

The identified TGFβ1-regulated proteins were then analyzed through systemic analysis to explain the diversity of TGFβ signaling. By performing functional clustering of the identified proteins we found that TGFβ1-affected proteins regulate a broad range of cellular activities, for example, motility, proliferation, differentiation, death and cell-cell communication. With dynamic clustering of proteins based on their expression we showed that 37 and 20 protein spots were down-regulated, and 34 and 22 protein spots were up-regulated upon TGFβ1 treatment in silver-stained and 35S-labeled gels, respectively. An oscillation pattern of changes with down-up-down was observed for
41 and 19, and up-down-up patterns for 32 and 12 protein spots in silver-stained and
\(^{35}\text{S}\)-labeled gels, respectively. Dynamic clustering of the identified proteins
representing a single functional class for example; cell proliferation, showed that 5 out
of pro-mitogenic proteins (MYCN, XRCC3, PAK1, TPD52L2 and CCT2) were
inhibited by TGFβ1 although with different time-frames, but 2 were up-regulated
(PLK1 and IMPDH2). Six proteins out of anti-proliferation stimuli (PLAG1, CTNNB1,
STMN1, STAT1, NME1 and PRDX1) were up-regulated. Thus, TGFβ initiates
multiple processes by triggering both positive and negative regulatory signals.

To unveil systemic properties of TGFβ signaling, we built a network which involves
TGFβ-regulated proteins. This network included also proteins and genes which are
known to interact physically and/or functionally with TGFβ-regulated proteins. An
analysis of the network of dependencies showed that it is a scale-free network and
TGFβ affected proteins in nuclei, cytoplasm, cell membrane and extracellular space. In
average, each component in the network is connected with 5 to 8 other components.
Species with highest number of connections were found to be p53, CTNBB1, pRB,
AKT1, EGFR, STAT1. We also observed that significant number of TGFβ1-regulated
proteins is also targeted by tyrosine-kinase receptor (TRK) initiated signaling, showing
a cross-talk between these two signaling pathways.

To confirm effectiveness of the network analysis, we explored TGFβ1 effect on MAP
kinase activation. We selected MAP kinase pathway, as it is the convergence point for
both growth inhibitor TGFβ and mitogenic factors, e.g. EGF, FGF, PDGF etc.
Activation of ERK1/2 was found to inhibit TGFβ/Smad-dependent signaling, as well as
being required for efficient TGFβ signaling \(^{38, 152, 153}\). Network analysis showed that
PLAG1 could regulate TGFβ dependent ERK1/2 activation. PLAG1 was described as
an oncogene which may have an impact on cell growth and apoptosis \(^{154-156}\). The
TGFβ/PLAG/ERK1/2 network includes CNK1, RASSFIA, Src, Ras and Raf proteins.
Plag1 is a zinc-finger protein which acts as a transcription factor for CNK1 \(^{157}\). We
explored the importance of these proteins in TGFβ1-dependent regulation of ERK1/2
activation. We observed that TGFβ1 stimulated the expression of Plag1 after 2hrs of
treatment of cells. CNK1 expression was enhanced after 8hrs and RASSF1A was
induced by TGFβ after 2hrs with peak of the expression after 8hrs. Phosphorylation of
the Src was enhanced already after 2h and was further increased after 8hrs and 24hrs
without affecting its expression level. We hence confirmed TGFβ1-regulated
expression of Plag1, CNK1, RASSF1A and Src phosphorylation.
For experimental interrogation of the model, overexpression (using corresponding overexpression vectors) and down regulation (using specific siRNAs) for Plag1, CNK1, RASSF1A and Src in various combinations were used. Totally 36 combinations were tested upon treatment of cell for 2hrs and 8 hrs. An impact of the manipulation was monitored by evaluation of ERK1/2 phosphorylation. We observed that the TGFβ1-dependent induction of Erk1/2 phosphorylation was detected predominantly at the endogenous levels of Plag1, CNK1 and RASSF1A. Effect of Plag1 on ERK1/2 activation was directly influenced by CNK1 which proves CNK1 as a target of Plag1 and a scaffold for Raf activation, and therefore validation of the model architecture. The scaffold role of CNK1 provides duality in its effect on Erk1/2 activation, i.e. prozone effect, where optimal responsiveness is observed only in a certain concentration range of a scaffold, and lower or higher levels of it would result in enhancement of the output signal and even loss of responsiveness. This indicated that Plag1 and CNK1 are required for maintaining the minimal level of basal Erk1/2 activity, as well as for the enhanced activation. We observed that manipulations of RASSF1A expression resulted in reversion of TGFβ1 into an inhibitor of Erk1/2 activation, when at the same time CNK1 endogenous expression was disturbed before addition of TGFβ1, i.e. when the levels of both RASSF1A and CNK1 were manipulated before addition of TGFβ1. We observed that Src alone enhanced Erk1/2 activation, as expected. Inhibition of Src kinase activity promoted TGFβ1 inhibitory effect even under conditions when basal Erk1/2 activation was observed, i.e. upon expression of Plag1 and CNK1. Measurement of cell proliferation showed that under conditions when TGFβ1 induced Erk1/2 activation, it also inhibited cell growth, e.g. for not-transfected cells. Under conditions when TGFβ1 inhibited Erk1/2 activation, we observed that it had no effect or stimulated cell growth. These results suggest a novel way of TGFβ-dependent regulation of ERK1/2 activation.

In this work we provide a comprehensive systemic analysis of TGFβ1 proteomics using MCF10A as cell model system and we identified a group of novel target proteins of this pathway that could explain dual specificity of TGFβ signaling in breast tumor development. We showed that Plag1 and CNK1 are required for maintaining the minimal level of basal Erk1/2 activity, which in turn is required for growth inhibitory activity of TGFβ and any alteration in expression of these proteins could contribute to the switch of TGFβ from a tumor suppressor into a tumor promoter.
Comparative proteome profiling of MCF10A and 184A1 human breast epithelial cells emphasized involvement of CDK4 and Cyclin D3 in cell proliferation.

Acquisition of the high proliferation rate is crucial for carcinogenic transformation of cells. Normal human mammary epithelial cells have a definite life span and do not undergo spontaneous immortalization where as during oncogenic transformation cells overcome their senescence check points which deliver them to multiply indefinitely- a phenomenon referred to as immortalization. The main aim of this study was to unravel the mechanism of cell immortalization by exploring and comparing proteomes of MCF10A and 184A1 breast epithelial cells. MCF-10A and 184A1 lines are examples of transformed HMEC with different proliferation potential. MCF-10A is fully immortalized cell lines and thus high proliferation rate whereas 184A1 lines are conditionally immortal cells with comparatively low proliferation rates. Thus MCF10A and 184A1 human breast epithelial cells are often used in explorations of early breast tumorigenesis. We generated protein expression maps (2D gels) for MCF10A cells and 184A1 cells. Using MALDI TOF/TOF mass spectrometry we performed systematic identification of all proteins detected in 2-D gels. 183 proteins in 184A1 cells and 318 proteins in MCF10A cells were identified. Identification of functional domains affected by the identified proteins was done by GoMiner software. Analysis of functional clusters of identified proteins showed significant similarity between 184A1 and MCF10A cells. Functional clustering indicated that more proteins representing cell cycle and cell proliferation functional class was identified in MCF10A cells in comparison with 184A1 cells. As we were interested in the difference between the proliferation rates of theses cell lines, we built proliferation-related network. Generation of the network concluded that components of mitogenic signaling such as PDGF, PI3K, Rac, Sos, MAP kinases are more represented in MCF10A cells as compared to 184A1. To validate this observation, we prepared proliferating and serum-starved non-proliferating 184A1 and MCF10A cells and detected proteins of interest by immunoblotting of total cell extracts with specific antibodies. We observed that CDK4, Cyclin D3, Cdc25B and p38γ were expressed at higher levels in proliferating cells (comparatively more in MCF10A than 184A1), and inhibition of proliferation by serum starvation led to decrease in protein expressions. To further confirm whether these proteins are involved in definition of the proliferation rate of cells, we checked the...
proliferation rate of these cells by down-regulating cdk4 and cyclin D3 expression with specific siRNA. Transfection of siRNAs in MCF10A and 184A1 cells resulted in inhibition of cell proliferation, as measured by incorporation of \([^3]H\)thymidine. These results suggest that maintaining of the high levels of CDK4 and Cyclin D3 in human breast epithelial cells is a significant contributing factor to the increased proliferation rate of cells. Thus, our results suggest that enhanced expression of cdk4 and cyclin D3 may be among the early events in the breast cancer progression.

In this study by generating proteome datasets of 184A1 and MCF10A we propose a comprehensive proteome profiling of human breast epithelial cells required for understanding of carcinogenesis.
Novel post-translational modifications of Smad2 identified by mass spectrometry

Every protein in a cell is modified and most of these modifications may account for a specific function of the protein. There are more than 200 known PTMs. Smad2 is a very important mediator of TGFβ signaling and plays a vital role in regulation of cell proliferation, differentiation and apoptosis. Apart from phosphorylation, acetylation and ubiquitination which are known to affect its activity, very little is known about other post translational modifications (PTMs) \(^{164-167}\).

The aim of this study was to identify novel PTMs of Smad proteins in breast carcinoma cell lines using mass spectrometry and to unravel the functional importance of these modifications.

MCF7 stable transfected cell lines expressing human Flag-Smad2 were generated. The genes were cloned in pMEP4 vector, as they could be under Cd\(^{2+}\)-inducible promoter. Cells expressing Flag-Smad2 and control cells were also treated with or without TGFβ. Smad proteins were detected by immunoprecipitation and SDS-PAGE. Smad protein identity was confirmed by in-gel digestion of the protein band with trypsin, and further analysis by mass spectrometry (as example). Peptides obtained was further analysed by mass spectrometer and protein sequence search engines (ProFound). Peptide mass fingerprinting (PMF) data from ProFound search indicated that that a significant number of detected Flag-Smad2 peptides were not matched to non-modified Smad2 peptides. We assumed that the presence of modifying groups changed molecular mass of a peptide, and therefore excluded a peptide from assigning it to the Smad2.

Potentially modified Smad2 peptides were identified by performing search with the FindMod tool (http://expasy.org/tools/findmod/) using PMF data. The FindMod tool identified 12 peptides in Flag-Smad2 mass list, which corresponded to peptides with PTMs. Some of the modifications which we identified were already observed such as acetylation at lysine \(^{164}\), phosphorylation at C-terminal peptide which contains receptor-phosphorylated serine residues \(^{165}\) and at the peptide containing PKC-phosphorylated serine110 \(^{168}\), thus confirming previous observations. Palmitoylation at Cys412 and Cys463, and geranylation at Cys463 was identified as novel PTMs. Possible modification such as methylation on Cys380, Lys43, Trp429 and Cys436, citrillination at Arg14, O-octanoylSer417 and S-nitrosylation at Cys381 was also observed. To further locate the modified amino acids, these unmatched peptides were
further sequenced using MALDI TOF/TOF. CAF modification of the peptides was also used for better detection of 'y' ions and improved sequencing required for the identifying peptide modifications. From the fragmentation of Smad2 peptides we were able to confirm hydroxylation of Trp18, and methylation on Gly122. We were also able to identify a 100 Da mass shift modification on Tyr128 and Tyr155 which was not previously reported. We assume that 100 Da mass shift on tyrosine is due to formation of a covalent link with S-cysteinyl (3'-(S-cysteinyl)-tyrosine), but this possibility has to be further explored. Thus we were able to indentify novel modifications on Smad2 proteins. Potential modification observed by us are located in the MH1 and MH2 domain of the Smad2 proteins which are involved in transcription activation and protein-protein interaction. These findings suggest that these modifications can play a role in TGFβ signaling and may be in transformation of cells.
XRCC3 inhibits TGFβ-dependent signaling

Ligand binding to the receptors leads to the phosphorylation and activation of R-Smads (Smad2 and Smad3) which then forms heteromeric complexes with Co-Smad (Smad4), and these complexes translocate into the nucleus, where they regulate transcription target genes. Regulation of transcription by Smad complexes require a number of binding partner proteins for review see 169.

From our previous studies we found that XRCC3 expression was downregulated by TGFβ1 after 2hrs and then increases back to the basal level after 8hrs in MCF10A. To confirm this observation we analysed the expression of XRCC3 upon TGFβ stimulation of cells. XRCC3 was found to be down regulated by TGFβ1 after 24hrs, whereas endogenous mRNA expression levels of XRCC3 was not altered upon TGFβ1 treatment, which was confirmed by Real Time PCR. These data suggests that expression of XRCC3 is affected by TGFβ1 post transcriptionally. As XRCC3 was previously known to form a complex at the DNA, we explored whether XRCC3 has a role in TGFβ1-induced transcription. We observed that co-transfection of full-length XRCC3 along with Smad3 resulted in a strong reduction of the CAGA reporter activation by TGFβ1 in luciferase assay. From these results we confirmed that XRCC3 downregulates TGFβ1-dependent transcriptional activation on CAGA promoter. We further went on to check the specificity of XRCC3 binding to CAGA promoter, and we observed that XRCC3 was specific to CAGA sequence, as it didn’t have any effect on transcriptional activity in p53-induced Gene-6 reporter in luciferase assay. As XRCC3 was found to inhibit Smad3-dependent transcriptional activation through CAGA promoter, we checked if XRCC3 forms a complex with Smad3 on CAGA probe by performing DNA precipitation assay. We observed that Smad3 and XRCC3 bind cooperatively with each other on the CAGA probe. To further analyse the specificity of interaction we performed DNA precipitation assay with scrambled CAGA probe, where Smad3 shows much lesser efficiency of binding 170. Smad3 showed a weak binding on scrambled CAGA probe and no affect of TGFβ1 stimulation was observed on the binding, where as XRCC3 didn’t show any binding on scrambled CAGA probe. Thus these results suggest that XRCC3 form a Smad3-dependent complex on CAGA probe. By performing EMSA using Escherichia coli-expressed GST fused full length XRCC3 and the MH2-deletion Smad3 (Smad3-ΔMH2), we were able to address that XRCC3 binding to CAGA probe is not direct.
Smad3-ΔMH2 bound directly to CAGA box containing probe. GST-fused XRCC3 failed to bind to CAGA box containing probe.

Since XRCC3 did not show any complex on CAGA probe, we tested if XRCC3 form a complex with Smad3 in vivo. By performing immunoprecipitation we observed that, XRCC3 was not found in a complex with Smad3. On the other hand XRCC3 was found to be in a complex with Smad4. And we also observed that the co-expression of Smad4 along with XRCC3 and Smad3 interfered with XRCC3 inhibition of Smad3-dependent transcriptional activation on CAGA promoter as confirmed by luciferase assay.

As XRCC3 was found to inhibit TGFβ1 transcriptional activity, we explored whether XRCC3 affects endogenous expression TGFβ1 responsive genes. We performed semi-quantitative RT-PCR analysis on a subset of genes by over expression of XRCC3. We found strongly induced expression of c-myc, ccnd1, ccne1, whereas the expression of c-fos and cdkn1a was downregulated.

Since XRCC3 showed a strong effect on the expression of the cell cycle genes, we analyzed XRCC3 affect on cell proliferation and apoptosis in MCF-7 cells. In vector-transfected control cells TGFβ1 inhibited proliferation of the cells as expected, whereas, overexpression XRCC3 increased cell proliferation considerably, as compared to control cells. XRCC3 overexpressing cell also didn’t show any TGFβ1-dependent growth inhibition. By performing apoptosis assay we observed that, overexpression of XRCC3 itself didn’t have an effect on apoptosis but was observed to bring down TGFβ1-dependent apoptosis. From these we conclude that XRCC3 has a growth-promoting effect and tends to bring down TGFβ1-dependent apoptosis in MCF7 cells.

In conclusion, we have identified here a XRCC3 as a novel target of the TGFβ pathway. We propose that XRCC3 which is known to be involved in DNA repair is a repressor of TGFβ activity. XRCC3 was shown to inhibit TGFβ1 dependent transcriptional activation. We also observed, XRCC3 binding on CAGA probe in Smad3 dependent way, even though further exploration is required for understanding the mechanism of the binding. We also observed a complex formation between Smad4 and XRCC3 in MCF7 cells and overexpression of Smad4 counteracted XRCC3 inhibition of Smad3-dependent transcription. We were also able to show that XRCC3 modulates endogenous expression of TGFβ1-responsive genes which have an important role in cell proliferation and cell cycle. Overexpression of XRCC3 in MCF7 cells has growth promoting effect. Thus, we delineated a XRCC3-dependent signaling pathway where TGFβ inhibited the expression of its own inhibitor.
GENERAL CONCLUSIONS

Based on the findings in this thesis, we can conclude that:

- Plag1 and CNK1 are required for maintaining the minimal level of basal Erk1/2 activity, which in turn is required for growth inhibitory activity of TGFβ and any alteration in expression of these proteins could contribute to the switch of TGFβ from a tumor suppressor into a tumor promoter.

- Enhanced expression of cdk4 and cyclin D3 may be among the early events in the breast cancer progression.

- Novel modifications on Smad2 proteins were identified and these modifications can play a role in TGFβ signaling and may be in transformation of cells.

- XRCC3 as a novel inhibitor of the TGFβ signaling pathway.
ACKNOWLEDGEMENTS

So finally we are here in the part of the thesis which most of you are interested in reading. I hope you have atleast gone through my abstract before coming to this page😊. Let see if I remember you…….

The work presented in my thesis was performed at the Ludwig Institute for Cancer Research and Karolinska Institutet, Department of Oncology and Pathology and Karolinska Biomics Center. I would like to thank many people who in one way or the other have inspired and contributed to my work and stay during these several years in Sweden.

Firstly, I acknowledge Serhiy Souchelnytskyi, for being my supervisor. I respect his endless enthusiasm, exceptional intelligence and educational talent. I am most grateful to you for encouraging me to think independently, choose my own projects and promote myself into an independent researcher.

Carl-Henrik Heldin, director of Uppsala branch of Ludwig, for your support, encouragement and always having faith in me. Thank you for all those wonderful discussions we had. I am indebted to you for standing by me in the most difficult times.

Jonas Bergquist, my co-supervisor during my PhD studies. Thank you for all your great comments and support. I appreciate all the discussions we had and also for inviting me for your Department meeting.

I am thankful to all current and former colleagues at Ludwig and KBC for the unforgettable work atmosphere. I acknowledge those people, who have contributed directly to my work and are co-authors on the papers, especially Kah Wai Lin and Aude Gautier. Even though you guys joined the group as my students, it turned out that I had to learn a lot from you. Lin thank you for all those helps to finish the experiments especially on Sundays. Thank you for the wonderful team work. Anna and Ihor, you guys definitely had been the backbone of the group. Thanks for the endless encouragement and creative discussion. Ihor I really enjoyed our “non scientific talks” (data not shown) which we used to have while walking back to the station from KI. Olena and Min for the wonderful and happy moments in the lab. Nazari, thank you for all your patience towards my stupid comments and tolerating the mess in the writing room and the lab. Sara, even though we are so different people, you have always been one of the my closest friend. Thanks to both of you for such a wonderful vacation in Egypt, I hope we would be able to do it once again. Maria Yakymovych thanks for your patience towards me and for being such a cheerful person.

All the group leaders at Ludwig for contributing to my work through several valuable discussions and supporting me during my good and bad time: Maréne Landström, Johan Ericsson (‘JoE’), Carina Hellberg, Aris (I hope I could be like you one day).
All the other members of Ludwig in Uppsala: Peter, we had some really great discussions and learned a lot from each other, especially during ‘aatha hai’. I hope we could get those times back; I owe you a lot and hope to return them sometime. Paolo Conrotto, I think we share some really great memories between us from Ludwig. We had some good time at Ulf’s summer place especially in the log house and blue berry fields. Piotr and Aga, I think you guys are the best couples I have seen. You guys have given me a lot of things to remember. I would never forget our trip to Duesseldorf. Michael and Marta thanks for all the help and support in all my good or bad decisions. Marta thank you for those wonderful advices, you always reminded of my mom at time☺. Christian and Maxi, though we met only a few times, I really enjoyed talking to you guys.

I would like to acknowledge other Ludwig corridor mates, Marcin and Kaska, Pacho (Namaste, for the wounderful canooing trip), Rosita, Markus Dahl (b..d©), Synisky, Anita, Sylvie (you are really good coach), Katerina, Maria E, Susann K, Olla (for those ping and pong games, which we never played), Maria S, Maite (for those lovely scientific and non scientific discussions and for including me in the electrophoresis “dream team”). Aino and Christer for great X’mas parties, Ulla, CPU Uffe, Eva hallin, Ingegärd.

Professor Ulf Hellman, my acknowledgement would never be complete without acknowledging you. You have been the person who has inspired me the most both scientifically and away from science. I have been always astonished by your vast technical knowledge and interest both in mass spectrometry and in bicycles. I am really sorry for not being good with the bikes. I would like to thank you for those amazing summer house vacations and blue berries, we still have them. Thank you!

All the great people at KBC and Karolinska Institutet, in particular:
Professor Rolf Lewensohn, Janne Lehtiö, Kristina Viktorsson, Ali Moshfegh, Åsa Wheelock (for all the help and good discussion and making my C.V. in a presentable format©), Johan Lengqvist, Hanna Eriksson, Maria Pernemalm, Jenny, Petra, Rao, Darawalee Wangsa (for those dull jokes and coffee time chat), Lukas Orre (for responding to all my stupid doubts©), Luigi, Birgitta, Susanne Becker (for listening to all my complaints), Sven (for a really nice lunch time), Patrik, Arnesto, Bettina, AnnSoffi, Anestis, the cleaning lady on the KBC 1st floor, the girls from the CCK shop, Dali and Betzy for making my stay at KBC really memorable one. Betzy, I should admit that I was really bad with the developing machine cleaning duty. Thank you for your patience.

My friends outside the lab, I am incredibly indebted to you all. Thank you for all the movies, dinners, late night talks, X’mas and New Year celebrations. I am sorry that I have not been able to be in touch with you guys lately, but I will make up for that. Döbelnsgatan boys, Vijay (thank you for those memorable good times, your “vishesh tipanies” and I will never forget the trip to Amsterdam with you and pandey and your swimming pool act), Jitu bhai, Kuljeet, Dharam , Tanmoy, Gaurav, Pandey (dada), Pratima and Prasanth© (for all great times and the time we spent in Amsterdam, we had our best time when staying at Flogsta).
Meena, Chandra and Neha (for all the great get together), Faizaan and Fatima, Noopur and Umashankar (for those lovely dinners and discussion), Pradeep and Madhuri, Nelson and PK singh/Uncle (for being a part of the family and for those wonderful Sunday lunches and Mutton curry), Vargese and Juli, Surajith and Brinda, Rajiv Sir and Renju chechi, Stephy,(for all the help and invaluable information regarding the discounts in Uppsala) Prakash, Mathu, Greta.

Uppsala Cricket team (Paul, Howard, Rahul bhai, Diggy, Shahid, Luke) for keeping me updated with cricketing skills and also for those wonderful matches which we lost.

Sandeep and Navya. Who would have thought that you will follow me till the end when we started studying together? Everytime I think of going away, you would be there; following me all the way from India to Uppsala, I was happy that I moved to KI and you moved even. Thank you for all these great years. Navya I still remember the offer you had for me when myself and Smitha were staying at your place. Appu, Sailesh, Kiran, Raghu, Rakesh for being friends for ever.

The most phenomenal bunch of the people around: Sarosh and Rashmi, Buppi and Sonchita, Onu and finally Chandu.. I have shared some of the most amazing time of my stay with you guys. Thank you for all those dinners and New Year parties. Sonchita for your smile and lovely hugs, I hope I will be able to return your lunch box one day, Buppi for you nice jokes which only we understood and those Punjabi dholak songs, Sarosh for lovely drives and cool moments and lovely mutton biryani’s, Rashmi for always irritating me☺ and concentrated “tuppa” foods. Onu for always smiling and talking loud. Chandu I will never forget your special dance on New Years Eve.

Professor. M.K Mathew (NCBS, Bangalore) for the allowing me to gain experience from his research lab and introducing me to the scientific world. Professor. N.H.Manjunath, Professor. N.Ramachandraswamy, Dr. V.R.Devaraj, Dr. C.S.Karigar, Dr. K.R.Siddalingmurthy my Master thesis guides.

Siju (the..i), for being besides me from the beginning till end. Thanks for everything.

Pappa, mummy, Amma, Achan, Unniattan, Sini, Sradha katty®, Sunilattan, Shree, I have kept you all till the end since I have to write from the bottom of my heart, thanks for always looking after me and for all the support along the way. Mummy, special thanks for all those beatings which I realize now that it was for my good.

Smikutty, I don’t know what to write for you. I always felt that I have been pretty lucky in my life, but I started believing in those feelings when you entered my life. You are the one who matters the most. All the bad times in my life matter me little as long as I know that you are around. Thank you for being there always.
REFERENCES

2. Stampfer, M.R. & Yaswen, P. Culture models of human mammary epithelial cell transformation. 
8. Zalewski, Z. [Obturation emphysema in a case of a foreign body left in the left upper lobe bronchus]. 
12. Callahan, J.F. *et al.* Identification of novel inhibitors of the transforming growth factor beta1 (TGF-beta1) type 1 receptor (ALK5). 
16. Schultz-Cherry, S. & Murphy-Ullrich, J.E. Thrombospondin causes activation of latent transforming growth factor-beta secreted by endothelial cells by a novel mechanism. 


102. Arteaga, C.L. et al. Anti-transforming growth factor (TGF)-beta antibodies inhibit breast cancer cell tumorigenicity and increase mouse spleen natural


