CATHEPSIN CLEAVAGE OF EPIDERMAL GROWTH FACTOR RECEPTOR (EGFR) AFFECTS SIGNALING PATHWAYS OF CANCER CELLS

Marija Grozdanić

Doctoral Dissertation
Jožef Stefan International Postgraduate School
Ljubljana, Slovenia

Supervisor: Prof. Dr. Marko Fonović, Jožef Stefan Institute, Ljubljana, Slovenia

Evaluation Board:
Prof. DDr. Boris Turk, Chair, Jožef Stefan Institute, Ljubljana, Slovenia, and Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, Slovenia
Prof. Dr. Janko Kos, Member, Jožef Stefan Institute, Ljubljana, Slovenia, and Faculty of Pharmacy, University of Ljubljana, Ljubljana, Slovenia
Prof. Dr. Polona Jamnik, Member, Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia

MEDNARODNA PODIPLOMSKA ŠOLA JOŽEFA STEFANA
JOŽEF STEFAN INTERNATIONAL POSTGRADUATE SCHOOL

Marija Grozdanić

CATHEPSIN CLEAVAGE OF EPIDERMAL GROWTH FACTOR RECEPTOR (EGFR) AFFECTS SIGNALING PATHWAYS OF CANCER CELLS

Doctoral Dissertation

VPLIV KATEPSINSKE CEPITVE RECEPTORJA ZA EPIDERMALNI RASTNI DEJAVNIK (EGFR) NA SIGNALIZACIJO RAKASTIH CELIC

Doktorska disertacija

Supervisor: Prof. Dr. Marko Fonović

Ljubljana, Slovenia, November 2023

To my family.

Acknowledgments

I would like to express my deepest appreciation and gratitude to my mentor, Prof. Dr. Marko Fonović, for the invaluable guidance, support, and continuous encouragement throughout my doctoral journey. Most of all, I am thankful for his support and endless optimism, which guided me even when it was difficult. I am grateful for him believing in me from the very beginning.

I would also like to thank Prof. DDr. Boris Turk for accepting me into his research group, for the evaluation of my thesis and all the important advice regarding my work. I would also like to thank him for stepping in when I needed it the most and for ensuring excellent working conditions.

I would also like to thank other members of the evaluation board, Prof. Dr. Janko Kos and Prof. Dr. Polona Jamnik, for the critical review of my dissertation.

I would like to acknowledge the financial support provided by the Slovenian Research Agency, which enabled me to pursue my research and complete this dissertation.

I would also like to thank current and former members of our proteomics group, especially Dr. Barbara Sobotič, Dr. Matej Vizovišek and Dr. Robert Vidmar for introducing me to proteomics as a scientific field and teaching me all of the techniques, but also to Tilen Sever and Matej Kolarić for all great discussions, collaborations and also for a great friendship.

I would also like to thank other members of the B1 Department for all the help, and brainstorm discussions but most of all for the beautiful friendships, and experiences that we had together. Special thanks to Maja, Andreja, Nežka, Janja, Petra and Jelena for such a beautiful friendship, support and understanding. I am grateful to consider them my close friends.

I am also thankful to my dear friends for their unwavering support, love, and encouragement throughout the journey. I am grateful for meeting Kaća, Tea, Zvezdan and Ana. My PhD journey became fulfilled, exciting and so much better with them.

I am also grateful for my friends from back home, Ivona, Tamara, Ana, Tijana, Mia, Boba, Uroš and Milan, for all the support and love that did not change over the years. I am thankful to them for being my home. I would also like to thank Miljana and Miloš for their love and support.

I would also like to thank my second family, Karmen, Marijo, Luka and Martin Uglešić, for being there for me in every step of this journey, and for their love and support.

And finally, I am thankful for my biggest support, my family, my person and sister Jelena, for Rene and my parents Stevan and Ljiljana Grozdanić. Their love and support made me believe that I can achieve everything that I put my mind to. Their belief in me, countless words of encouragement, and understanding during the challenging times have been a constant source of strength. Their presence in my life has made this achievement even more meaningful, and I am profoundly grateful for their support. They are my rock stars!

Abstract

Epidermal growth factor receptor (EGFR) belongs to the ErbB family of receptor tyrosine kinases and plays an essential role in cell differentiation, migration, proliferation, and metabolism. Alterations in EGFR signalling were found in a number of cancers like lung, breast, colorectal cancer and gliomas, shown to be more aggressive and resistant to therapeutics. Consequently, EGFR has been intensively studied as an important anticancer therapeutic target. Signalling through EGFR is commonly triggered by ligand binding. However, deletions in the extracellular region of EGFR can also cause constitutive activation, such as in the case of the EGFRvIII variant. Such deletions can influence receptor activation and downstream signalling cascades. Also, due to its constitutive phosphorylation, cancer cells expressing this variant are highly tumorigenic. EGFR was also identified as a substrate of extracellularly present cysteine cathepsins, known to cleave ectodomains of membrane proteins, including receptors, cytokines, and adhesion proteins.

In this study, we confirmed extracellular cathepsin L-mediated cleavage of epidermal growth factor receptor (EGFR) and identified the cleavage site in the extracellular domain after R224. To further evaluate the relevance of this cleavage, we cloned and expressed a truncated version of EGFR, starting at G225, in HeLa cells. We confirmed the constitutive activation of the truncated protein in the absence of ligand binding and determined possible changes in intracellular signalling. Furthermore, we determined the effect of truncated EGFR protein expression on HeLa cell viability and response to the EGFR inhibitors erlotinib and cetuximab. Our data reveal the nuclear localization and phosphorylation of EGFR and signal transducer and activator of transcription 3 (STAT3) in cells that express the truncated EGFR protein and suggest that these phenomena cause resistance to EGFR inhibitors. Also, we determined the effects of truncated EGFR expression on HeLa cell migration. However, our findings open new questions about EGFR and could possibly lead to more effective strategies in anticancer therapy.

Povzetek

Receptor epidermalnega rastnega faktorja (EGFR) spada v družino receptorskih tirozin kinaz ErbB in igra bistveno vlogo pri celični diferenciaciji, migraciji, proliferaciji in metabolizmu. Spremembe v signalizaciji EGFR so odkrili pri številnih vrstah raka, kot so pljučni rak, rak dojke, črevesni rak in gliomi, za katere se je izkazalo, da so bolj agresivni in odporni na terapevtike. Posledično so EGFR intenzivno preučevali kot pomembno terapevtsko tarčo proti raku. Signalizacijo prek EGFR običajno sproži vezava liganda, vendar lahko delecije v zunajcelični regiji EGFR povzročijo tudi konstitutivno aktivacijo, kot na primer v primeru različice EGFRvIII. Takšne delecije lahko vplivajo na aktivacijo receptorjev in navzdolnje signalne kaskade. Prav tako so zaradi svoje konstitutivne fosforilacije rakave celice, ki izražajo to varianto, zelo tumorogene. EGFR so identificirali tudi kot substrat zunajcelično prisotnih cisteinskih katepsinov, za katere je znano, da cepijo ektodomene membranskih proteinov, vključno z receptorji, citokini in adhezijskimi proteini.

V tej študiji smo potrdili cepitev receptorja epidermalnega rastnega faktorja (EGFR), ki jo povzroči zunajcelični katepsin L, in identificirali mesto cepitve v zunajcelični domeni po aminokislinskem ostanku R224. Za nadaljnjo oceno pomembnosti te cepitve smo klonirali in izrazili skrajšani EGFR z začetkom pri G225 v celicah HeLa. Potrdili smo konstitutivno aktivacijo skrajšanega proteina v odsotnosti vezave liganda in določili možne spremembe znotrajceličnega signaliziranja. Poleg tega smo določili učinek izražanja skrajšanega proteina EGFR na sposobnost preživetja celic HeLa in odziv na inhibitorja EGFR erlotinib in cetuksimab. Naši podatki razkrivajo jedrna lokalizacijo in fosforilacijo EGFR ter prenašalca signala in aktivatorja transkripcije 3 (STAT3) v celicah, ki izražajo skrajšani protein EGFR, in kažejo, da ti pojavi povzročajo odpornost na inhibitorje EGFR. Določili smo tudi učinke izražanja skrajšanega EGFR na migracijo celic HeLa. Naše ugotovitve odpirajo nova vprašanja o EGFR in bi lahko vodile do učinkovitejših strategij pri zdravljenju raka.

Contents

List of Figures

List of Tables

Abbreviations

ADAM	...	A disintegrin and metalloproteinase
ALCAM	...	Activated Leukocyte Cell Adhesion Molecule
Akt	...	Protein kinase B
APCs	...	Antigen presenting cells
AREG	...	Amphiregulin
ATP	...	Adenosine triphosphate
AXL	...	Tyrosine-protein kinase receptor UFO
BAD	...	Bcl2-associated agonist of cell death
Bax	...	Bcl2-associated X protein
Bcl2	...	B-cell lymphoma 2
BIM	...	Bcl2 interacting mediator of cell death
BTC	...	Betacellulin
C-Myc	...	Myc proto-oncogene
CAM	...	Cell adhesion molecule
CD8+	...	CD8-positive cytotoxic T cells
COX2	...	Cyclooxygenase-2
CT	...	C-terminal tail
DAG	...	Diacylglycerol
DNA	...	Deoxyribonucleic acid
ECM	...	The extracellular matrix
EGF	...	Epidermal growth factor
EGFR	...	Epidermal growth factor receptor
EMA	...	European Medicines Agency
EMT	...	Epithelial-mesenchymal transition
EPGN	...	Epigen
ER	...	Endoplasmic reticulum
ERK	...	Extracellular signal-regulated kinase
FDA	...	Food and drug administration
GAGs	...	Glycosaminoglycans
GAPDH	...	Glyceraldehyde 3-phosphate dehydrogenase
GBM	...	Glioblastoma multiforme
GDP	...	Guanosine diphosphate
Grb2	...	Growth factor receptor-bound protein 2
GTP	...	Guanosine triphosphate
HB-EGF	...	Heparin-binding EGF-like growth factor
HER	...	Human epidermal growth factor receptor
IFI27	...	Interferon alpha-inducible protein 27
IFIT1	...	Interferon induced protein with tetratricopeptide repeats 1
IL	...	Inteleukin
IP3	...	Inositol triphosphate
IRF3	...	Interferon regulatory factor 3

JM	...	Juxtramembrane domain
JNK	...	c-Jun N-terminal kinase
KRAS	...	Kristen rat sarcoma viral oncogene homolog
M6P	...	Mannose 6-phosphate
mAb	...	Monoclonal antibody
MAPK	...	Mitogen-activated protein kinase
MCOLN1	...	Mucolipin-1
MEK	...	Mitogen-activated protein kinase kinase
MET	...	Hepatocyte growth factor receptor
mRNA	...	Messenger ribonucleic acid
mTOR	...	Mammalian target of rapamycin
mTORC2	...	Mammalian target of rapamycin complex 2
NF_KB	...	Nuclear factor kappa-light-chain-enhancer of activated B cells
NK cell	...	Natural killer cell
NRAS	...	Neuroblastoma rat sarcoma viral oncogene
NRG	...	Neuregulin 1
NSCLC	...	Non-small cell lung cancer
PDGFR	...	Platelet-derived growth factor receptor
PDK1	...	Phosphoinositide-dependent kinase-1
PI3K	...	Phosphoinositide 3-kinase
PIP2	...	Phosphatidylinositol 4,5-bisphosphate
PIP3	...	Phosphatidylinositol 3,4,5-trisphosphate
PKB	...	Protein kinase B
PKC	...	Protein kinase C
PLCγ	...	Phospholipase C gamma
PLD	...	Phospholipase D
PM	...	Plasma membrane
PTB	...	Phosphotyrosine-binding domain
RaI	...	Retinoic acid-induced protein 1
Raf	...	Rapidly accelerated fibrosarcoma kinase
Ras	...	Rat sarcoma virus GTPase
RTK	...	Receptor tyrosine kinase
SCLC	...	Small cell lung cancer
SH2	...	Src homoly 2 domain
Shc	...	Src homology and collagen family
SIF-A	...	STAT3 homodimer
SIF-B	...	STAT1/STAT3 heterodimer
Sos	...	Son of sevenless
Src	...	Proto-oncogen tyrosine-protein kinase FSI
STAT	...	Signal transducers and activators of transcription
TAMs	...	Tumor-associated macrophages
TGFα	...	Transforming growth factor alpha
TKD	...	Tyrosine kinase domain
TKI	...	Tyrosine kinase inhibitor
TM	...	Transmembrane domain
TME	...	Tumor microenvironment
TNBC	...	Triple-negative breast cancer
TRAIL	...	Tumor necrosis factor-related apoptosis inducing ligand

Standard Amino Acid Abbreviations

Amino acid	Three-letter abbreviation	One-letter abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Phenylalanine	Phe	F
Glycine	Gly	G
Glutamine	Gln	Q
Glutamic acid	Glu	E
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Proline	Pro	P
Serine	Ser	S
Tyrosine	Tyr	Y
Threonine	Thr	T
Tryptophan	Trp	W
Valine	Val	V

Chapter 1

Introduction

1.1 Proteases

Proteases are enzymes which catalyze the hydrolysis of the peptide bond. Proteolysis is one of the essential irreversible post-translational modifications [1]–[3]. Proteases have been classified based on amino acids responsible for their catalytic activity, and we can differentiate metallo, serine, threonine, aspartic, and cysteine proteases. Proteases could also be classified as amino- or carboxy-peptidases if protease cleaves N- or C- terminus of the substrate, respectively, or endopeptidase if protease cleaves in the middle of the substrate [2], [4]. Proteolytic processing can be non-selective and important in general protein turnover but also very selective. Some proteases can cleave at a specific site on the protein and change protein activity, cellular localization, and structural integrity. Consequently, proteases represent important regulators of various biological processes such as cell differentiation and proliferation, apoptosis, tissue remodelling and wound healing. Dysregulation of this proteolytic processing leads to various diseases, like cancer, cardiovascular and neurodegenerative diseases [5], [6].

1.2 Cysteine Cathepsins

Among cathepsins in humans, we can differentiate serine cathepsins A and G, aspartic cathepsins D and E, and the largest group of 11 cysteine cathepsins B, C, F, H, K, L, O, S, V, X and W. Cysteine cathepsins belong to the CA clan of cysteine proteases, more specifically to the C1 (papain-like) family. They are small monomeric proteins with molecular mass in the 25-30 kDa range, except for cathepsin C, which is a tetramer with a molecular mass of around 200 kDa. Most of the cysteine cathepsins are ubiquitously expressed in human organisms, except for several cathepsins with restricted tissue distribution [3], [7], [8]. Cathepsin K is predominantly expressed in osteoclast and most epithelial cells, while cathepsin V is restricted in the thymus and testis [9]–[11]. Additionally, cathepsin S is highly expressed in the antigen-presenting cells (APCs) (dendritic cells and B-cells), and cathepsin W is predominately expressed in natural killer cells (NK cells) and CD8+ lymphocytes [12], [13].

The cysteine cathepsins are mainly endopeptidases. However, few are exceptions, cathepsin C acts as an aminopeptidase, while cathepsin X is a carboxypeptidase [3], [14]. The exceptions are also exopeptidases cathepsin B and H since they can also act as endopeptidases [15], [16]. For optimal activity, cysteine cathepsins require a slightly acidic environment. Consequently, cysteine cathepsins are predominately found in endosomes and lysosomes [17]. For a long time, cysteine cathepsins were considered responsible for non-specific protein degradation in the endosomal-lysosomal system. However, their presence

in other cellular compartments and their specific roles in various physiological processes have become increasingly recognized. Cysteine cathepsins were also found in the cell cytoplasm, nucleus and extracellular space [18]–[21].

1.2.1 Cathepsin's structure and specificity

All cysteine cathepsins share the same papain-like fold, composed of two domains. The left domain (L-) is composed of three α helices, and the right domain (R-) is a domain with a β barrel structure (Figure 1) [7]. Domain L- and R- together form an active site cleft responsible for proteolytic activity. Two active site residues, Cys25 and His159 (papain numbering), create a thiolate-imidazolium ion pair in the active site cleft.

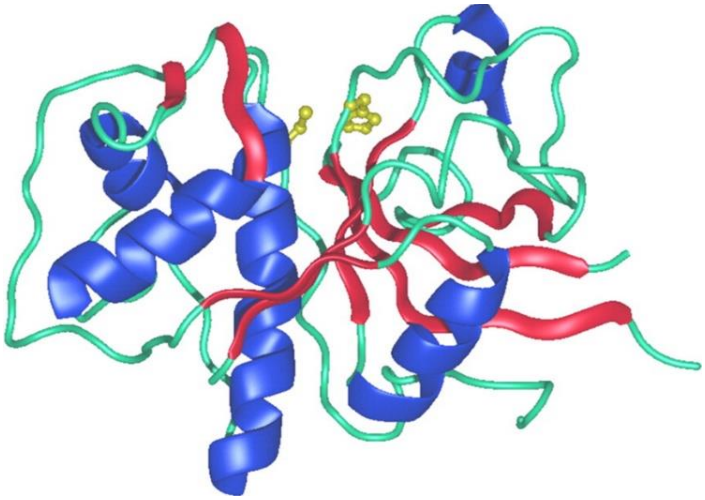


Figure 1: The molecular structure of cathepsin L. Secondary structure elements are presented in blue colour for α-helices and red colour for β-sheets. The active site residues are coloured yellow [7].

The binding surface responsible for cleavage is called the protease binding cleft [22]. The protease binding cleft is composed of subsites, S1-Sn (non-primed subsites) towards the N-terminus from the scissile bond (cleavage site) and S1'-Sn' (primed subsites) towards the C-terminus from the scissile bond. Subsites within the binding cleft recognize and bind substrate residues, P1-Pn and P1'-Pn', according to the same nomenclature (called Schechter-Berger nomenclature (Figure 2) [23]. In cysteine cathepsins, both domains contain loops creating the active site cleft. While endopeptidases have the active site cleft extended throughout the whole interface between domains L- and R-, exopeptidases contain additional features which reduce the number of binding sites [7], [24].

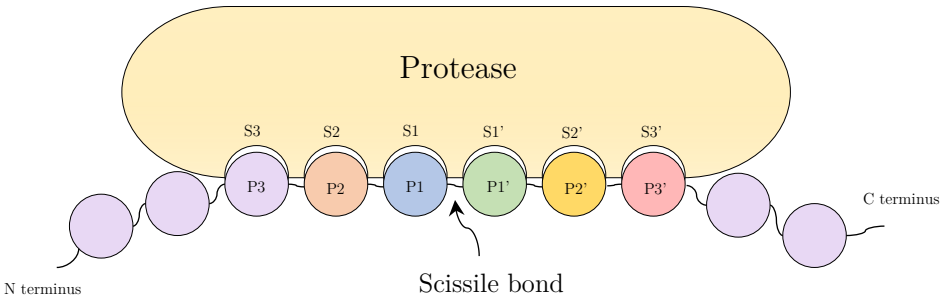


Figure 2: Schematic representation of substrate binding to the protease according to Schechter-Berger nomenclature [23].

Cysteine cathepsins have broad specificity, contrary to some members of cysteine proteases, specifically legumain and caspases, which cleave after one or two specific amino acids. Within cysteine cathepsins, only the S2 site forms a defined pocket and, together with S1 and S1' binding sites, are recognized as major substrate recognition sites. Together, the results from proteomic and combinatorial peptide library approach research show that cysteine cathepsins in position P2 prefer small hydrophobic amino acids like Leu, Ile and Val, even though aromatic amino acid residues like Phe and Tyr can also be recognized. Some exceptions exist. Cathepsin K has a preference for Pro in P2 and Gly in the P3 position, while cathepsin B prefers Arg in the P2 position [22], [25]–[28].

1.2.2 Activation and regulation of cathepsins

Like most enzymes, cysteine cathepsins are synthesized as preproenzymes. Preproenzyme is synthesized in the rough endoplasmic reticulum (ER), and signal peptide is removed proteolytically during the passage through ER lumen. The proenzyme is next transported to the Golgi apparatus for proper folding and post-translational modifications (N-linked glycosylation and mannose-6-phosphorylation). Further, proenzyme is recruited into vesicles for transport to the late endosomes/lysosomes [29], [30]. Mature, proteolytically active cathepsin is released after propeptide removal. Propeptide regulates enzyme activity and prevents their early activation and unwanted proteolysis. Propeptide folds through the active site cleft of the cathepsins and inhibits the enzyme's activity. It has been shown that propeptides are very potent and specific inhibitors of their cognate enzymes, and the amino acid sequences show some similarities [31]. Propeptide could be removed from the enzyme either autocatalytically under acidic conditions or by proteases like cathepsin D, legumain and similar lysosomal proteases. After propeptide removal, mature and active cathepsin can create a chain reaction and rapidly activate other procathepsins. Contrary to endopeptidases, cathepsin C and X cannot be activated autocatalytically. These exopeptidases require endopeptidases for their activation, such as cathepsin S and L, [21], [32], [33].

Molecules with an important role in cysteine cathepsins regulation are glycosaminoglycans (GAGs) and other negatively charged surfaces. It has been shown that GAGs facilitate the autocatalytic activation of cysteine cathepsins. GAGs binding to the procathepsin cause conformational changes, leading to unleashed interaction between propeptide and mature enzyme and, consequently, easier processing of another procathepsin [34]–[36].

1.3 Extracellular Cysteine Cathepsins

As already mentioned, cysteine cathepsins are mainly located in the endo-lysosomal compartment. However, cysteine cathepsins have several different transport mechanisms intra- and extracellularly. These transport mechanisms are cell type-dependent and depend on the pathophysiological status of the cells [20], [37]. In some epithelial cells, and upon specific signalling, proteolytically active cysteine cathepsins can be secreted into the extracellular space. They can be recruited into the transport vesicles, translocated from the endosomes/lysosomes to the plasma membrane (PM), and secreted into the extracellular space. Another transport mechanism is lysosomal exocytosis, which represents the secretory pathway of lysosomes. During lysosome exocytosis, lysosomes are moved toward the plasma membrane in a kinase-dependent manner. After docking to the plasma membrane, lysosomes are fused with PM by triggering an increase in Ca^{2+} through the endo/lysosomal cation channel mucolipin 1 (MCOLN1). Fusion further leads to the secretion of the lysosomal components into the extracellular space [38]–[40]. In different cells, like macrophages and fibroblasts, cysteine cathepsins are secreted as zymogens due to impaired mannose-6 phosphate (M6P) signalling. Recycling of M6P to the Golgy could be disrupted due to pH changes, and M6P deficiency could lead to the rerouting of the cysteine cathepsin zymogens to the extracellular space [41]. Since cysteine cathepsins are stable under acidic conditions, secretion as zymogens can prolong their stability at neutral pH found in extracellular space. However, cysteine cathepsins could also be secreted in a mature form. In this case, vacuolar-type H+ ATPases (V-ATPase) provide local acidity for prolonged cysteine cathepsin activity in the extracellular milieu [42]. GAGs regulation of the cysteine cathepsins also occurs at a pH close to neutral and contributes to cysteine cathepsin activity in extracellular space [43].

1.3.1 Extracellular cysteine cathepsins in physiological processes

Secretion of cysteine cathepsins into the extracellular space was found to be involved in several physiological processes like bone remodelling, and wound healing. Cysteine cathepsins regulate this processes through the processing of the extracellular matrix (ECM) proteins [21], [43]. The extracellular matrix represents a matrix of proteins secreted by surrounding cells within the tissues and organs. ECM is composed of structural proteins such as fibrous proteins (elastin, collagen), proteoglycans, hyaluronan and matricellular proteins (tenascin, osteopontin), and glycosaminoglycans [44]. ECM homeostasis is regulated by the constant balance between protein synthesis and degradation [45]. During bone remodelling, osteoclasts are the specialized cells responsible for bone resorption and remodelling. Cathepsin K is secreted by osteoclasts and is crucial in collagen type I and elastin degradation during bone resorption. Cathepsins are responsible for processing two more bone ECM proteins, osteonectin, a protein involved in cell-matrix interaction, and osteocalcin, an ECM protein implied in bone formation and insulin metabolism [46]–[49]. Cathepsin B is involved in keratinocytes migration during wound healing. It is secreted by keratinocytes to degrade ECM proteins and facilitate the migration of the keratinocytes [50], [51].

In addition, secreted cysteine cathepsins were found to be involved in prohormone processing. Cathepsins B, L and K are secreted by thyroid epithelial cells, where they can process extracellular thyroglobulin and release thyroid hormone [52]. Extracellular cathepsins also have an important role in releasing peptide neurotransmitters during the cell-cell communication of neurons and endocrine cells. Cathepsin L and V were found involved in releasing the neuropeptide Y, cholecystokinin, encephalin and dynorphins. This

process is facilitated by the fact that the processing of prohormones occurs in secretory vesicles, where the pH is acidic [53], [54].

1.3.2 Extracellular cysteine cathepsins in pathological conditions

Dysregulations of the cysteine cathepsins expression, localization, or activity have been connected with several pathological conditions, like arthritis, osteoporosis, cardiovascular diseases, various lung pathologies and cancer (Figure 3) [4], [55]. For example, osteoclast hyperactivity is responsible for bone mineral density loss during osteoporosis. As mentioned, cathepsin K is one of the most abundant proteases secreted during bone resorption [56]–[59]. In the development of arthritis, various matrix metalloproteases and cysteine cathepsins were connected with cartilage degradation. During disease progression, the pH at the cartilage surface changes from pH 7 to 5, and as metalloproteases are mainly active in neutral pH, cysteine cathepsins take over the important role in the later stages of arthritis [60]–[63].

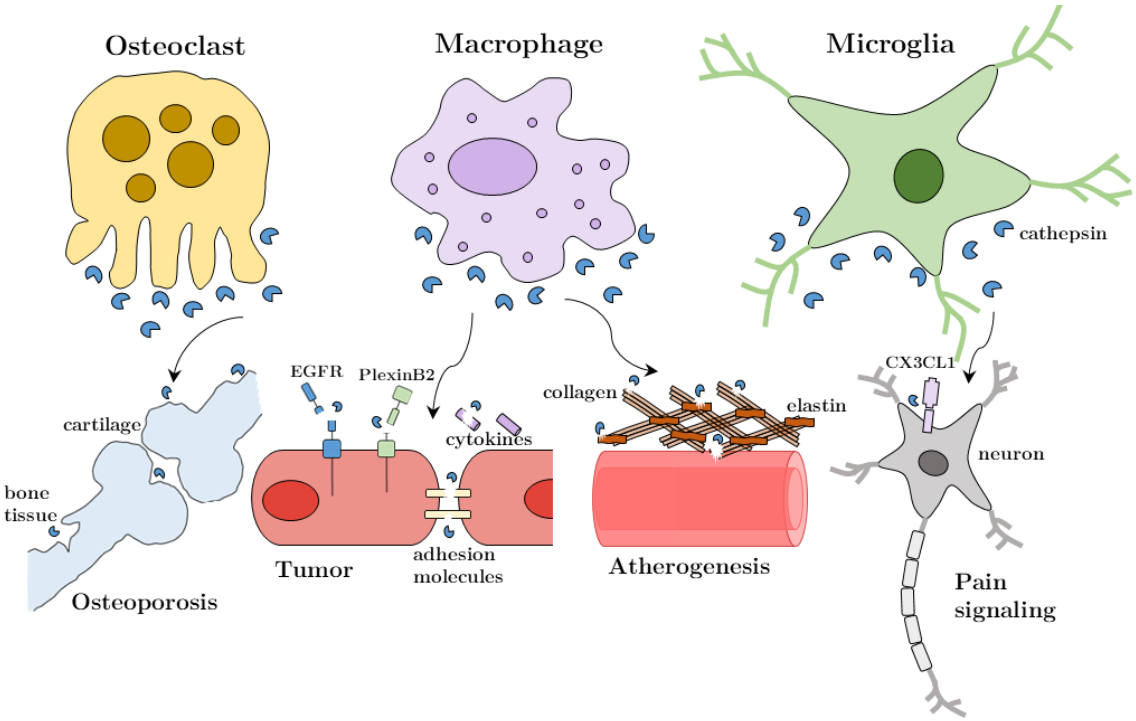


Figure 3: Graphical presentation of cysteine cathepsins secreted in various pathological conditions. Cathepsins could be secreted: by osteoclast and promote osteoporosis and arthritis; by tumour cells, and tumour-associated macrophages and process different signalling molecules involved in tumour progression; by macrophages promoting atherosclerosis; and by microglia cells and be involved in neuropathic pain signalling [6], [55].

Extensive degradation of the ECM is one of the characteristics of several cardiovascular disorders, such as atherosclerosis, cardiomyopathy, coronary and valve disease, and aortic aneurysms [64]–[67]. Extensive ECM protein degradation by cathepsins K, S and V leads to destabilization of artery structure or valve dysfunction [68]. Macrophages secrete cathepsins S, B and L within atherosclerotic lesions, where these cathepsins are upregulated and responsible for elastolytic activity in the atherosclerosis plaque [69]–[71]. Cysteine cathepsins are secreted by intestinal macrophages and have an important role in

inflammatory bowel disease [72]. Cathepsin K was linked to psoriasis since it was upregulated in the psoriatic lesions [73]. In addition, the T-cells and several cytokines were found to trigger psoriatic keratinocytes expression of the cathepsin S, where it can cleave IL-36γ, involved in psoriatic inflammation [74], [75]. Finally, in the central nervous system, cathepsin S is secreted by microglia, where it can cleave fractalkine (CX3CL1) and have an important role in neuropathic pain signalling [76]. However, one of the first determined and most extensively studied pathological roles is cysteine cathepsins' role in various types of cancer.

1.3.3 Role of the extracellular cysteine cathepsins in cancer

1.3.3.1 Cysteine cathepsins in cancer

Cysteine cathepsins are highly upregulated in various cancers, like lung, breast, pancreatic, brain cancer and melanoma. Cathepsins have been connected with various stages of cancer progression, growth and metastasis [4], [77]–[79]. Genetic ablation of cathepsin B, L, H and S significantly abolished tumour growth and progression in the mammary gland and/or pancreatic islet mouse model [6], [80]–[82]. However, in the skin cancer mouse model, genetic ablation of cathepsin L showed increased tumorigenesis [83]. Using selective inhibitors, cathepsin B was found to be involved in bladder cancer since its inhibition abolished tumour bladder cancer growth [84]. Inhibition of cathepsin X showed to inhibit migration of the prostate cancer cell model [85]. Furthermore, inhibition of cathepsin K was found to inhibit bone metastasis in the prostate and breast cancer models [86]. All these studies suggest that cathepsins have an important but very different role in cancer, depending on the type.

1.3.3.2 Cysteine cathepsins in the tumour microenvironment

It is increasingly recognized that the surrounding tumour microenvironment (TME) has an important role in carcinogenesis. TME is composed of surrounding immune cells, blood vessels, lymphocytes, fibroblasts, bone marrow-derived inflammatory cells, extracellular matrix (ECM) and various signaling molecules. Interaction within TME influences different stages in cancer development but also determines response to anticancer therapy [87]. Cysteine cathepsins are secreted into the TME from various types of cells, mostly from tumour-associated macrophages (TAMs) but also from fibroblast, epithelial and cancer cells [55], [88], [89]. Several mechanisms induce abnormal expression and secretion of cysteine cathepsins in TME. Cytokines, interleukin 4 (IL-4), IL-6 and IL-10, are responsible for the regulation and stimulation of TAMs, and TAMs enhanced secretion of cathepsins in several cancers [6], [88]. Signal transducer and activator of transcription 3 and 6 (STAT3 and STAT6) activation were found to promote overexpression and secretion of cysteine cathepsins B, C, S, and Z, mostly from macrophages [90]. Cathepsin L was found to be secreted by fibrosarcoma due to its poor affinity toward the M6P receptor [91].

One of the major roles of cysteine cathepsins in the tumour microenvironment is ECM degradation [43]. Cleavage of fibronectin by cathepsins B, S and L was connected with cell growth, migration and differentiation. The same cathepsins were found to cleave laminin and have a proangiogenic effect [92]–[94]. Another cysteine cathepsins substrate among ECM proteins is E-cadherin, a cell-cell adhesion molecule, and cleavage of this protein was connected with increased tumour invasiveness in the pancreatic islet mouse model [81]. Additionally, cathepsin K cleavage of periostin was connected with breast cancer bone metastasis, and a decrease of periostin intact in the C-terminal was proposed as a biomarker for osteolytic lesions in breast cancer bone metastasis [95], [96].

Cysteine cathepsins L and S were found to cleave extracellular domains (ectodomains) of several transmembrane proteins in mouse pancreatic cancer models, *in vitro* and *in vivo*. Using mass spectrometry-based approach, several cell adhesion molecule (CAM) proteins and transmembrane receptors were identified as substrates. Additionally, in cathepsin S deficient mice, cleavage of these substrates was highly decreased. Among identified proteins were CAM proteins like ALCAM, L1CAM and nectin-like protein 5 and transmembrane receptors like Plexin A1, Plexin B2 and epidermal growth factor receptor (EGFR) [97]. Interestingly, the majority of mentioned transmembrane proteins were already established as mediators of cancer progression.

1.4 Receptor Tyrosine Kinases

Receptor tyrosine kinases (RTKs) are transmembrane receptors responsible for cell-cell, and cell-environment communications, contributing to normal cell development and homeostasis. RTKs belong to one of the largest gene families, the protein kinase family. Members of this family are enzymes which catalyze the reaction below:

$$MgATP^{-1} + protein—OH \rightarrow Protein—OPO_3^{2-} + MgADP + H^+$$

The protein kinases facilitate the transfer of the phosphate group from the phosphate donor as adenosine triphosphate (ATP) to the specific substrate. Based on the nature of the —OH group, they are classified as protein-serine/threonine kinases, protein tyrosine kinases and tyrosine-kinases-like proteins. Tyrosine kinases can be receptor and non-receptor kinases [98]. There are nineteen different subfamilies of RTKs, similar in basic structure, where all have an N-terminal extracellular domain (ECD), a single transmembrane domain, and an intracellular kinase domain with a C-terminal tail region. The signalling of RTKs is activated by ligand binding to ECD. The major difference between RTKs is in ECD, where they differ in the composition of different subdomains and, consequently, in the capability to bind structurally different ligands. However, alteration in receptors abundance, activity, or cellular distribution could be linked to various diseases, like cancer, diabetes, atherosclerosis and angiogenesis [99], [100].

1.5 Epidermal Growth Factor Receptor

Epidermal growth factor receptor (EGFR) belongs to a family of receptor tyrosine kinases, to the human epidermal growth factor receptor family (ErbB/HER). It was first identified as ErbB protein, discovered in 1962 by Cohen, and it is one of the most studied signalling proteins [101]. The ErbB family includes four family members, EGFR/HER1/ErbB1, HER2/ErbB2, HER3/ErbB3, and HER4/ErbB4. ErbB family can bind various ligands, create dimers (homo- and hetero-dimers) or higher oligomers, and activate intracellular signalling. ErbB2 is unique since it does not bind any known ErbB receptor ligand [102]. Additionally, ErbB3 can bind a ligand, but its activation shows very weak activation [103].

1.5.1 EGFR structure

EGFR is 170 kDa transmembrane glycoprotein composed of an N-terminal extracellular domain (ECD), a single transmembrane domain (TM), and juxtamembrane (JM) domain, intracellular tyrosine kinase domain (TK), followed by C-terminal tail (CT) (Figure 4) [98], [104]–[106]. ECD is composed of four subdomains, including domain I (L1), domain II (CR1), domain III (L2), and domain IV (CR2).

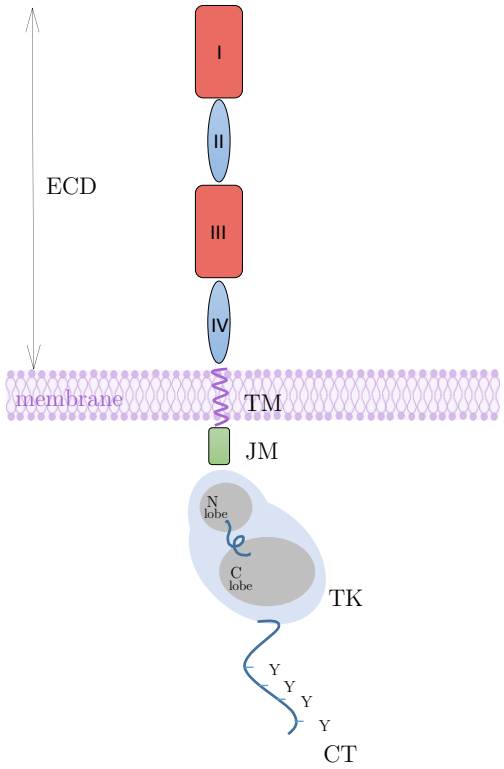


Figure 4: Graphical representation of the EGFR domain composition. EGFR extracellular domain (ECD) is composed of four subdomains, domains I and III (presented in red), and domains II and IV (blue). The structure continues with the transmembrane domain (TM) (violet helix) and the juxtamembrane domain (JM) (green). Tyrosine kinase domain (TK) (light blue) is presented together with its amino-terminal N and carboxy-terminal C-lobe (grey). And finally, C-terminal (CT) tail (blue line) with phosphorylation sites (Y).

Domains I and III are leucine-rich domains responsible for ligand binding, while domains II and IV are responsible for receptor dimerization since they contain numerous cysteine residues engaged in disulfide bond formation. TM domain is a single α-helical transmembrane peptide, self-associated in lipid membrane through one or more GxxxG dimerization motifs. TM domain continues with the JM domain, with a critical role in EGFR phosphorylation since its replacement did not affect ligand binding or receptor dimerization but abolished EGFR phosphorylation [107]. Furthermore, the JM domain interacts with negatively charged lipids in the membrane and contributes to receptor dimerization and signalling [108]. The tyrosine kinase domain is composed of a small amino-terminal lobe (N-lobe) consisting of one αC helix and five β-sheet strands (β1-5) and a large carboxy-terminal lobe (C-lobe) consisting of five α-helices. In a cleft between the two lobes is located the ATP-binding site [98], [106], [109], [110]. And finally, the C-terminal tail has almost all phosphorylation sites [111], [112].

1.5.2 EGFR ligands

Signalling of the EGFR highly depends on the presence of the EGF-like ligands [113]–[116]. EGF-like ligands are small molecules and can be classified into four subgroups:

1. ligands which specifically bind just ErbB1 or EGFR (Figure 5) (epidermal growth factor (EGF), tumour growth factor-α (TGF-α), amphiregulin (AREG), and epigene (EPGN));
2. ligands binding to ErbB1 and ErbB4 (heparin-binding EGF-like growth factor (HB-EGF) and betacellulin (BTC));
3. ligands binding to ErbB3 and ErbB4 (neuregulins 1 and 2 (NRG1 and NRG2));
4. and ligands binding just ErbB4 (NRG3 and NRG4).

Each ligand contains at least one EGF-like domain (Figure 5), or EGF motif, responsible for RTK binding and activation. EGF is unique and has nine EGF motifs, however, only the one adjacent to the membrane has this function and binds EGFR [114]. Each ErbB ligand is produced as an inactive, membrane-bound precursor [117]. Membrane-bound precursor ligand is composed of the extracellular segment, transmembrane and small intracellular segment. The active ligand, containing the EGF motif, is released from the cell membrane via proteolytic processing. Members of the ADAM (A disintegrin and metalloproteases) family are responsible for the cleavage of the membrane precursor ligand ectodomain, where active ligand occurs [118]. The released ligand can bind and activate receptors on distant cells (endocrine), neighbour cells (paracrine) or on the cell from which it originates (autocrine)[114].

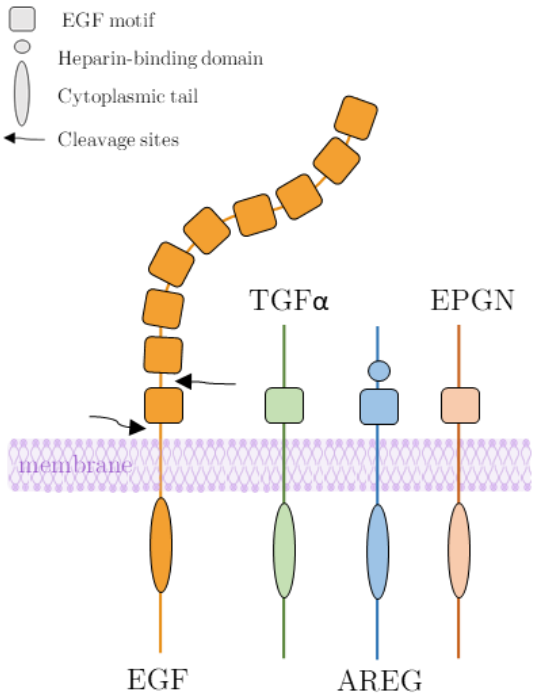


Figure 5: Graphical representation of the EGF-like ligands, which specifically bind EGFR: EGF; TGF α ; AREG; EPGN. Membrane-bound precursor ligand is composed of the extracellular, transmembrane and small intracellular segments. The active ligand is composed of one or more EGF motifs, which are released from the cell membrane via proteolytic processing. EGF is unique and has nine EGF motifs, and only the one adjacent to the membrane has this function and binds EGFR.

Human EGF is a 53 amino acids long ligand composed of six cysteines arranged in three disulfide bridges, forming three loops (Figure 6). Loop A from C1 and C3, loop B from C2

and C4, and loop C from C5 and C6. Besides receptor activation and signalling, ligands are involved in processes like protein sorting to different membrane localization, exosomal signalling, and cross-talk with other receptor signalling systems [119]–[122].

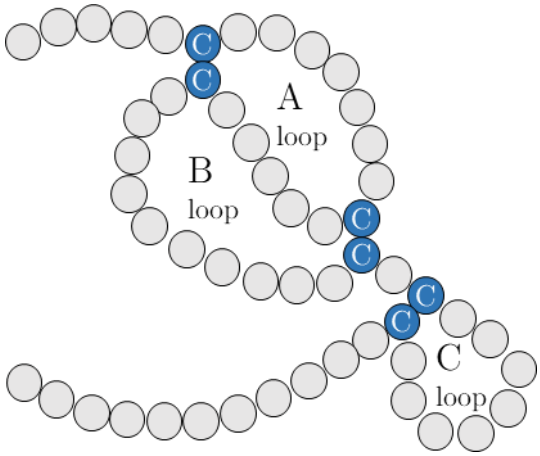


Figure 6: Graphical representation of the human EGF ligand, with six cysteines (C) arranged in three disulfide bonds, forming three loops, loop A, B and C.

1.5.3 EGFR dimerization induced by EGF binding

The crystal structure of all four ErbB receptors has been determined, either bound to the ligand or without one [104], [109], [123], [124]. Different studies have shown that EGF binding causes EGFR receptor conformational change, dimerization and tyrosine kinase activity in the intracellular domain (Figure 7). In the absence of the ligand, when the receptor is in closed or tethered conformation, the dimerization arm (β-hairpin-loop from domain II) is buried in domain IV, while dimerization is autoinhibited. Ligand binding promotes a significant conformational change in ECD, exposing a dimerization arm for interaction with the second ligand-bound receptor. Domain I of EGFR interacts with the EGF B-loop, while EGF A- and C- loops interact with EGFR domain III. This interaction pulls domains I and III together, breaks the tether of domains II and IV, and exposes the dimerization arm. ECD of the EGFR dimerizes, bringing the TKDs into proximity. TKDs create asymmetric dimer, the carboxy lobe of the first, the activator TKD, causes allosteric changes in the amino lobe of the second, the receiver TKD, and rotates into the active position. Active kinase further catalyzes the phosphorylation of tyrosine residues, creating a docking site for proteins or enzymes that drive downstream signaling [98], [105], [125], [126].

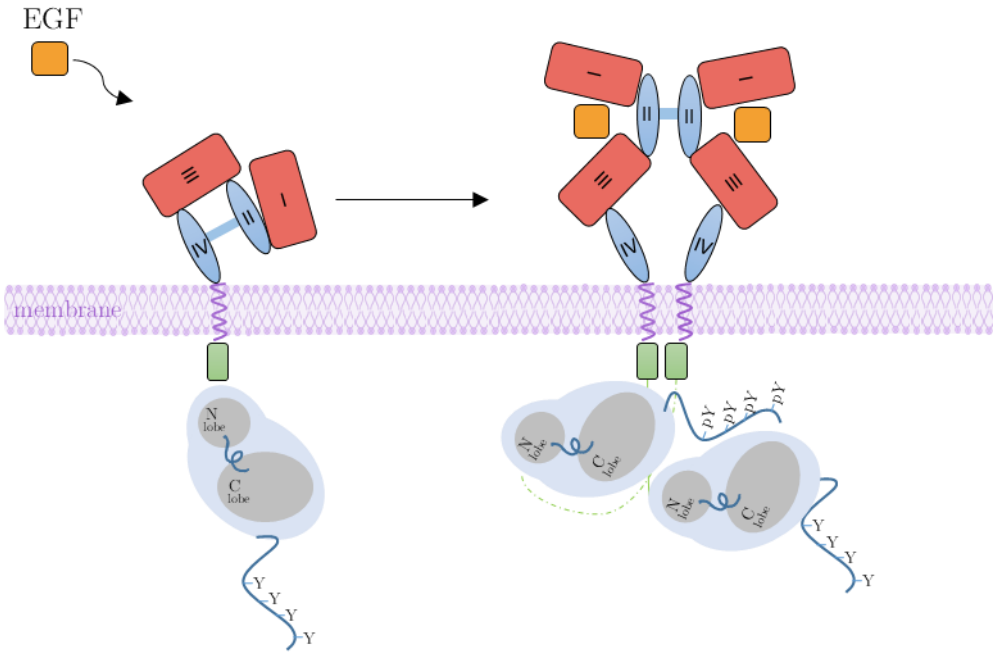


Figure 7: A model for EGF-induced dimerization of the EGFR.

1.5.4 Ligand-dependent signaling

In EGFR-mediated cellular signaling, the first step is tyrosine phosphorylation of cellular substrates. Substrate phosphorylation is mediated directly as an EGFR substrate or indirectly through EGFR-dependent phosphorylation and activation of other cellular kinases. Phosphorylation of the substrate can also be achieved through other members of the ErbB family as an EGFR partner, forming a heterodimer with EGFR. Phosphorylation of the EGFR C-domain provides specific docking sites for Src homology 2 (SH2) or phosphotyrosine binding (PTB) domain of intracellular signal transducers, further connecting with various signaling partners (Figure 8). EGFR can activate several pathways, pathways like Ras/Raf/MEK/ERK1/2 pathway, phosphatidylinositol 3-kinase PI3K/Akt pathway, and phospholipase C-γ (PLCγ) pathway [98], [127], [128].

In the Ras/Raf/MEK/ERK1/2 pathway, the significant role has the Grb2 adaptor protein. Grb2 binds to the EGFR either directly, the SH2 domain of Grb2 binds to the EGFR Y1068 or Y1086, or indirectly through the EGFR-associated Shc adaptor protein. The PTB domain of Shc binds to EGFR, phosphorylates, and recruits Grb2 to the receptor. Grb2 bound to the EGFR facilitate Ras association with Sos protein, resulting in the exchange of GDP for GTP and Ras activation. Ras further activates Raf-1, leading to activation and nuclear translocation of the extracellular signal-regulated kinases ERK1 and ERK2. ERK1/2 substrates are various kinases and transcription factors that promote cell proliferation [129], [130].

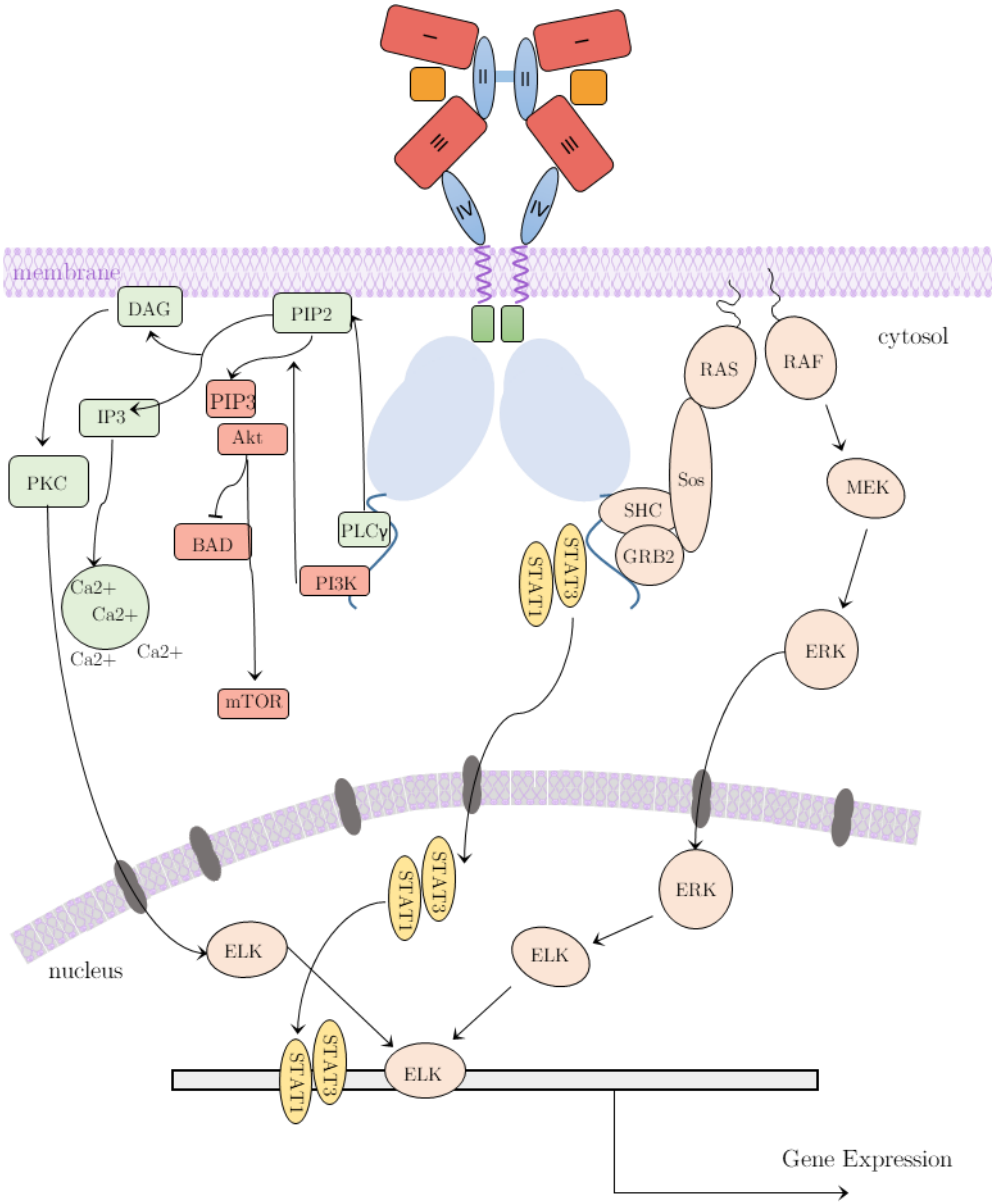


Figure 8: Signaling pathways activated by ligand binding to the EGFR: Ras/Raf/MEK/ERK1/2 pathway presented in orange; signal transducers and activators of transcription (STATs) pathway in yellow; phosphatidylinositol 3-kinase PI3K/Akt pathway in red, and phospholipase C- γ (PLC γ) pathway in green.

Activation with EGF affects cellular phospholipid metabolism. Phosphatidylinositol-3-kinase (PI3K), phospholipase C-γ (PLCγ), and phospholipase D (PLD) are some of the direct EGFR substrates [127]. PI3K is important in cellular processes like proliferation, adhesion, migration and cell survival. ErbB3, the binding partner of the EGFR, binds the p85 subunit of the PI3K via the SH2 domain and activates PI3K. Activated PI3K enzyme catalyzes the phosphorylation of phosphatidylinositol 4,5-biphosphate (PIP2) to form phosphatidylinositol 3,4,5-triphosphate (PIP3). PIP3 translocates Akt kinase (also known as protein kinase B (PKB)) to the plasma membrane, where it can be activated by phosphoinositide-dependent protein kinase 1 (PDK1) and mammalian target of rapamycin complex 2 (mTORC2), and phosphorylated at T308 and S473, respectively. Activated Akt kinase can catalyze the phosphorylation of mTOR (mammalian target of rapamycin)

and/or phosphorylation and inhibition of Bcl2-associated death promoter (BAD) and consequently be involved in processes like cell survival [131]–[133].

Activation of PLCγ is obtained by PLCγ binding to EGFR and its phosphorylation at Y783. Active PLCγ catalyzes the hydrolysis of PIP2 to form inositol 1,4,5-triphosphate (IP3) and diacylglycerol (DAG). IP3 mediated release of Ca^{2+} and can activate Ca^{2+}-dependent pathways such as RaI and nuclear factor κB (NFκB), while DAG activates protein-serine/threonine kinase C (PKC) with the wild spectre of downstream substrates involved in pathways like MAPK and JNK pathways. With phosphorylation of its substrates, PKC affects cellular processes like proliferation, migration, adhesion and cell death [127], [134], [135].

Another important signalling pathway activated downstream of EGFR signalling are signal transducers and activators of transcription (STATs) pathways. STAT family is composed of seven proteins, and among them, STAT1, STAT3 and STAT5 were found to interact with EGFR directly through the SH2 domain and phosphorylate at Y701, Y705, or Y694, respectively. After phosphorylation, STAT proteins dimerize. STAT3 can form homodimers, or heterodimers with STAT1. The activated complex then translocates to the nucleus, where it can regulate transcription by binding to various target genes or binding to other regulatory proteins like e-fos and c-jun. Finally, they activate transcription of the genes involved in processes like proliferation, differentiation and survival [136], [137].

EGF stimulation of the EGFR results in the stimulation of various intracellular signaling pathways. However, it should be considered that these signalling pathways are very often interlinked.

1.5.5 EGFR signalling in cancer

Alteration in EGFR signalling has been connected with the development and progression of many cancers, like lung, breast, colorectal cancer and gliomas (Figure 9). Alteration can be due to cross-talks with other cell-surface receptors, receptor and/or ligand overexpression, several gene alterations leading to constitutive activation, and due to the deficiency in receptor downregulation [106], [117], [138]–[140].

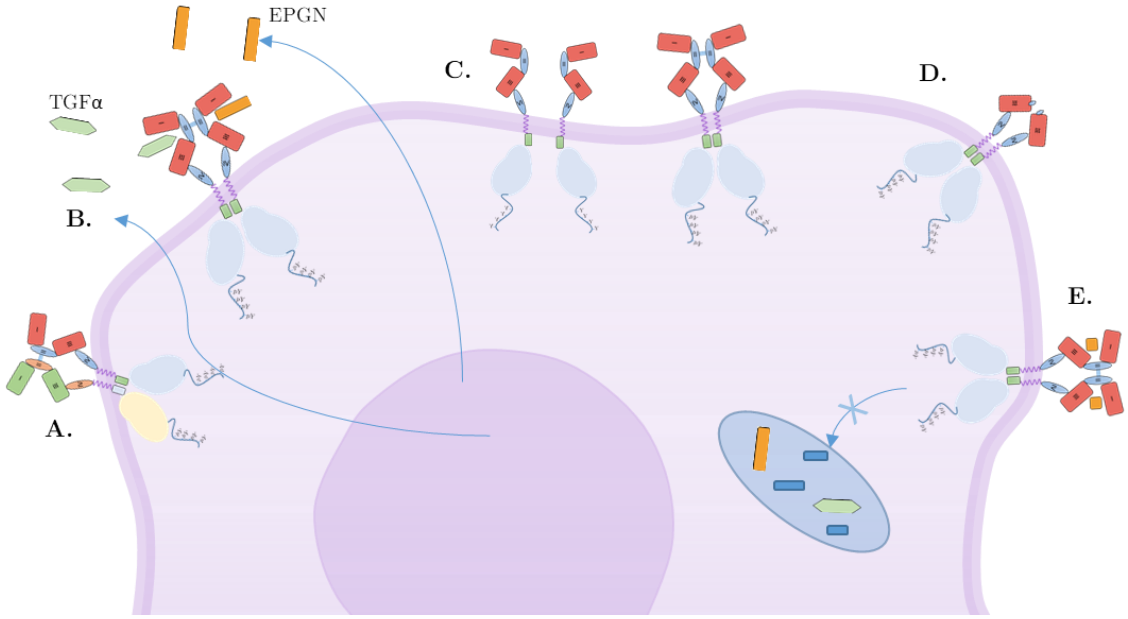


Figure 9: Mechanisms of the EGFR oncogenic signalling: A. Cross-talk with other receptors; B. Autocrine EGF-like ligands production; C. EGFR overexpression; D. Gene

alteration leading to constitutive activation of the receptor, as for EGFRvIII variant; D. impaired internalization and downregulation of the EGFR.

The EGFR cross-talk with other receptors was shown to have an important role in cancer cells (Figure 9A) [140]. EGFR can cross-talk with other receptors from the ErbB family but also with other members of RTKs, cytokine receptors, some cell adhesion molecules and some ion channels. The heterodimer of EGFR with ErbB2 was shown as one of the most tumorigenic among EGFR homo- and heterodimers [141]–[143]. Cell adhesion molecules like integrins are involved in EGFR activation in cancer cells and consequently, in the increase of cancer cell proliferation and survival [140], [144].

Overexpression of the EGFR is achieved either by amplification of the *EGFR* gene or due to changes in transcription level (Figure 9C). For example, the p53 protein, wild type and mutant, binds to its promoter and activates *EGFR* transcription [140], [145]. Gene amplification occurs in 30% of squamous cell carcinoma and 15% of adenocarcinomas. About 60% of the NSCLC showed the overexpression of EGFR. In 15-30% of breast cancers, EGFR was overexpressed, and connected with poor clinical outcomes and large tumour size. EGFR overexpression is also found in triple-negative breast cancer (TNBC) and represents the essential therapeutic target in TNBC [117].

Once activated, EGFR is internalized into the cell by endocytosis and eventually sorted either for lysosomal degradation and termination of the signalling, or recycling back to the cell membrane (Figure 9D) [146], [147]. Impaired receptor downregulation leads to prolonged and uncontrolled receptor signalling, associated with cancer. For example, enhanced recycling of the EGFR back to the plasma membrane was connected with the high level of EGFR in benign epithelial tumors caused by papillomavirus infection [148].

Various alterations of the EGFR gene were discovered. We can differentiate them based on their location, on the extracellular and intracellular (Figure 10) [117], [138]. There are several deletions in the extracellular domain of the EGFR gene, usually found in glioblastomas (GBMs). Type I variant (EGFRvI) is tumorigenic and lacks most of the extracellular domain (ΔM1-L566), type II (EGFRvII) has an in-frame deletion of exons 14 and 15, encoding amino acids residues in ECD subdomain IV (ΔE545-G627) [149], [150]. Since this deletion disables EGFRvII from forming tethered (closed) conformation, EGFRvII is forced into the extended conformation and shows ligand-independent activation.

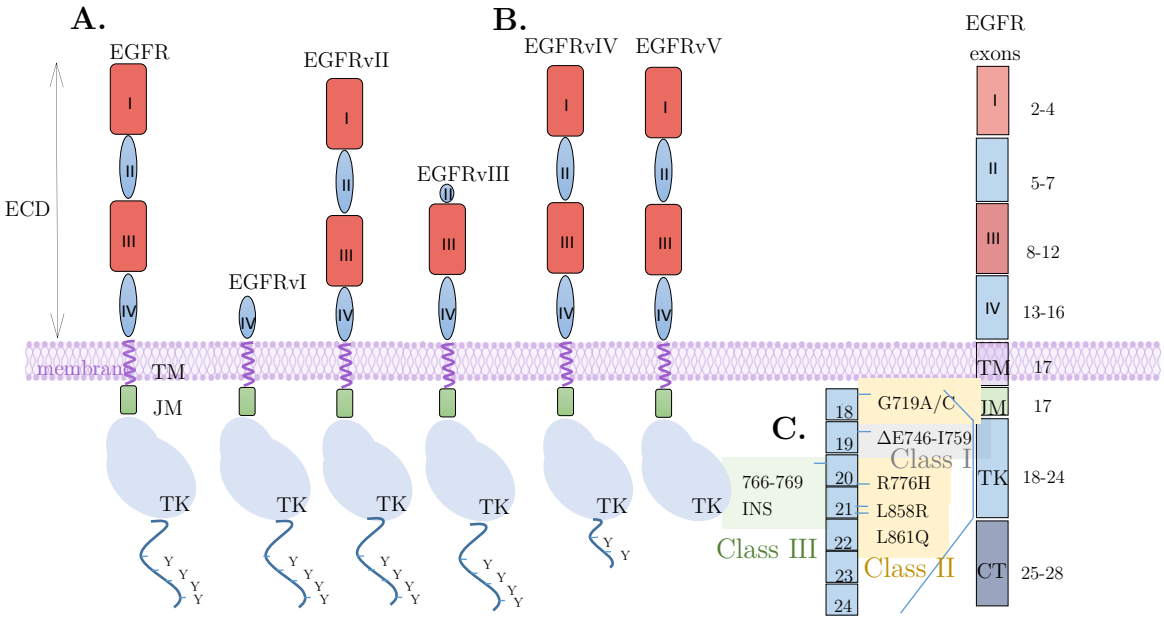


Figure 10: Frequent EGFR mutations located in the extracellular or intracellular region of the receptor. From left to right: EGFR deletions in A. the extracellular part (EGFRvI, EGFRvII and EGFRvIII); B. deletions in the intracellular part (EGFRvIV and EGFRvV); C. exon encoding human EGFR, and enlarged exons 18-24 with shown tyrosine kinase domain small mutations.

Type III variant (EGFRvIII) has an in-frame deletion of exons 2-7, and it lacks ECD subdomain I and two-thirds of subdomain II (ΔE30-T297). EGFRvIII is highly tumorigenic because of its constitutive phosphorylation, despite its lack of ligand binding [151], [152]. There are also some intracellular mutations in GBM. Type IV EGFR variant (EGFRvIV) lacks exons 25-27, while type V (EGFRvV) is truncated, and lacks exons 25-28, from amino acid 958 to the rest of the C-terminal [150], [153]. In non-small cell lung cancer (NSCLC) patients, various mutations were reported, most of them in the tyrosine kinase domain of the EGFR [154]–[156]. This mutation could be classified into three classes. Class I mutations are short deletions encoded by exon 19, causing the loss of four to six amino acid residues between E746 and S759. Class II mutation represents the substitution of a single amino acid residue, occurring between exon 18 to 21, while class III mutations involve duplications or insertions, mostly occurring in exon 20. Most of the mutated EGFRs show increased tumorigenicity and resistance to inhibitors used clinically [106], [156].

The complexity of the EGFR signalling in cancer cells makes it one of the most important and intensively studied anticancer targets.

1.5.6 Ligand-independent signaling

Activation of the EGFR signalling can also occur without ligand binding. This constitutive signalling was observed in cancer cells with EGFR overexpression, or cancer cells with EGFR mutations, usually in the extracellular part of the receptor.

1.5.6.1 EGFR overexpression

The high local concentration of the EGFR in a plasma membrane can cause receptor phosphorylation even without ligand binding. The relative phosphorylation level per receptor increases with the density of the EGFR on the plasma membrane [157]. When the density of receptors is similar to that found in cancer cells overexpressing EGFR, the level of EGF-independent phosphorylation becomes the same as EGF-dependent phosphorylation. Glioma cells overexpressing full-length EGFR without ligand stimulation showed increased mRNA expression for 66 genes [158]. When compared with the ligand-induced activation of the EGFR, overexpression of the EGFR showed phosphorylation of the EGFR tyrosines but also activation of different or noncanonical EGFR signaling pathways [159]. Overexpression of EGFR did not activate canonical signalling pathways like ERK or Akt signalling pathways. Constitutive activation of overexpressed full-length EGFR triggered different, noncanonical pathways, and activated transcription of IRF3 and downstream genes like TRAIL, IFIT1 and IFI27. However, additional ligand treatment of the overexpressed full-length EGFR using EGF resulted in canonical pathways' reactivation and the noncanonical pathway's termination. This study suggested that these two pathways are mutually exclusive [160].

1.5.6.2 EGFRvIII signaling

Among the EGFR mutants, one of the most common and most studied is EGFRvIII (Figure 10) [151], [152], [160]. EGFRvIII is found in GBM, NSCLC, breast and prostate cancer. As mentioned before, it lacks exons 2 to 7, encoding subdomain I and two-thirds of subdomain II of the EGFR extracellular domain [149]. Loss of most of the subdomain II is thought to prevent tethered, inactive conformation, favouring extended and active EGFR conformation. Also, since it lacks domain I, EGFRvIII cannot bind EGFR family ligands [161]. EGFRvIII is constitutively active, but its kinase activity is much weaker than the ligand-induced activity of the EGFR. However, this weak activity was shown to be sufficient to be tumorigenic. One of the reasons is unique membrane stability and impaired endocytosis and degradation of the mutant receptor. EGFRvIII remains active in the plasma membrane longer than full-length EGFR [162]. EGFRvIII is found to promote cell proliferation, invasion and angiogenesis, and reduction in apoptosis [152], [160], [163].

1.5.6.3 Proteolytic processing of the EGFR

As in the case of genetic deletion, complete or partial removal of the EGFR ectodomain by proteolytic cleavage can also influence the activity and functionality of the receptor. It was shown that the matriptase-prostasin cascade of serine proteases proteolytically cleaves the N-terminal region of EGFR and triggers its constitutive activation [164]. As mentioned, the cysteine cathepsins L and S can cleave the extracellular domain of EGFR [97]. Moreover, cathepsin S was also found to regulate EGFR signal transduction by mediating its endosomal degradation [165]. These suggest that cysteine cathepsins may be an important modifier/modulator of EGFR activity. Their role could be particularly important in pathological states such as carcinogenesis, where cathepsins are overexpressed and often secreted into the extracellular milieu [6].

1.5.7 Nuclear localization and function of the EGFR

Besides the well-known EGFR role and signalling from the plasma membrane, EGFR was also found to translocate to the cell nucleus [98], [166], [167]. After the ligand binding, EGFR goes through clathrin-dependent or clathrin-independent endocytosis, further

leading to receptor: (i) downregulation; (ii) degradation: (iii) recycling back to the plasma membrane; or (iv) translocation to the nucleus [168]. To pass a nuclear double membrane, proteins with molecular weight higher than 40kDa form the nuclear pore complex [169]. Around 30 proteins, known as nucleoporins, are responsible for nuclear pore complex formation. The importin α/β mechanism regulates EGFR translocation to the nucleus [170]. However, this mechanism represents one of the proposed mechanisms, all reviewed in [171]. Nuclear EGFR has been connected with several cancer types, including NSCLC, breast cancer and head and neck squamous cell cancer [170], [172], [173]. EGFR localized within the nucleus can function as a positive regulator of transcription, DNA replication, and DNA damage repair, resulting in the proliferation and spread (metastasis) of cancer cells [98], [174]. The C-terminal of the EGFR was found to be involved in intrinsic gene transactivation. Nuclear EGFR activates transcription of genes encoding cyclin D1, increasing cell proliferation, but also genes encoding aurora kinase A, inducible nitric oxide synthetase, and STAT1, involved in cell differentiation, proliferation, and survival [166], [167], [175]. EGFRvIII was also found in the nucleus in GBM, where together with EGFRwt activates STAT3 to promote cyclooxygenase-2 (COX-2) gene expression, the gene with an essential role in cancer progression [176]. Nuclear EGFR was connected with the resistance of cancer cells to various therapeutic regimes, and it is proposed to be a potent biomarker for therapy response [166].

1.5.8 Inhibitors of the EGFR

Two therapeutic strategies have been employed for targeting EGFR in various human cancers, monoclonal antibodies (mAb) and small molecular weight tyrosine kinase inhibitors (TKIs) [167], [177], [178]. These two strategies differ in inhibition mechanism. The mAbs target the extracellular domain of the EGFR and block ligand binding to the receptor. Consequently, dimerization and activation of intracellular signaling pathways of the EGFR are inhibited. Cetuximab and panitumumab are commercially available mAbs for EGFR, targeting colon, head, and neck cancers. Contrarily, TKIs target the intracellular tyrosine kinase domain and bind to the ATP-binding pocket, preventing autophosphorylation of the receptor and activation of intracellular signaling pathways. TKIs can bind to the ATP binding pocket reversibly or irreversibly. Reversible inhibitors compete with the ATP for binding pocket, while irreversible covalently bind to the kinase active site and show prolonged clinical effects. The first generation of TKIs, gefitinib and erlotinib, are Food and Drug Administration (FDA) approved inhibitors with great results in NSCLC, breast, head and neck, and ovarian cancers. However, various EGFR mutations and other mechanisms of resistance to treatment led to the II, III and IV generation of TKIs.

1.5.8.1 Cetuximab

Cetuximab (Erbitux^{TM}) is a human/murine chimeric monoclonal antibody, which binds EGFR to its extracellular domain, and competitively inhibits the binding of EGF-like ligands (Figure 11) [178]–[180]. The binding of cetuximab to the EGFR prevents proper exposure of the dimerization arm in subdomain II and prevents dimerization with the members of the ErbB family. Also, cetuximab binding prevents phosphorylation while promoting the internalization of the EGFR. Treatment reduces cancer cell proliferation by inducing a cell cycle arrest in the G1 phase and cancer cell metastasis by downregulation of matrix metalloproteases and pro-angiogenic factors. Moreover, it induces apoptosis by upregulation of some pro-apoptotic factors like Bax, downregulation of Bcl2 and activation

of caspases. Treatment with the Cetuximab antibody also induces the recruitment of immune cells and antibody-induced cytotoxicity.

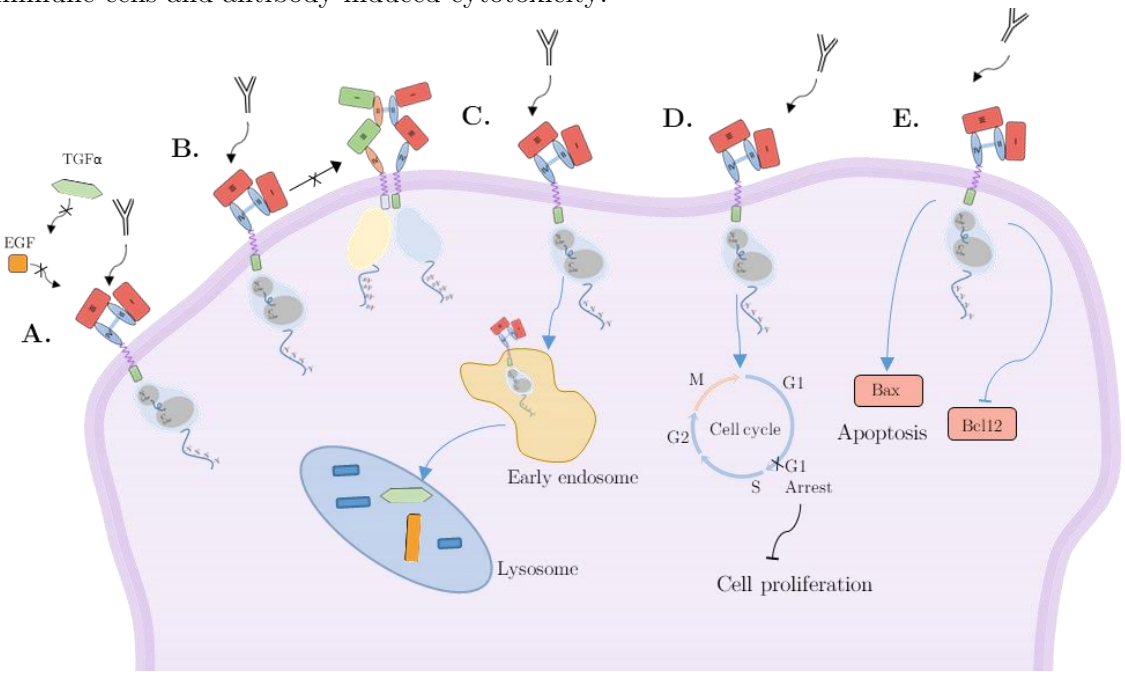


Figure 11: Mechanisms of action of Cetuximab binding to the EGFR: A. Blocks binding of the EGF-like ligands; B. Prevents dimerization with the ErBb family members; C. Promotes receptor internalization and degradation; D. Induces cell cycle arrest; E. Induces apoptosis [178].

Cetuximab was approved by the Food and Drug Administration (FDA) and the European Medicines Agency (EMA) in 2004. Cetuximab treatment showed promising antitumor activity in clinical trials as monotherapy and combined with radiation and/or chemotherapy in cancers like metastatic colorectal cancer and head and neck squamous cell carcinomas.

1.5.8.2 Erlotinib

Erlotinib (Tarceva™) is a first-generation TKI, chemically known as N-(3-rthynylphenyl)-6,7-bis(2-methoxyethoxy)-4-quinazolinamine, monohydrochloride [178], [181], [182]. Erlotinib selectively and reversibly inhibits EGFR activity by competing with ATP for the ATP binding pocket within the EGFR kinase domain. After Erlotinib binding, EGFR cytoplasmic tyrosines cannot be phosphorylated, and intracellular signaling was not initiated. Erlotinib was approved for treating NSCLC patients by the FDA in 2004 and by EMA in 2005. In combination with gemicitabine, erlotinib is also used in pancreatic cancer treatment.

1.5.9 Resistance to EGFR inhibitors

Both treatments, mAbs and TKIs, were approved and demonstrated to be effective anti-cancer drugs with acceptable toxicity. Unfortunately, patients who initially responded well to EGFR inhibitors almost inevitably developed resistance.

Several mechanisms of resistance to anti-EGFR antibodies have been reported [183]. In colorectal cancer, mutations in KRAS and NRAS are known to induce resistance to anti-EGFR antibodies [184], [185]. Resistance could be achieved with bypass signalling,

activated during the treatment. Some activated bypass signallings are amplification or mutations of other RTKs like HER2, MET and PDGFR and activation of alternative pathways like AXL, MEK and Src family kinases [186]–[188]. Expression of EGFRvIII mutant was also shown to induce Cetuximab resistance [189]. Cetuximab also failed to prevent ligand-independent signaling in cells with full-length EGFR overexpression.

Additionally, intracellular and further nuclear translocation of the EGFR during Cetuximab treatment was also proposed as one of the possible mechanisms for cancer cell survival. Moreover, nuclear STAT3 was proposed as a mediator of resistance. In patients with metastatic colorectal cancer, only 13% of patients with positive nuclear staining of phospho-STAT3 in tumour tissue responded to Cetuximab treatment, while 40.8% of patients without nuclear phopho-STAT3 staining responded to the treatment [190].

EGFR resistance to TKIs can be classified into four groups: (i) secondary EGFR gene mutations; (ii) activation of alternative pathways; (iii) resistance to apoptotic cell death, and (iv) phenotypic transformations. Among secondary mutations, T790M, a mutation in the ATP binding pocket, is the most frequent cause of resistance [191]–[193]. This mutation causes conformational change and a decrease in the binding of first-generation TKIs. Several more mutations in the ATP binding pocket, named gatekeeper mutations, are found to be involved in EGFR resistance [194], [195]. To overcome resistance to gatekeeper mutations, second and third generation of TKIs have been developed.

Another cause of TKI resistance, like in mAbs, is the development of several alternative pathways. Resistance was frequently observed when the hepatocyte growth factor receptor (MET) gene was amplified [196]. Moreover, combining EGFR and MET inhibitors was proposed to overcome resistance [197]. Upregulation of some other members of the ErbB family, like HER2 and HER3, was also connected with resistance to TKIs [198]–[200]. RTKs like AXL, insulin-like growth factor receptor and fibroblast growth factor receptor are also known as bypass signallings involved in resistance [201]–[203]. Among downstream signalling pathways, MAPK and PI3K/Akt were found to have an important role in resistance. Moreover, enhanced activity of STAT3 was also connected with EGFR TKIs resistance. Inhibition of STAT3 during the treatment was proposed to overcome resistance [190].

TKIs induce cell death mainly through apoptosis. Usually, a pro-apoptotic member of the Bcl-2 family, Bcl-2-like 11 (BIM), is phosphorylated at Ser69 by ERK and degraded through the ubiquitin/proteasome pathway. However, TKI treatment inhibits ERK and BIM degradation, causing apoptosis by increasing BIM intracellular concentration. Tumours with decreased expression of BIM were more resistant to TKIs [204]–[206].

Altogether, EGFR signalling and resistance to EGFR inhibitors in cancer are characterized by the complicated genetic network, communication with other RTKs, and constitutive activation of various downstream molecules. Therefore, a better understanding of EGFR signalling is required to overcome their involvement in tumour progression and resistance to treatment.

Chapter 2

Aims and Hypothesis

EGFR was recently identified as a novel substrate of cysteine cathepsins. This dissertation aims to determine the exact position of this interaction and whether cleavage affects the activation of EGFR and its downstream signalling cascade. Also, it is important to investigate how truncation of EGFR ectodomain influences different processes of cancer cells, such as migration, invasion and proliferation. Moreover, examining whether this cleavage contributes to anticancer drug resistance is necessary. Additionally, this dissertation aims to provide new insight into the roles of cysteine cathepsins in cancer progression and to help us determine the novel molecular mechanism through which proteolysis influences cancer development.

In the course of our work, we tested the following hypotheses:

- Cathepsin L cleaves the EGFR ectodomain as a specific site creating a stable proteolytic product.
- Cathepsin-based truncation of EGFR increases its autophosphorylation, distinct from ligand-dependent autophosphorylation.
- The presence of truncated EGFR activates a downstream phosphorylation cascade of signaling proteins.
- Presence of truncated EGFR changes cellular physiology.
- The presence of truncated EGFR influences cancer cell resistance to anticancer drugs.

Finally, a better understanding of the interaction between cathepsins and EGFR could provide important insight into signalling mechanisms that govern various aspects of cancer development. Furthermore, knowledge obtained in the proposed thesis has the potential to open new venues in the development of anticancer therapy.

Chapter 3

Materials and Methods

3.1 Materials

3.1.1 Chemicals

- ABC (ammonium bicarbonate)
- AcD3-NHS, made in house [207]
- ACN (acetonitrile)
- BrdU cell proliferation ELISA (Roche)
- Bradford reagent (Bio-Rad)
- BSA (Sigma-Aldrich)
- DMSO (Merck)
- DPBS, (Biowest)
- DTT (Sigma-Aldrich)
- E-64 (Bachem)
- EDTA (Serva)
- FA (formic acid)
- FBS (Sigma-Aldrich)
- Glutamine (Sigma-Adrich)
- Hanks based enzyme-free dissociation solution (Millipore)
- IAM (iodoacetamide)
- Methanol (Merck)
- NaCl (Serva)
- NP-40 (Roche Diagnostics)
- Penicillin/Streptomycin (Sigma-Adrich)
- Prestained Protein Ladder PageRuler (MBI Fermentas)
- PolyJet™ in vitro DNA transfection reagent (SignaGen Laboratories)
- Phosphatase arrest cocktail (G-Biosciences)
- Protease inhibitor cocktail (Sigma)
- SDS (Thermo Fisher Scientific)
- Sequencing-grade porcine trypsin (Promega)
- TEMED (Merck)
- TRIS (Serva)
- Triton X-100
- Trypan Blue (Sigma-Aldrich)
- TrypleSelect trypsin (Gibco, Invitorgen)
- Tween-20 (Sigma-Aldrich)
- Q5 hot start DNA polymerase (New England BioLabs (NEB))

- Western Blotting Detection Reagents ECL™ (GE Healthcare)
- Urea

3.1.2 Ligand and Inhibitors

- Cetuximab (Selleck Chemicals)
- E-64 (2S,3S)-trans-epoxysuccinyl-L-leucylamido (4-guanidino) butane)(Peptide Institute)
- EDTA (ethylenediaminetetraacetic acid)
- Epidermal growth factor EGF (Sigma-Aldrich)
- Erlotinib (Selleck Chemicals)
- PMSF (phenylmethylsulfonyl fluoride)

3.1.3 Commercial kits

- 12.5% SDS-PAGE Precast gel (Lonza)
- Human Phospho-Kinase Array (R&D System)
- Annexin V-FITC Apoptosis Detection Kit (eBioscience™, Invitrogen)
- BrdU Cell Proliferation ELISA Kit (chemiluminescent) (Abcam)

3.1.4 Buffers, solutions and media

- PBS

6.6 mM Na2HPO4
13.4 mM NaH2PO4
150 mM NaCl
pH 7.4

- TBS

20 mM Tris
500 mM NaCl
pH 7.5

- TBST

TBS
0.05 Tween 20

- Lysis buffer 1

20mM TRIS/HCl pH 7.4
150mM NaCl
1mM EDTA
1mM Na-orthovanadate
10 mM NaF
10mM β-Glycerophosphate
1% (v/v) NP-40
10% glycerol

- Lysis buffer 2
50 mM TRIS pH 7.8
150 mM NaCl
1 mM EDTA
1% NP-40
0.1% SDS
0.5% Na-deoxycholate

- Solution for SDS-PAGE stacking gel
0.625 mL 40% acrylamide-bisacrylamide
1.25 mL 0.5 M Tris/HCl (pH 6.8)
3.075 mL H2O
0.1 mL 10% SDS
0.02 mL 10% APS
0.01 mL TEMED

- Solution for SDS-PAGE (10%) resolving gel
2.5 mL 40 % acrylamide-bisacrylamide
2.5 mL 1 M Tris-HCl (pH 8.8)
4.9 mL H2O
0.1 mL 10% SDS
0.04 mL 10% APS
0.02 mL TEMED

- SDS-PAGE loading buffer (6x)
375 mM Tris-HCl
6% SDS
48% glycerol
100 mM DTT
0.03 % bromophenol blue
pH 6.8

- SDS-PAGE buffer
25 mM Tris
192 mM L-Gly
0.1 % SDS

- Western transfer buffer
25 mM Tris
192 mM L-Gly
15 mM SDS
10% methanol

- Laemmli buffer (4x)
240 mM Tris-Cl
8% SDS
40% glycerol
20% β-mercaptoethanol
0.04% bromphenol blue

- Staining solution
  0.5% Coomassie Brilliant Blue R-250
  50% methanol
  10% acetic acid

- Destaining solution
  30% ethanol
  10% acetic acid

- Reduction buffer
  10 mM DTT
  25 mM ABC

- Alkylation buffer
  55 mM IAM
  25 mM ABC

- Trypsinisation solution
  25 mM ABC
  1 µg sequencing-grade porcine trypsin

- Extraction solution
  50% ACN
  5% FA

- Culture media
  DMEM
  10% FBS
  1% penicillin/streptomycin
  1% glutamine

- Freezing media
  cDMEM
  40% FBS
  10% DMSO

3.1.5 Recombinant proteins

- Recombinant human cathepsin L was expressed in Pichia Pastoris, methylotrophic yeast expression system, and purified using a protocol previously described in [**208**].

3.1.6 Plasmids

- pcDNA4HisA (Invitrogen)
- pcDNA4TM myc-His(pcDNA4) (Thermo Fisher Scientific)

3.1.7 Antibodies

- C-Myc (#R950-25, Invitrogen)
- cytosolic EGFR (#4267, Cell Signaling Technology)
- EGFR Leu25-Ser645 (#MAB1095, R&D System)

- EGFR Tyr1173 (#AF1095, R&D System)
- EGFR Tyr1086 (#MAB8967, R&D System)
- EGFR Tyr1068 (#22375, R&D System)
- EGFR Tyr1045 (#MAB3570, Cell Signaling Technology)
- EGFR Tyr1045 (#MAB3570, Cell Signaling Technology)
- EGFR Tyr1197 (#PA5-37553, Thermo Fisher Scientific)
- GAPDH (#2118S and #97166S, Cell Signaling Technology)
- STAT3 (#ab68153, Abcam)
- STAT3 Y705 (#ab76315, Abcam)
- STAT3 S727 (#ab86430, Abcam)
- HRP-conjugated goat anti-rabbit IgG (#111-035-045, Jackson Immunoreshearch)
- HRP-conjugated goat anti-mouse IgG + IgM (#115-035-068, Jackson immunoresearch)
- HRP-conjugated donkey anti-goat IgG (#ab6885, Abcam)

3.1.8 Mammalian cell lines

- HeLa, human cervical adenocarcinoma cell line
- MDA-MB-231, human breast adenocarcinoma cell line

3.1.9 Equipment

- 1.5 mL and 2 mL micro tubes (Sarstedt)
- 15 mL and 50 mL centrifuge tubes (BD Biosciences)
- 96-well microtiter cell culture black plate with transparent bottom (Corning)
- Amicon 0.5 Ultra Centrifugal filters (Millipore)
- Automatic pipettes (Eppendorf)
- Cell line incubator (Binder)
- Centrifuges (Eppendorf)
- Concentrator 5301 (Eppendorf)
- Empore/C18 disks (Varian)
- FACSCalibur flow cytometer
- Gel imaging system G:Box (Syngene)
- G: BOX Chemi XRQ gel dock system (Syngene)
- Laminar air flow cabinet for cell culture BSB 4A (Gelaire Flow Laboratories)
- Nitrocellulose membrane
- Orbitrap LTQ Velos mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nanoLC II HPLC unit (Thermo Fisher Scientific)
- pH-meter MP 220 (Mettler Toledo)
- Power supply Western Lightning Volt™ Power Supply OSP-300 (Thermo Scientific)
- PHOS-Select iron-affinity gel (Merck)
- SigmaPrep spin columns (Sigma-Aldrich)
- Tecan Infinite M1000 plate reader (Tecan)
- TiO2 beads (Titansphere Phos-TiO 10 μ m, GL Science)
- Vertical Electrophoresis Lab Apparatus MiniProtean Tetracell (Biorad)
- Western transfer apparatus TE22 Mighty Small Transphor Unit (GE Lifesciences)

3.2 Methods

3.2.1 Cell Culture

Cell lines were chosen based on their endogenous epidermal growth factor receptor expression. HeLa is easy to transfect cell line and has low endogenous expression levels of the EGFR, while the MDA-MB-231 cell line expresses high levels of endogenous EGFR.

3.2.1.1 Cell growing and passaging

The human HeLa cervical cancer cell line and human MDA-MB-231 breast cancer cell line were grown to confluence in DMEM supplemented with 10% fetal bovine serum FBS, 1% penicillin/streptomycin, and 1% glutamine. Cell lines were grown in a humidified incubator at 37 ºC and 5% CO2. When cells reached a confluency of about 80-90%, cells were washed twice with PBS, and TrypleSelect trypsin was added to detach them. After incubation, around 5 min at 37 ºC, cDMEM was added to inactivate trypsin. Cells were diluted 1:10 in cDMEM and then transferred to the new plate.

3.2.1.2 Cryopreservation of the cells

After trypsinization and trypsin inactivation, the cell suspension was centrifuged for 5 min at 400 rcf. After centrifugation, supernatant was removed, and cell pellet was resuspended in freezing media to 1-2 x 10^6 cell/ml concentration. The 1 ml of cell suspension was then transferred to the vial. A Cryofreezing container was used to transfer the vials to the -80 ºC freezer.

3.2.1.3 Thawing of the cells

The cell vials taken from the liquid nitrogen or -80 ºC freezer were first thawed in a water bath for around 1 min at 37 ºC and then resuspended in 10 ml of cDMEM. The cell suspension was then centrifuged 5 min at 400 rcf, and the supernatant containing DMSO was removed. The cell pellet was resuspended in cDMEM and transferred to the plate.

3.2.2 Construction of plasmids and transfection methods

EGFR cDNA was obtained from the Addgene plasmid 11011 (gift from Matthew Meyerson) (Greulich et al., 2005). cDNA was amplified by Q5 hot-start DNA polymerase using the following nucleotide primers: TT-CTC-GAG-ATG-CGA-CCC-TCC-GGG (sense primer containing the XhoI restriction site and starting with the EGFR signal peptide), and AA-CTC-GAG-TGC-TCC-AAT-AAA-TTC-ACT-GCT (antisense primer containing XhoI restriction site). cDNA was cloned into pcDNA4™myc-HisA. Truncated EGFR cDNA, which encodes a protein that matches the cathepsin cleavage product, was generated by adding the EGFR signal peptide to truncated EGFR cDNA, starting with G225. The prepared cDNA construct was cloned into the pcDNA4™myc-HisA vector in the same way as the full-length EGFR.

HeLa cells were transiently transfected with pcDNA4™myc-HisA, pcDNA4 EGFR, and pcDNA4 t-EGFR using PolyJet™ in vitro DNA transfection reagent following the manufacturer's instructions. The cells were incubated with a transfected reagent for a minimum of 6h, after which the medium was replaced with the fresh, completed medium. Depending on the experiment, the transfected cells were washed with Dulbecco's Phosphate Buffered Saline, DPBS, and either collected for cell lysis or further treated.

3.2.3 Cell treatments

3.2.3.1 EGF treatment

The HeLa cell line was treated with EGFR ligand EGF. Before treatment with EGF ligand, HeLa cells were starved overnight in a serum-free medium. After starvation, the cells were incubated with EGF (10 or 100 ng/ml) for 8 min.

3.2.3.2 Erlotinib and cetuximab treatment

24h after transfection, HeLa cells expressing empty vector, full-length EGFR or truncated t-EGFR were starved overnight in a serum-free medium. Starved cells were treated with tyrosine kinase inhibitor (TKI) erlotinib (0 or 1μM) for 1 hour. After the treatment, cells were additionally treated with 0 or 100 ng/ml EGF for 8 min. Cells were lysed and prepared for Western blot analysis. The same experiment was performed using four different concentrations of erlotinib (0, 10, 100 or 1000 nM). For negative control, cells were treated with serum-free media containing DMSO. Similarly, after transfection of the HeLa cell line, cells were treated with serum-free or serum-free media containing cetuximab (50 μg/ml) for 48h. After the treatment, cells were additionally treated with o or 100 ng/ml EGF for 8 min. Cells were lysed and prepared for immunological analysis.

3.2.4 Cell lysate preparation

Cells were washed twice using DPBS and lysed on-plate using Lysis buffers 1 (20mM TRIS/HCl, 150mM NaCl, 1mM EDTA, 1% (v/v) NP-40, 10% glycerol, 1mM sodium orthovanadate,10 mM NaF, 10mM β-Glycerophosphate, pH 7.4). Samples were centrifuged at 16 000 × g for 15 min and 4 °C. Lysates were collected, and their protein concentrations were determined using the Bradford assay.

3.2.5 Bradford assay

The Bradford assay was used for the determination of the total protein concentration. This assay is based on the fact that the absorption maximum of Coomassie Brilliant Blue G-250 shifts from 465 nm to 595 nm after binding to aromatic amino acid residues. As an external standard, bovine serum albumin (BSA) was used in the 1 – 24 μg/mL concentration range. BSA stock solution was prepared in a concentration of 100 μg/mL. Appropriate volumes of BSA stock solution and diluted cell lysates (1:1000 dilution) were pipetted into clear 96-well plate wells, and dH2O was added to the final volume of 160 μL (Table 1). In each dilution, 40 μL of Bradford reagent was added, and A595 was measured using the Tecan plate reader Infinite M1000. The total protein concentration of the lysates was calculated from the standard curve acquired by measuring the BSA standard.

Table 1: Volumes used for standard curve mesurment:

BSA concentration (µg/mL)	Volume of the BSA stock solution (µl)	dH2O volume (µL)	Volume of the Bradford reagent (µL)
0	0	160	40
2	4	156	40
4	8	152	40
6	12	148	40
8	16	144	40
10	20	140	40
12	24	136	40
14	28	132	40
16	32	128	40
20	40	120	40
24	48	112	40

3.2.6 REAP protocol for cellular fractionation

Nuclear and cytoplasmic fractions were fractioned using Rapid Efficient And Practical (REAP) nuclear/cytoplasmic separation protocol (Figure 12) as previously described in [209]. HeLa cells were transfected with plasmid constructs pcDNA4, pcDNA4 EGFR or pcDNA4 t-EGFR, as previously explained. The cells were counted 32 hours after transfection, and the same number of cells per condition was used for further fractionation. Cells were washed in cold 1ml of PBS in a 1.5 ml micro-centrifuge tube, then centrifuged for 10s at 10000 rpm (pop-spin). The supernatant was removed and cell pellets were resuspended in 900 µL of ice-cold 0.1% NP-40 in PBS and triturated for 5 min using a 1ml pipet tip. For the whole cell lysate, 300 µL was removed, and 100 µL of 4 x Laemmli buffer was added and then kept on ice. The remaining samples were then pop-spined, and 300 µL of the supernatants were removed and saved as a cytosolic fraction. 100 µL of 4 x Laemmli buffer was added to the cytosolic fraction, and fractions were further boiled for 1 min. The remaining supernatants were removed, and pellets were resuspended in 1ml of ice-cold 0.1% N-40 in PBS and pop-spined. Supernatants were removed, and the remaining pellets were resuspended in 180 µL of 1 x Laemmli buffer, representing nuclear fractions. Whole-cell lysates and nuclear fractions were then sonicated. Samples were then loaded on SDS-PAGE, using 10 µL, 10 µL and 5 µL of whole cell lysate, cytoplasmic and nuclear fraction, respectively. After the electrophoresis, proteins were transferred to the nitrocellulose membrane and incubated with antibodies. As a control for nuclear fraction Histon H3 was used, while GAPDH was used for the cytoplasmic fraction.

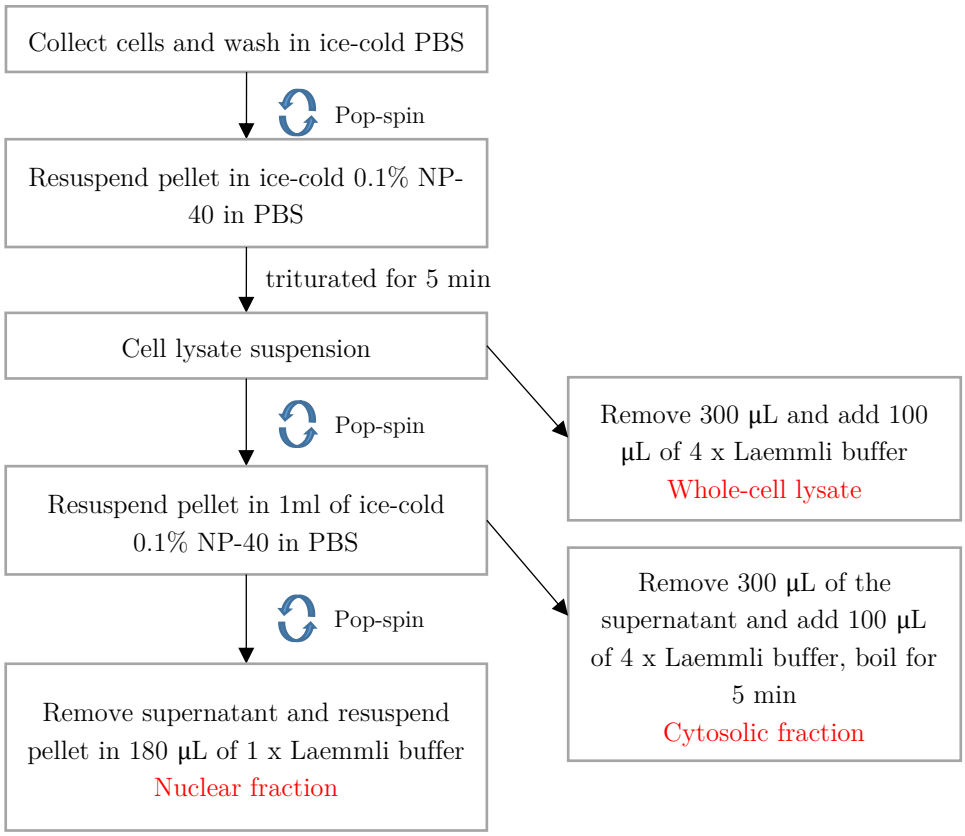


Figure 12: Schematic representation of the REAP protocol for cellular fractionation.

3.2.7 SDS-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis was used for protein separation. Equal amounts of proteins were boiled at 95 °C for 6 min in the presence of 6 x loading buffer. Proteins were separated on the 5% SDS stacking gel following the 10% SDS resolving gel. Electrophoresis ran in an SDS-PAGE buffer at a constant current of 30 mA/gel until the standards were separated. Pre-stained protein ladder, PAGE Ruler, was used as a standard.

3.2.8 Western transfer and immunological detection

The western transfer was used for protein transfer from the gel to the nitrocellulose membrane. The transfer ran for 3h at the current 250 mA in a Western transfer buffer containing 10% methanol. After the transfer, membranes were blocked in a 5% skim milk solution in TBST for 1 hour and then incubated with primary antibodies overnight at 4 °C. Membranes were washed with TBST 3 times for 15 min and then incubated with HRP-conjugated secondary antibodies for 1 h at room temperature. After additional washing steps, bands were visualized using enhanced chemiluminescent ECL Western Blotting Detection Reagent and G: BOX Chemi XRQ gel dock system.

3.2.9 Human Phospho-Kinase Array

HeLa cells were transfected with plasmid constructs pcDNA4, pcDNA4 EGFR or pcDNA4 t-EGFR, as previously explained. 32 hours after transfection, the completed growth

medium was replaced with a serum-free medium, and cells were starved overnight. Transfected HeLa cells were lysed and incubated with Phospho-Kinase Array membranes according to the manufacturer's instructions. In short, cell lysates were incubated with membranes from the array overnight at 4 °C, after which the membranes were washed and incubated with the cocktail of antibodies for 2 hours. After additional washing, membranes were incubated with Streptavidin-HRP (streptavidin conjugated to the horseradish-peroxidase) for 30 min and again washed three times. Dots were visualized using ECL Western Blotting Detection Reagent and G: BOX Chemi XRQ gel dock system. Pixel densities were determined using ImageJ's extension for the Dot Blot Analyzer. The average signal of the paired dots was determined, while the referent spots were used for alignment between different membrane sets. Relative phosphorylation levels for each dot pair were determined by dividing the EGFR or t-EGFR sample values with pcDNA4 sample values.

3.2.10 Cell viability assay

Cell viability was determined using the combination of Annexin V and propidium iodide (PI) staining. This protocol enables us to distinguish between viable, apoptotic and necrotic cells. Annexin V binds phosphatidylserine (PS), a lipid normally restricted to the inner leaflet of the plasma membrane, but during apoptosis, it becomes exposed on the outer leaflet of the plasma membrane. This enables Annexin V to bind PS and, in this case, to label apoptotic cells. On the other hand, the ability of PI to stain cells depends on cell membrane permeability since it needs to pass through the plasma membrane, intercalate with nucleic acids and display fluorescence. PI can only label late apoptotic and necrotic cells, cells with disrupted cell membranes.

In our experiment, one day after transfection, HeLa cells were seeded in density $2x10^5$ cells per well in a 12-well plate. Cells were grown in DMEM-completed media or in the same media containing 10 μM Erlotinib for an additional 48h. Staurosporine-treated samples were treated with 1 μM Staurosporine for the last 16 hours. Cell suspension and adherent cells were collected and treated with Annexin V and PI. Cell viability was investigated by flow cytometry using a previously described protocol [210]. Samples were analysed using the FACSCalibur flow cytometer and FlowJo software. Annexin V and PI-negative cells were considered viable.

3.2.11 Cell proliferation assay

Proliferation was analyzed using the BrdU (5-bromo-2'-deoxyuridine) proliferation assay. BrdU is a synthetic analogue of thymidine which incorporates into newly synthesized DNA in proliferating cells. The BrdU can be detected with specific antibodies and used for the proliferation measurement.

24 hours after transfection, 5000 cells/well were seeded into black cell culture 96-well plates with a clear bottom in serum-free cDMEM or cDMEM for 40 hours. BrdU nanoparticles were added in the last 20 hours, and a proliferation assay was then used according to the manufacturer's instructions. In short, cells were fixed, anti-BrdU antibodies were added, and after 90 min, cells were washed, and the substrate was added. Luminescence was then measured with the Tecan Infinite M1000 plate reader.

3.2.11 Cell migration assay

Migration assay was performed using trans-well migration protocol, also known as Boyden chamber assay [211]. One day after transfection, cDMEM culture media was replaced with serum-free cDMEM, and cells were incubated for additional 24 hours. After incubation,

cells were detached using Hanks-based enzyme-free dissociation solution, counted, and 5 x 10^4 cells were seeded into the 24-well insert, trans-well with 8µm porose membrane. Cell migrated from the trans-well containing serum-free cDMEM, through the porose membrane of the trans-well, towards the well containing 2%, or 10% FBS cDMEM culture media, for 24 hours (Figure 13). In the erlotinib-treated sample, cells migrated from the serum-free cDMEM culture media containing 0.01 µM erlotinib towards the 10% FBS cDMEM culture media. The trans-wells were then washed using PBS, and the cells that had not migrated from the top of the membrane were carefully removed using cotton tips. Migrated cells were fixed by adding 600-700 µl of the 70% ethanol into a well for 10 minutes. After 10 minutes, ethanol was removed, and trans-wells were left to dry. Cells were then stained in 0.2% Crystal Violet for 5 min at room temperature, after which the excess staining was washed using distilled water. After drying the trans-well, the migrated cells were observed under the microscope. The number of cells which migrated on the other side of the trans-well was counted using ImageJ in three different fields of view to acquire an average sum of the migrated cells.

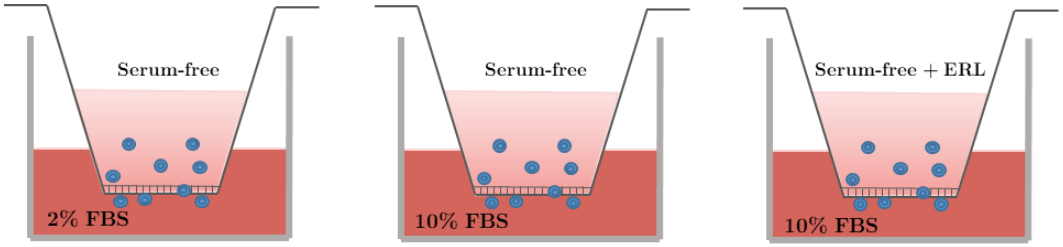


Figure 13: Migration assay. Experimental workflow of the assay and conditions in which cells migrated.

3.2.12 Treatment of cells with recombinant cathepsin L

MDA-MB-231 cells were grown to the confluence, after which the cell plates were washed twice with DPBS. The cells were then detached using Hanks-based enzyme-free dissociation solution (Millipore, Billerica, MA, USA). Collected cells were centrifuged at 200 × g for 5 min and rinsed in PBS buffer (pH 6.0). Ten million cells per condition were incubated with 1 µM cathepsin L in a PBS buffer (pH 6.0) containing 0.5 mM DTT, 0.5 mM PMSF (phenylmethylsulfonyl fluoride) and 1 mM EDTA (ethylenediaminetetraacetic acid) at 37 °C for 5, 15, 30, and 60 min. For negative control, cells were incubated with E64 (25 µM of final concentration) inhibited cathepsin L in the same buffer. After incubation, the residual activity of recombinant cathepsin L was blocked with the addition of E64 (20 µM of final concentration). Using centrifugation at 500 × g for 5 min, cell pellets and supernatants were separated and collected. Collected supernatants were centrifuged once more at maximum speed for 5 min. Again, the supernatants were collected and prepared for Western blotting analysis.

3.2.13 Determination of cathepsin L cleavage site on EGFR ectodomain

MDA-MB-231 cells were grown to confluence and detached with Hanks-based enzyme-free dissociation solution (Figure 14). Thirty million of the MDA-MB-231 cells per condition were incubated for 1 h at 37 ºC in 500 µl of PBS (pH 6.0 supplemented with 0.5 mM DTT, 0.5 mM PMSF and 1 mM EDTA) with 0.2 µM recombinant human cathepsin L, or with E64 (25 µM of final concentration) inhibited cathepsin L as a negative control. After incubation, the residual activity of recombinant cathepsin L was blocked with the addition

of E64 (20 µM final concentration). Cells were centrifuged at 500 × g, for 5 min, and supernatants were collected and centrifuged once more at maximum speed for 5 min. Amicon 0.5 Ultra Centrifugal filters) were used to concentrate supernatants to the final volume of 40 µl. The same amount of supernatants was separated on 12.5% SDS-PAGE Precast gel.

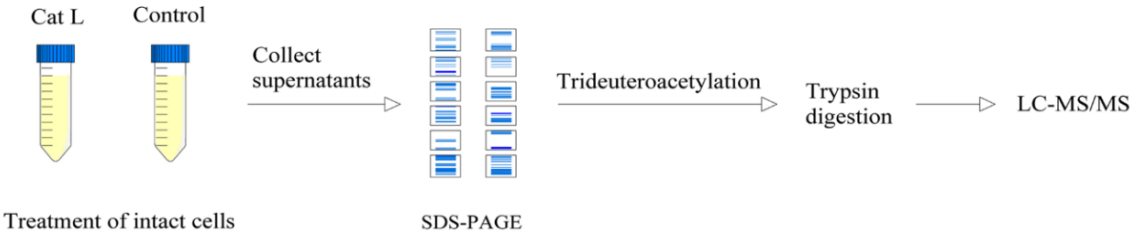


Figure 14: Experimental workflow representing determination of cathepsin L cleavage site on EGFR.

Separated proteins were fixed and stained using Coomassie Brilliant Blue for 1 h, followed by the overnight-incubation with destaining solution. The gel was rinsed in water, and sample lanes were cut from the gel. Each lane was cut into six bands, representing six protein samples per condition. The bands were cut into 1 mm^3 gel cubes and further destained using 150 µl of 50% ACN (acetonitrile) in 25 mM ABC (ammonium bicarbonate) for 30 min, at 25 ºC and 1,200 rpm, twice. The gel cubes were dehydrated in 100% ACN and vacuum-dried. Free N-termini, created by cathepsin L cleavage, were labelled using N-hydroxysuccinimide ester of trideutero-acetate (AcD3-NHS) [207]. AcD3-NHS solution was prepared by dissolving 20 mg of AcD3-NHS in 2 ml of PBS buffer (0.1 M, pH 8.5). The gel cubes were incubated in 150 µl of AcD3-NHS solution for 1 hour, after which the gel cubes were dehydrated, and the labelling step was repeated. To reduce cysteine residues, the dehydrated gel cubes were incubated with 150 µl of the reduction buffer for 1 h at 56 ºC, while alkylation of the same residues was performed using 150 µl of the alkylation buffer for 30 min at 30 ºC in the dark. The unreacted iodoacetamide(IAM) was quenched by an additional incubation with 150 µl of 20 mM DTT in 25 mM ABC for 30 min at 37 ºC. The gel cubes were then washed with 150 µl MS grade water and again completely dehydrated using 150 µl of 100% ACN and vacuum-dried. The proteins were digested in-gel. The gel bands were incubated with around 1µg of sequencing-grade porcine trypsin in 25 mM ABC at 37 ºC overnight. The next day, trypsin was inactivated by adding formic acid FA (final concentration of 5%), and peptides were extracted from the gel by adding 150 µl of extraction solution. The samples were further concentrated by vacuum drying to 50 µl and desalted using a C18 tip. The C18 tip was prepared by loading 4 Empore/C18 disks into the 200-µl pipet tip and further activated as described in [26]. Samples were applied to the C18 tip and washed using 0.1 % FA. The peptides were eluted using 300 µl 60% ACN with 0.1% FA in three subsequent steps. Acetonitrile was removed by vacuum drying the samples to a total volume of 12 µl and sent for the LC-MS/MS analysis.

3.2.14 Phosphoproteomics

HeLa cells transiently expressing pcDNA4, pcDNA4 EGFR, or pcDNA4 t-EGFR were lysed on plate, 48 h after transfection (using the Lysis buffer 2, including phosphatase arrest cocktail). The cell lysates were transferred to the test tube and left for 30 min to shake on ice. After shaking, samples were centrifuged at 14 000 × g, for 15 min at 4 ºC, and the whole cell lysates were collected. The total protein concentration was determined using the

Bradford assay. 2mg per sample was used for further phospho-peptide analysis. Urea was added to the samples in the final concentration of 6M, and proteins were further reduced by incubating samples with a DTT in a 10mM final concentration at room temperature for 1 h. IAM in a final concentration of 32mM was incubated with samples for 1 h at room temperature in the dark to alkylate cysteine residues. To suppress unreacted iodoacetamide, samples were incubated with 38mM DTT for 1 h at room temperature. Urea was diluted to a concentration of 1M using 0.1 % SDS with 50 mM Tris at pH 8.0. Sequencing grade trypsin was added to the samples in a 1:100 trypsin to substrate ratio (w/w), and incubated at 37 ºC overnight. To inactivate trypsin, the pH of the samples was lowered to pH 6.0 by adding acetic acid. Also, Urea was added to the final concentration of about 6M. The C18 stage tips were prepared by loading 20 C18 discs into the 200 µl pipet type. After the C18 tips activation, the samples were applied to the C18 tips and washed with 0.1 % (v/v) water solution of FA. The peptides were eluted using 300 µl 60% ACN with 0.1% FA in three subsequent steps, after which samples were concentrated by vacuum drying.

The phospho-peptides were enriched using, subsequently, PHOS-Select iron-affinity gel and TiO2 beads. The concentrated samples were diluted using 0.1% TFA in 50% ACN to the concentration of approximately 4 µg/µl. The samples were then added to the tubes containing PHOS-Select iron affinity gel equilibrated in the same buffer, vortexed, and left to mix end over end for 1 h at room temperature. After incubation, samples with iron affinity gel were transferred into the SigmaPrep spin columns and spin down in a centrifuge. The flow-through was collected as an unbound fraction. After several washing steps with 0.1% TFA in 35% ACN, the peptides were eluted from the iron affinity gel using 150mM NH4OH, 25% ACN. Both elution and unbound fractions were concentrated using vacuum drying. Concentrated eluted peptides were loaded on the C18 tip (20 discs), while concentrated unbound fractions were added to the TiO2 beads.

The unbound fractions were diluted using a buffer containing 1M glycolic acid, 80% ACN, and 5% TFA and then incubated with TiO2 beads for 1 h at room temperature. The TiO2 beads were previously equilibrated in 80% ACN and 0.4% TFA. The TiO2 beads were centrifuged for 2 min at 500 × g, the supernatants were transferred to the fresh tube and incubated again with equilibrated beads for another hour at room temperature. TiO2 beads were transferred to the C18 tips in the same way after the second incubation. Tips were washed with 100µl of 30% ACN 0.5% TFA and then with 80% ACN 0.5 TFA.

The eluted fractions were fractioned and eluted from the C18 tip subsequently by adding:

- 100 µ l of 0.1% FA and centrifugation at 600 × g for 5 min;
- 100 µ l of 2% ACN 0.1% FA centrifuged at 900 × g for 5 min;
- 100 µ l of 5% ACN 0.1% FA centrifuged at 1100 × g for 5 min;
- 100 µ l of 8% ACN 0.1% FA centrifuged at 1400 × g for 5 min;
- 100 µ l of 10% ACN 0.1% FA centrifuged at 1400 × g for 5 min;
- 100 µ l of 40% ACN 0.1% FA centrifuged at 1400 × g for 5 min.

The eluted fractions were concentrated by vacuum drying to a volume of 20 µl and sent for the LC-MS/MS analysis.

The unbound fractions were fractioned and eluted from the C18 tip subsequently by adding:

- 100 µ l of 15% ammonia at 600 × g for 5 min;
- 100 µ l of 15% ammonia 2% ACN centrifuged at 900 × g for 5 min;
- 100 µ l of 15% ammonia 5% centrifuged at 1100 × g for 5 min;
- 100 µ l of 15% ammonia 8% ACN centrifuged at 1400 × g for 5 min;

- 100µl of 15% ammonia 10% ACN centrifuged at 1400 × g for 5 min;
- 100µl of 15% ammonia 40% ACN centrifuged at 1400 × g for 5 min.

The unbound fractions were concentrated by vacuum drying to a volume of 20 µl, and sent for the LC-MS/MS analysis.

3.2.15 Analysis by LC-MS/MS

The samples were analyzed using an Orbitrap LTQ Velos mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) coupled to an EASY-nanoLC II HPLC unit (Thermo Fisher Scientific, Waltham, MA, USA). The peptide samples dissolved in solvent A (0.1 % FA (v/v) in water) were loaded on a C18 trapping column, Proxeon EASY-ColumnTM (Thermo Fisher Scientific), and separated onto C18 PicoFritTM AQUASIL analytical column (New Objective, Woburn, MA, USA). The peptides were eluted with a 90 min linear gradient of 5 to 50% of solvent B (100% CAN in 0.1% FA) at a 300 nl/min flow rate. Using Orbitrap mass analyzer, the full MS mass spectra were obtained at a resolution of 30,000, with a mass range of 300-2,000 m/z. The nine most intense ions from full MS spectra were fragmented by HCD fragmentation. MS/MS spectra were recorded at a resolution of 7,500. For MS/MS fragmentation, only the precursor ions with assigned charge (>1) were chosen. Dynamic exclusion was enabled with a repeat count of 1, repeat duration of 30s, and exclusion duration of 30s.

3.2.16 Proteomic data analysis

Tandem mass spectra were analyzed using the PEAKS studio software (version 10.6, Bioinformatics Solutions Inc., Waterloo, Canada). The search was performed against the Homo sapiens UniProt database, using trypsin cleavage with a maximum of 3 mixed cleavages. Parent and fragment mass error tolerance was 15.0 ppm and 0.02 Da, respectively. Carbamidomethylation was used as a fixed modification, and for variable modifications oxidation, acetylation and trideutero-acetate (heavy, +3amu) labelling were used. The resulting protein list was filtered by 1% false discovery rate (FDR).

In the phosphoproteomic experiment, raw data were analyzed using MaxQuant software (version 1.6.17.0, Max-Planck Institute for Biochemistry, Martinsried, Germany). The search was performed using the Andromeda search engine against the Homo sapiens UniProt database. Besides default parameters, STY phosphorylation was added as a variable modification. The label-free quantification (LFQ) was enabled. The Perseus software (version 1.6.15.0) was used for further analysis of MaxQuant data. Before further analysis, potential contaminants, reverse hits, and phosphosites with a localization probability of less than 0.75 were removed. Data were further filtered in a way that phosphosites that were not present in at least two out of three biological samples were removed. The filtered phosphoproteomic data were used for the t-test, and the phosphosites with two-fold difference and p-value less than 1% were considered significant.

3.2.17 Statistical analysis

GraphPad Prizm (Dotmatics) software was used for visualization and statistical analysis. Quantitative data are presented as means ± standard error. Error bars show the standard deviations based on duplicate values of each dataset for the Human Phospho Kinase Array, representing two technical values from one biological sample. In cell viability, cell proliferation and cell migration experiment, error bars show the standard deviations based on triplicate values of each dataset, from three biological samples.

Differences were compared using two-way ANOVA test. When p-values were 0.05 or less, differences were considered statistically significant.

Chapter 4

Results

4.1 Determination of Cathepsin Cleavage Site

To study the influence of the extracellular cysteine cathepsins on EGFR signaling, MDA-MB-231 cancer cells were treated with recombinant cathepsin L and the supernatant containing proteins released from the cell surface was collected (Figure 12). For the negative control, cells were treated with E64-inhibited cathepsin L.

Proteins in the supernatant were N-terminally labelled with trideuteroacetylation. Trideuteroacetylation was used to label free N-termini and differentiate them from the N-termini created by trypsin during digestion. Primary amines, amino groups at protein N-terminus and Lys side chains, were labelled, but also the *neo*N-terminus, newly created by cathepsin L cleavage. After labelling, proteins were digested and identified by mass spectrometry (MS).

Among the identified peptides (Table 2), we observed two deuteroacetylated N-termini at aminoacid residues L25 and G225, which indicates that these termini were not generated by trypsin. While proteolytic cleavage at L25 corresponds to the signal peptide cleavage site, cleavage at G225 is generated by cathepsin L. Furthermore, EGFR peptide with a non-tryptic C-terminus at H585 was also identified, which could represent the second putative cathepsin cleavage site.

Table 2: List of EGFR peptides identified by mass spectrometry. The peptides with cathepsin L cleavage site are marked with red lines. The table presents the significance score (-10lgP) value, the average sequence length, precursor mass error (ppm), mass over charge value (m/z), the charge state (z), and posttranslational modifications (PTM).

Peptide	-10lgP	Length	ppm	m/z	z	PTM
R.GK(+45.03)SPSDC(+57.02)C(+57.02) HNQC(+57.02)AAGC(+57.02)TGPR.E	98.75	20	2.0	7556427	3	Acetate labelling reagent (heavy, +3amu); Carbamidomethylation
R.FSNNPALC(+57.02)NVESIQWR.D	87.66	16	2.4	9679648	2	Carbamidomethylation
R.NYVVTDHGSC(+57.02)VR.A	71.77	12	1.1	7038285	2	Carbamidomethylation
R.NLQEILHGAVR.F	62.78	11	2.4	6253557	2	
R.GPDNC(+57.02)IQC(+57.02)AH.Y	58.74	10	2.0	5862357	2	Carbamidomethylation
R.IPLENLQIIR.G	58.66	10	2.9	6048734	2	
A.L(+45.03)EEK(+45.03)K(+45.03)VC(+57.02)Q GTSNK(+45.03)LTQLGTFEDHFLSLQR.M	43.03	29	4.7	11962971	3	Acetate labelling reagent (heavy, +3amu); Carbamidomethylation
R.G(+45.03)K(+45.03)SPSDC(+57.02) C(+57.02)HNQC(+57.02)AAGC(+57.02)TGB) PR.E	41.24	20	-0.9	7706503	3	Acetate labelling reagent (heavy, +3amu); Carbamidomethylation
R.ESDC(+57.02)LVC(+57.02)R.K	39.25	8	1.0	5197213	2	Carbamidomethylation

MS analysis of the supernatants showed that EGFR was only identified in the cathepsin-treated sample and that all identified peptides were located in the extracellular region of EGFR, mainly within domains I and II (Figure 15).

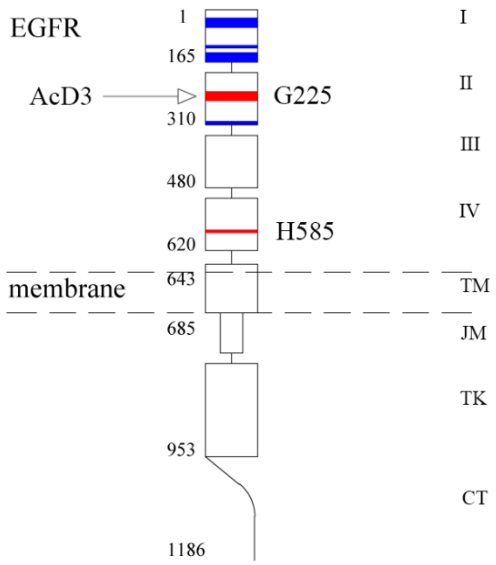
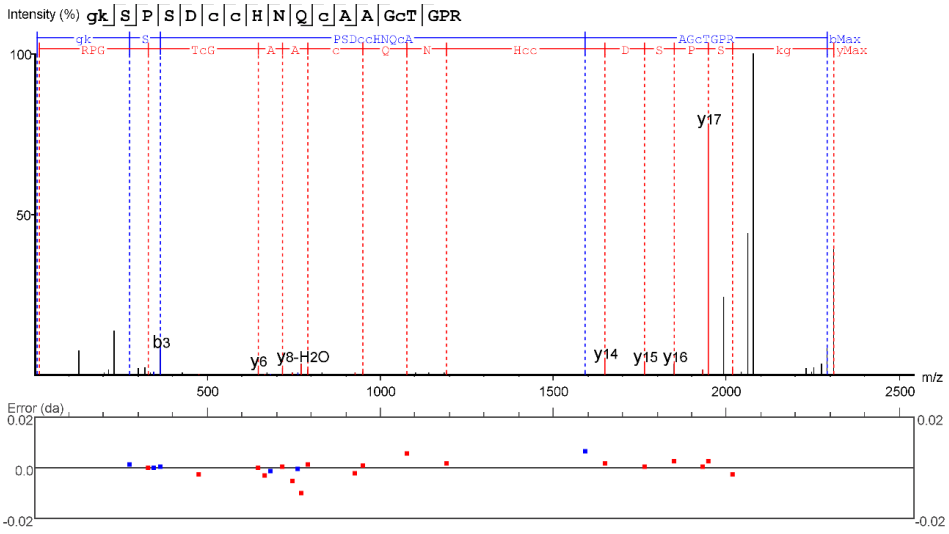


Figure 15: Graphical representation of peptides (red and blue lines) identified by mass spectrometry within the EGFR extracellular region. Red lines represent peptides created by cathepsin L.

Here we presented the MS/MS spectra of the peptides generated by cathepsin L (Figure 16, A and B). Tandem mass spectra were analyzed using PEAKS studio software. Parent and fragment mass error tolerance was 15.0 ppm and 0.02 Da, respectively.

A)



B)

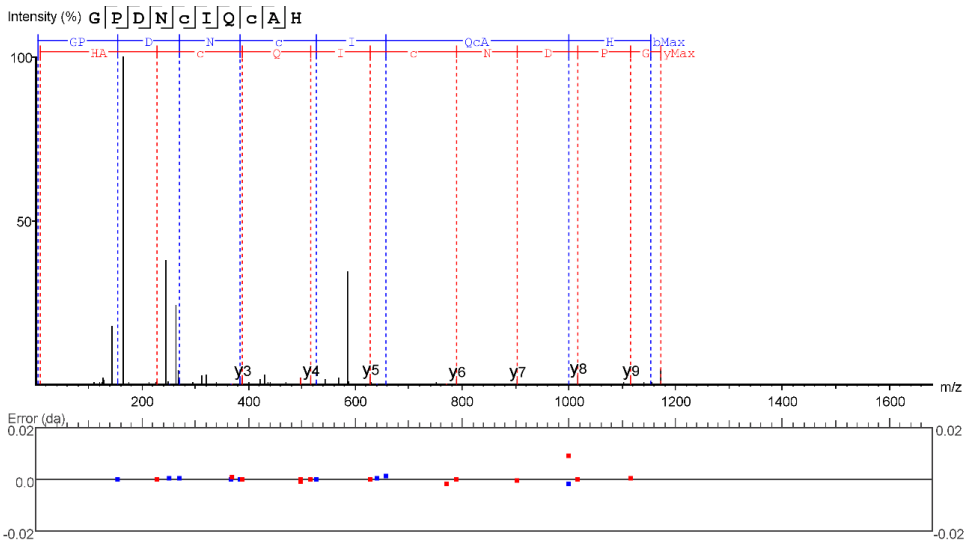


Figure 16: MS/MS spectra of the cathepsin L-generated peptides of EGFR. A) deuteroacteylated N-termini at amino acid residues G225, which indicate that these termini were not generated by trypsin, also deuteroacetylated at K226, B) second putative cathepsin L cleavage site, non-tryptic C-terminus at H585.

To further evaluate these results, immunological detection of cleaved EGFR ectodomain in the supernatant of cathepsin L-treated cancer cells was performed. This experiment revealed the formation of two stable EGFR fragments (Figure 17). The main fragment was observed at 40 kDa, while the less intense cleavage product was observed at 130 kDa. Cells were treated for 0, 5, 15, 30 and 60min, and consequently, the bend density of the EGFR fragment increased with the time of treatment. Molecular masses of both cleavage products observed on Western blot correspond to identified putative cleavage sites considering the determined glycosylation pattern of the EGFR ectodomain [212], [213].

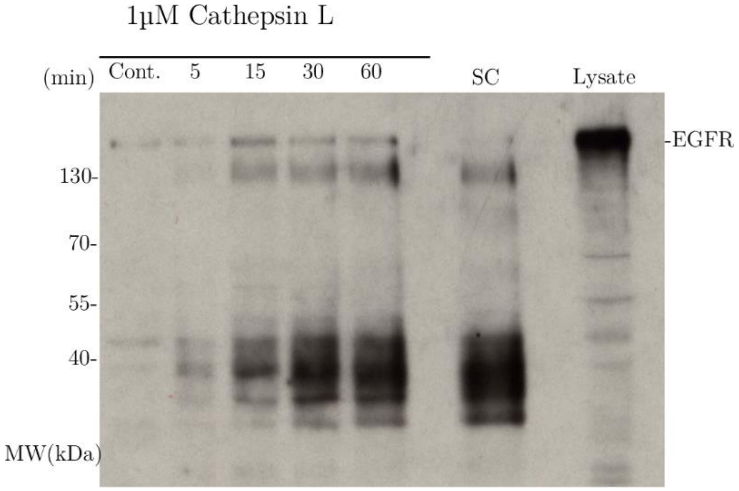


Figure 17: Immunological detection of EGFR ectodomain fragments in the supernatant of the MDA MB231 cells treated with cathepsin L. Supernatants were collected at 5, 15, 30 and 60 minutes, while cells treated with inhibited cathepsin L were used as negative control. The sample used for mass spectrometry determination of the cathepsin L cleavage site was

used as a shedding control (SC). The MDA-MB-231 cell lysate was used as a control for the hole-length EGFR.

4.2 Autophosphorylation of the truncated EGFR

To further evaluate the properties of cathepsin L cleavage and its possible physiological implications, we constructed the N-terminally truncated form of EGFR, which corresponded to the cathepsin cleavage product starting at G225 amino acid residue. The hole-length and truncated EGFR cDNA were cloned into pcDNA4^{TM}myc-HisA vector, a mammalian expression vector with two useful tags, 6xHis and C-Myc, for simpler protein purification or detection using antibodies.
To find a proper expression system, endogenous levels of the EGFR were tested in two cancer cell lines, MDA-MB-231 and HeLa cell lines (
Figure 18A). The MDA-MB-231 cell line showed overexpression of the endogenous EGFR and constitutive activation of the receptor
Figure 18B).

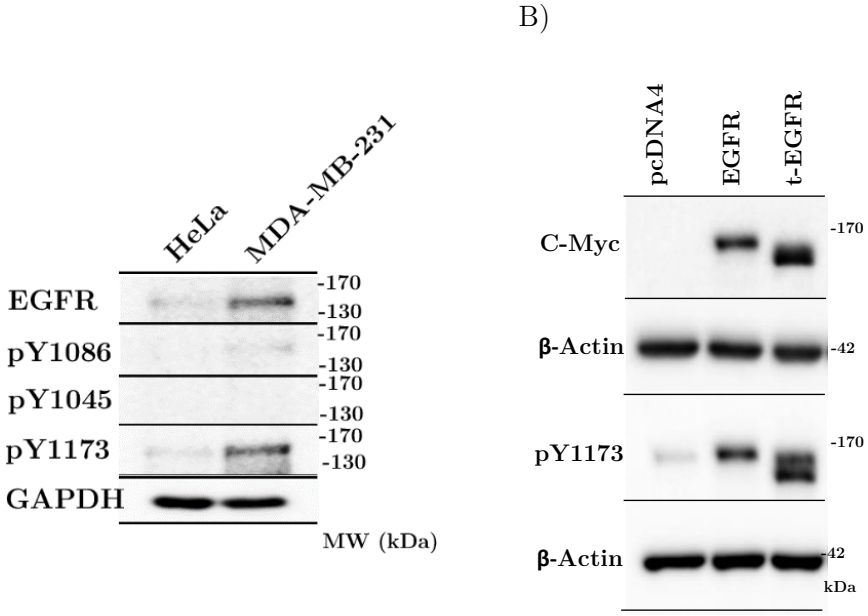


Figure 18: A) Expression level of the endogenous EGFR in HeLa and MDA-MB-231 cell lines. Detection of the EGFR phosphorylated tyrosines Y1086, Y1045 and Y1173 residues was used for phosphorylation level determination, while GAPDH was used as a loading control. B) Immunological detection of the C-Myc and EGFR phosphorylated Y1173 residue in MDA-MB-231 cells expressing empty vector, full-length EGFR or t-EGFR. The detection of β-actin was used as a loading control.

However, the HeLa cell line expressed much lower levels of endogenous EGFR and no basal phosphorylation. Consequently, to determine its autophosphorylation potential, we overexpressed the truncated EGFR (t-EGFR) in a HeLa cell line (Figure 19).

Immunological detection of several known EGFR phosphorylation sites showed that t-EGFR causes constitutive phosphorylation of tyrosine residues Y1173, Y1045, Y1197, Y1086 and Y1068. We also showed that autophosphorylation of truncated EGFR was independent of EGF treatment. As expected, endogenous EGFR from HeLa cells expressing empty vector showed no basal level of phosphorylation, while overexpression of the full-

length EGFR caused some level of constitutive phosphorylation. Phosphorylation of five different phosphorylated tyrosine residues of endogenous and overexpressed EGFR increased in an EGF concentration-dependent manner.

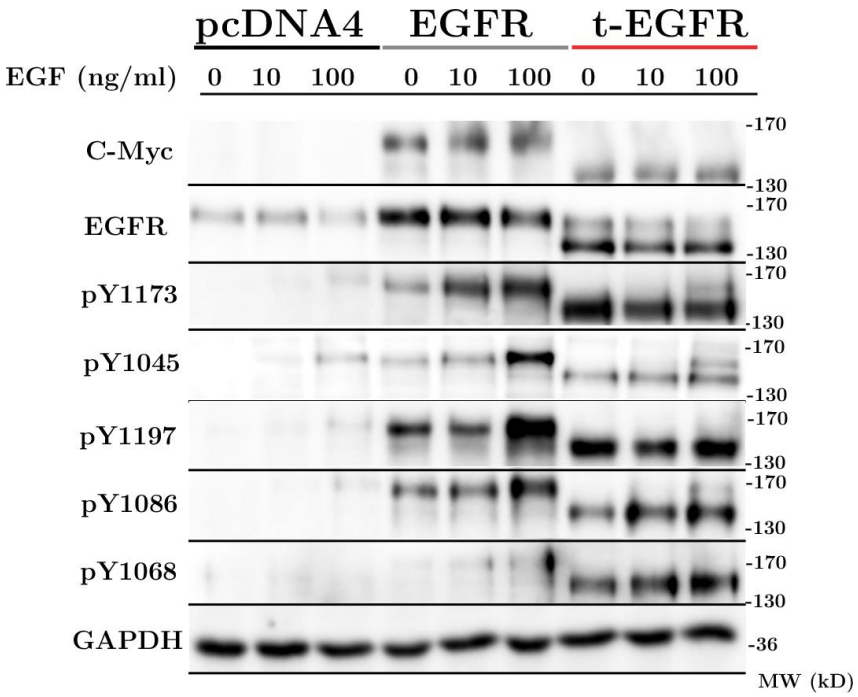


Figure 19: Immunological detection of C-Myc, EGFR, and several EGFR phosphorylated Tyr residues in HeLa cells expressing empty vector, full-length EGFR or t-EGFR, treated with EGF (0, 10 or 100 ng/ml) for 8 min. The detection of C-Myc was used as a transfection control to ensure the same levels of transfected proteins, band represents fusion protein C-Myc with EGFR/t-EGFR. The detection of GAPDH was used as a loading control.

Based on quantitative analysis of Western blots, at the basal level, truncated EGFR showed at least two times higher phosphorylation, as for Y1086, compared to the overexpressed full-length protein (Figure 20).

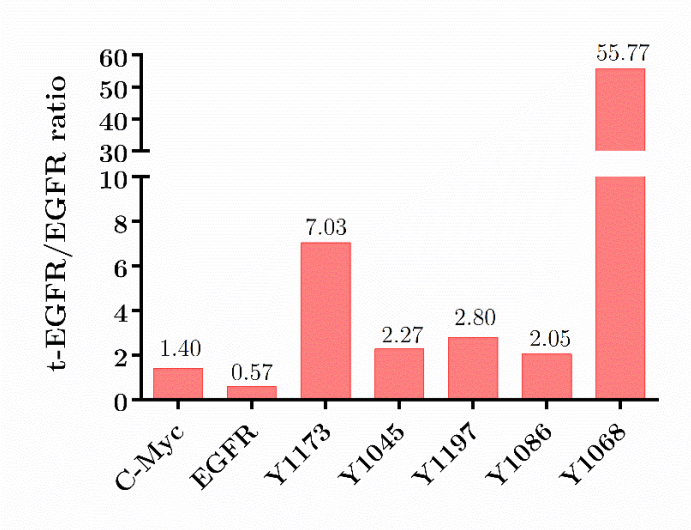


Figure 20: Quantitative analysis of Western blots, bars represent the ratio between t-EGFR and full-length EGFR at the basal level, without EGF treatment.

4.3 Phosphorylation Profile of HeLa Cells Overexpressing Truncated EGFR

We used the Human Phospho-Kinase array to determine possible differences between downstream phosphorylation events in cells expressing full-length and truncated EGFR. Array enabled the identification of 37 kinase phosphorylation sites (Figure 21).

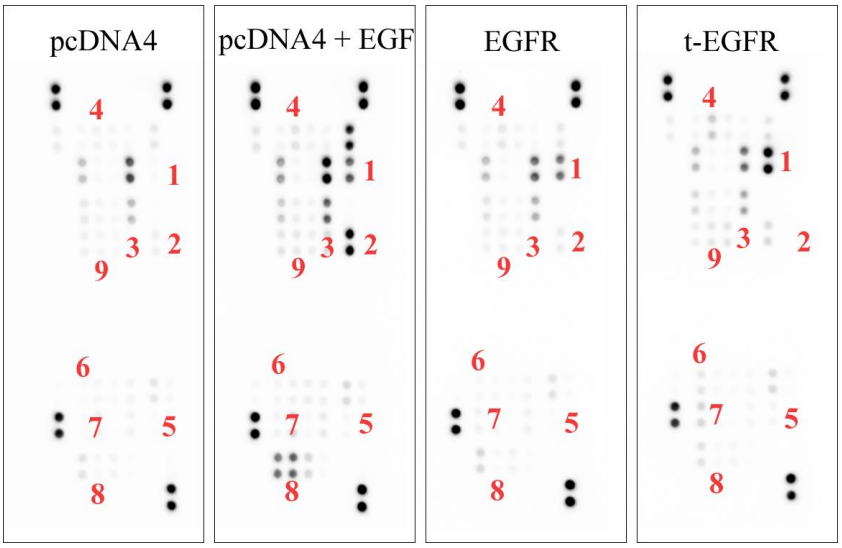


Figure 21: Human Phospho-Kinase Array antibody blots, showing changes in the phosphorylation profile of the Hela cells expressing empty vector as a negative control,

empty vector with EGF stimulation, cells expressing full-length EGFR, and expressing t-EGFR. Numbered phosphorylation sites had significantly increased phosphorylation in cells expressing truncated EGFR compared to the negative control.

After quantification of the dot blot densities, obvious differences in activated signalling pathways were identified. Cells expressing truncated EGFR showed around two-fold higher phosphorylation of the Y1086 than those with overexpressed full-length EGFR (Figure 22).

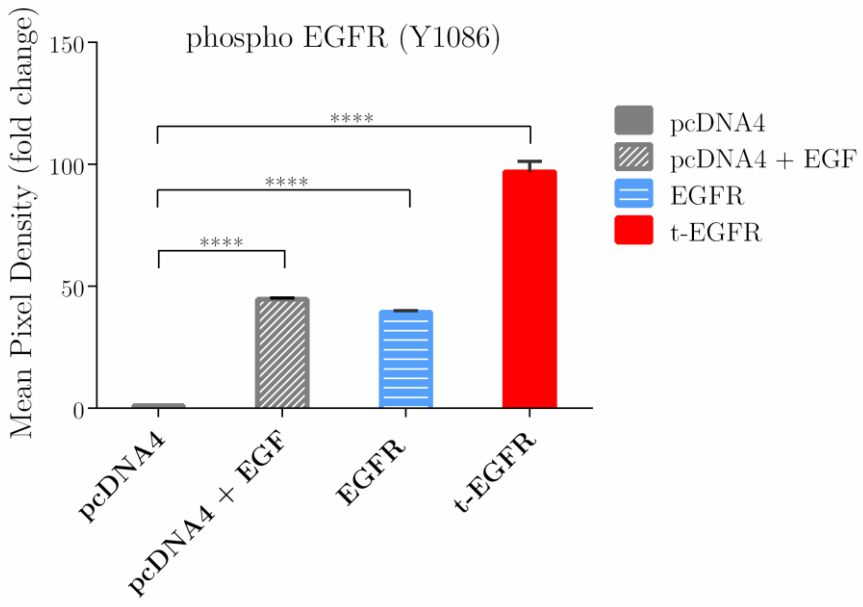


Figure 22: Quantified phosphorylation of the EGFR Y1086 in all 4 dot blots, normalized to the negative control. ****$p < 0.0001$, compared to negative control. Error bars show the standard deviations based on duplicate values of each dataset.

Besides EGFR Y1086, constitutive activation of the truncated EGFR, compared to the negative control, significantly increased the phosphorylation of 10 protein phosphosites from the array (Figure 23). Expression of truncated EGFR significantly increased phosphorylation of ERK1/2, MSK1/2, p38 α, c-Jun, STAT1, STAT3(S705), RSK1/2/3, Src and p70 S6 kinase.

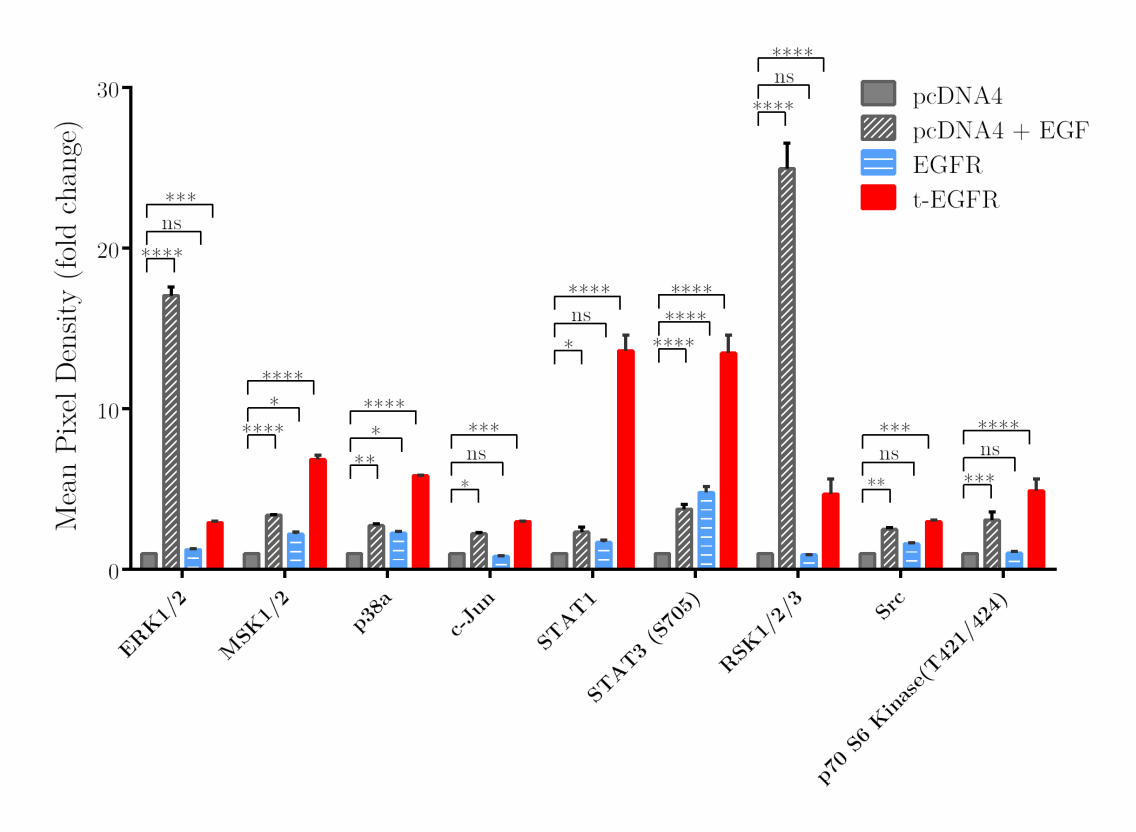


Figure 23: Quantified phosphorylation of the phospho-sites with significantly increased phosphorylation (in t-EGFR expressing HeLa cells compared to negative control). All sites were normalized to the negative control. Dot blots were quantified using ImageJ's extension for the Dot Blot Analyzer. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ****$p < 0.0001$, compared to negative control. Error bars show the standard deviations based on duplicate values of each dataset.

Together with EGFR, the interaction of significant phosphosites was analyzed using String (Figure 24). To determine pathways triggered by this constitutive activation, using the FDR value of less than 0.05, 256 significant pathways were enriched using Gene Ontology (GO), 108 using KEGG, and 82 using the Reactome database.

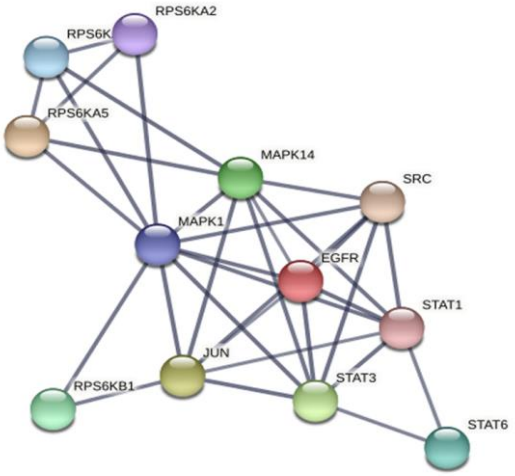


Figure 24: Protein-protein interaction network of the phospho-sites with significantly increased phosphorylation (in t-EGFR expressing HeLa cells compared to negative control), retrieved by String analysis.

Signalling pathways which were statistically more relevant were shown in (Figure 25) for GO, (Figure 26) for KEGG and (Figure 27) for Reactome enrichment. Among enriched pathways were EGFR signalling and signalling pathways such as MAPK family signalling pathways, JAK-STAT and PI3K-Akt signalling pathways, but also pathways involved in resistance to tyrosine kinase inhibitors and regulation of cell death and apoptosis. These results revealed new leads for further research.

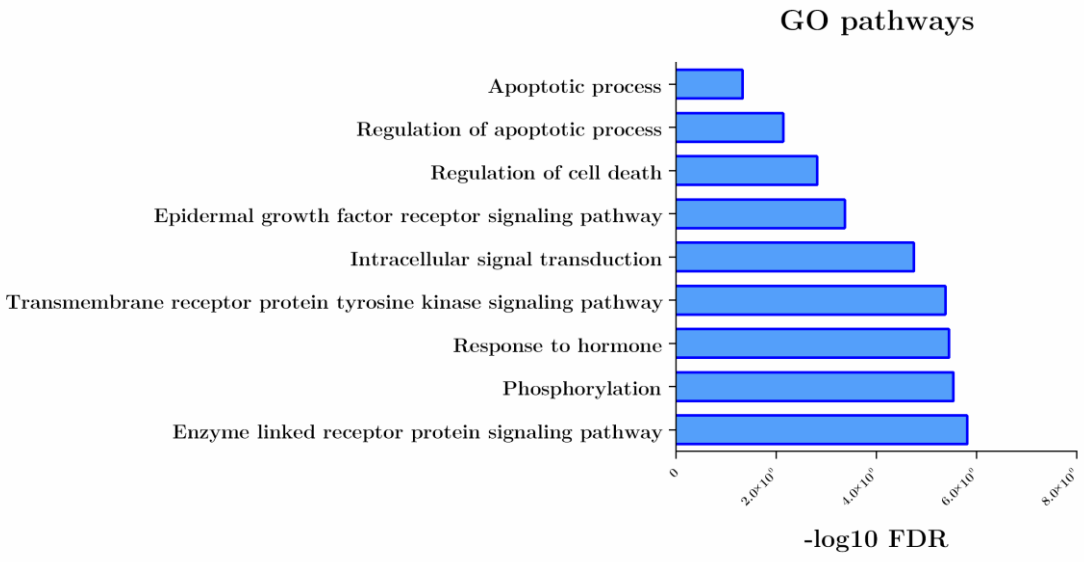


Figure 25: The –log10FDR values for the enriched Gene Ontology (GO) pathways.

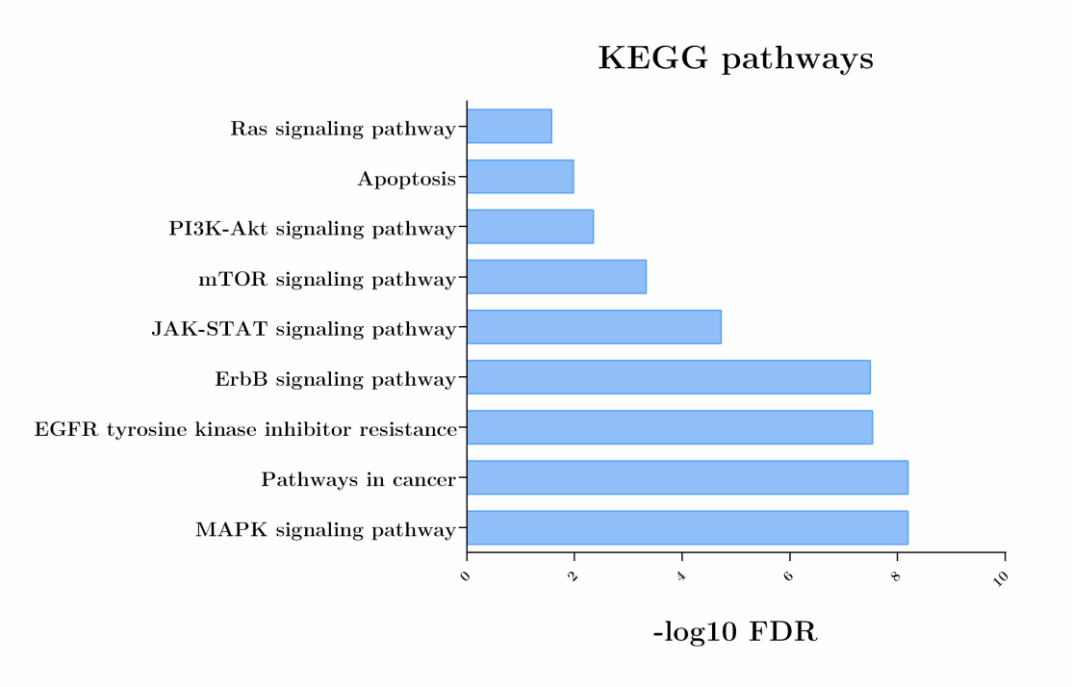


Figure 26: The –log10FDR values for the enriched KEGG pathways.

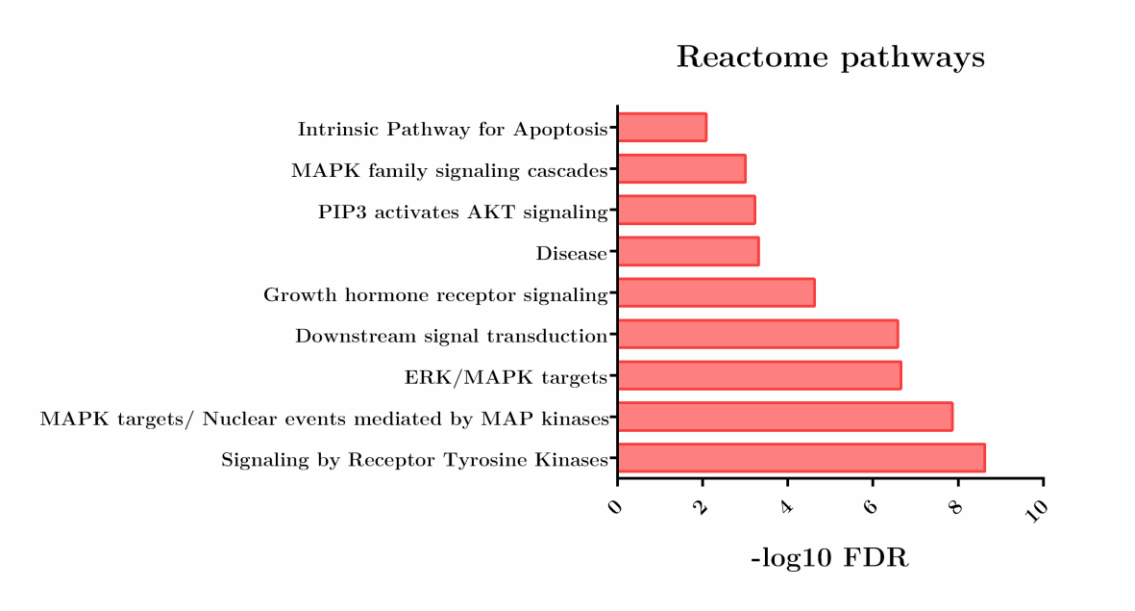


Figure 27: The –log10FDR values for the enriched Reactome pathways.

4.4 Phosphoproteomic Analysis

Active EGFR is involved in the further activation and regulation of various cellular signaling pathways, mainly through direct or indirect phosphorylation. Consequently, determining the cell phosphoproteome could help understand how the expression of truncated EGFR in HeLa cells influences intracellular signaling. We used mass spectrometry-based phosphoproteomics, which enabled us to identify a large number of phosphorylated peptides simultaneously. To increase the abundance of phosphorylated

peptides, we combined two methods for the phosphopeptide enrichment, using PHOS-Select iron-affinity gel and TiO2 beads. HeLa cells expressing pcDNA4, pcDNA4 EGFR and pcDNA4 t-EGFR were collected and prepared using this protocol in three biological samples. This protocol enabled us to identify 2786 phosphorylated peptides present in at least two biological samples in at least one group. The t-test was used to determine significantly differentiated phosphopeptides between two groups, or peptides with at least a two-fold difference and p-value less than 1%. Volcano plots were used for the visualization of the results.

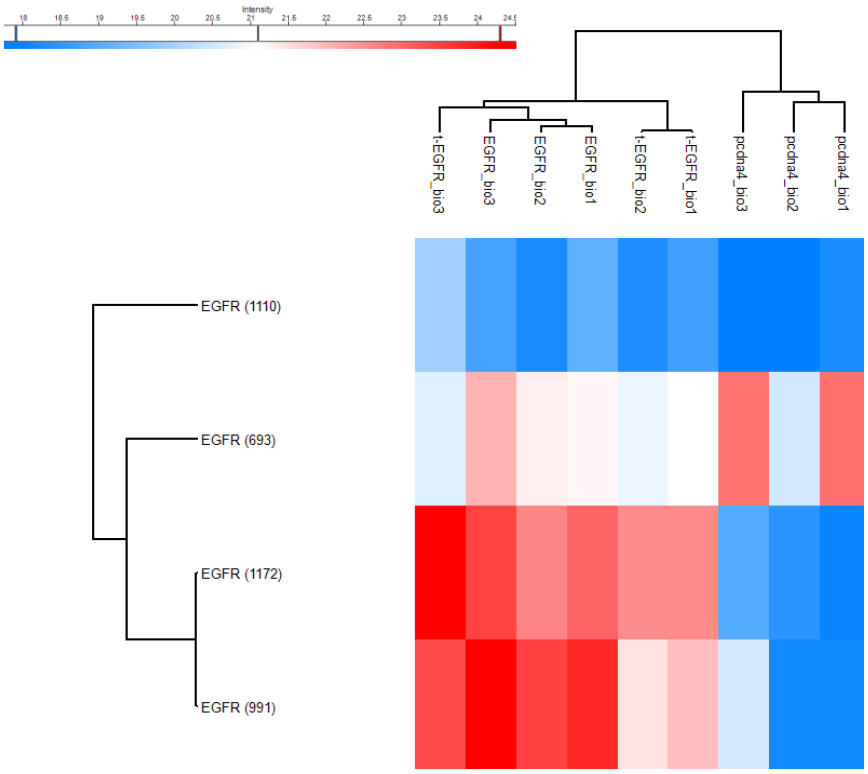


Figure 28: Heat map representing the values for four different EGFR phosphorylated peptides identified in three sample groups, Y1100, T693, S991 and Y1172. The groups represent HeLa cells expressing pcDNA4, pcDNA4 EGFR and pcDNA4 t-EGFR in three biological samples.

Among all identified phosphorylated peptides, four different phosphorylated EGFR peptides were identified, Y1110, T693, S991 and Y1172 (Figure 28). The figure represents values for the EGFR peptides in all samples. Missing values were replaced with the values from the normal distribution.

Figure 29: Visualization of the phosphoproteomics result using the vulcano plot. The volcano plot represents significantly differentiated phospho-peptides between the pcDNA4 and pcDNA4 t-EGFR groups. The x-axis represents the log2 fold chain, and the y-axis represents the –log10 of the p-value. The chosen threshold was at a two-fold change, and the p-value was less than 1%.

When we compared Hela cells expressing pcDNA4 and pcDNA4 t-EGFR, 77 significantly differentiated phosphopeptides were identified. Among them, 17 were significantly up-regulated, and 60 were significantly down-regulated in the t-EGFR group (Figure 29). Among the significantly up-regulated was EGFR phosphorylated at the Y1172, confirming constitutive activation of the truncated EGFR. To better understand processes influenced by t-EGFR expression, the significantly up-regulated phosphopeptides were analyzed using String. The String was used to create a network that can elucidate triggered signalling pathways and cellular compartments involved in these pathways.

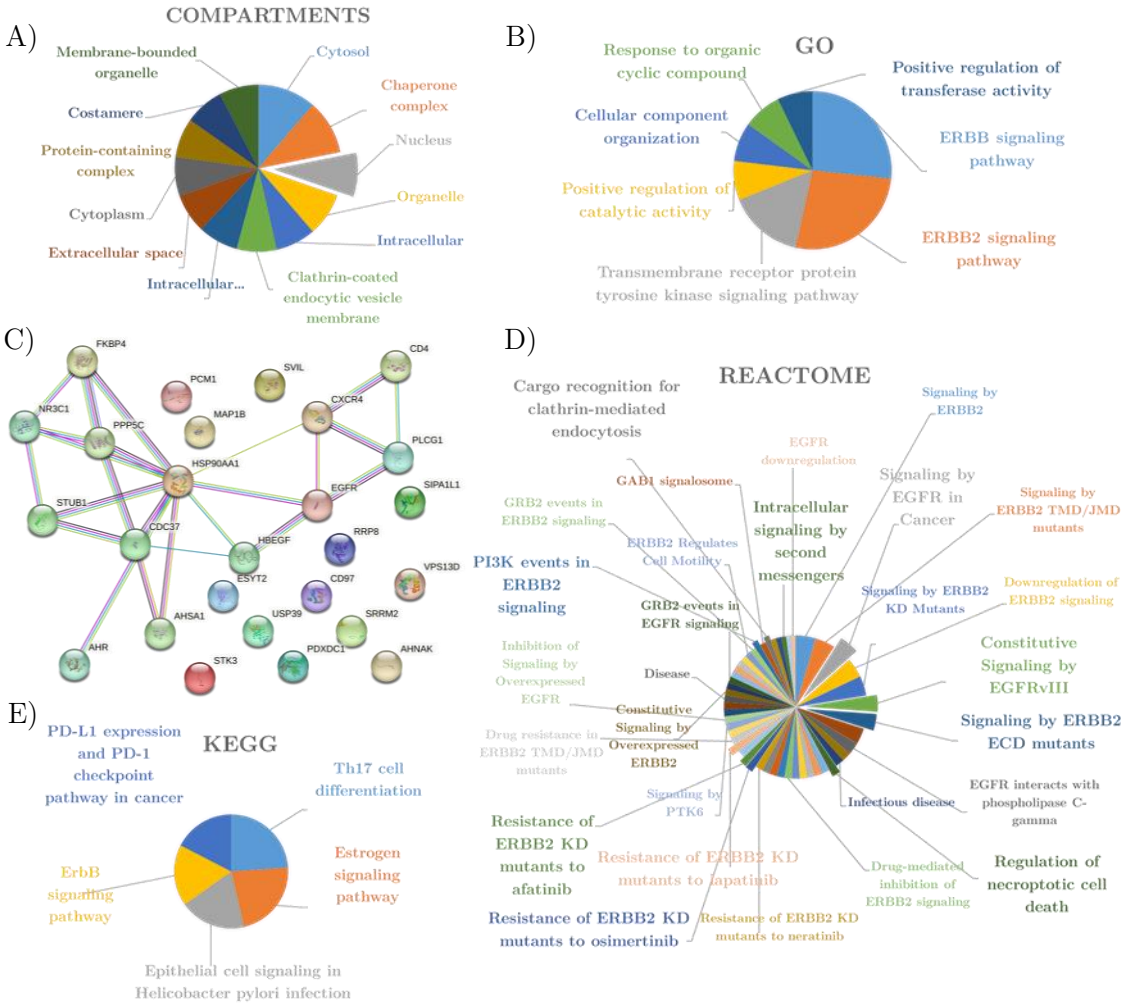


Figure 30: The String analysis of the phosphopeptides significantly up-regulated in HeLa cells expressing t-EGFR when compared to the control group. A) The pie chart of the –log10FDR values for the enriched cellular compartments; B) The pie chart of the –log10FDR values for the enriched Gene Ontology (GO) pathways; C) String network with the confidence of 0.9 and maximum 10 interactions; D) The pie chart of the –log10FDR values for the enriched Reactome pathways; E) The pie chart of the –log10FDR values for the enriched KEGG pathways.

Signalling pathways triggered by t-EGFR expression in the HeLa cell line, which were statistically relevant, were shown in (Figure 30). Among enriched pathways was preferentially ERBB signalling in cancer and the resistance of these receptors to various tyrosine kinase inhibitors. Among significant pathways were the regulation of necroptotic cell death, cellular component organization, and clathrin-mediated endocytosis. Also, cellular compartments involved were, as expected, membrane and cytosol, but also clathrin-coated endocytic vesicle membrane and nucleus.

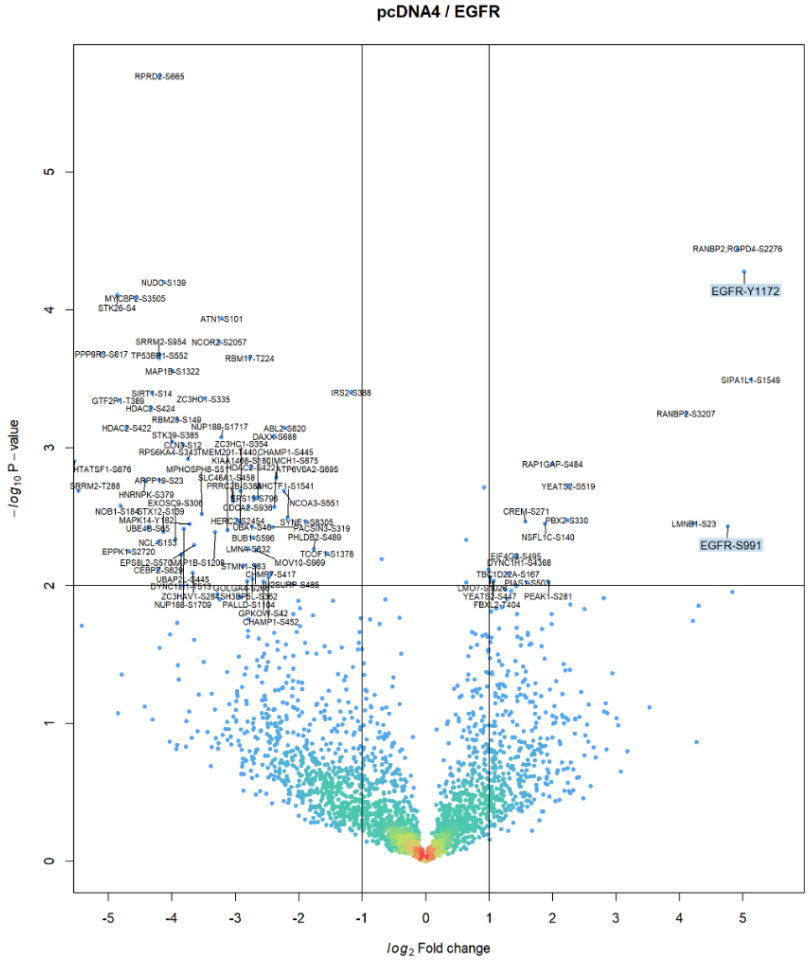


Figure 31: Visualization of the phosphoproteomics result using the vulcano plot. The volcano plot represents significantly differentiated phospho-peptides between the pcDNA4 and pcDNA4 EGFR groups. The x-axis represents the log2 fold chain, and the y-axis represents the –log10 of the p-value. The chosen threshold was at a two-fold change and p-value less than 1%.

When we compared Hela cells expressing pcDNA4 and pcDNA4 EGFR, 93 significantly differentiated phosphopeptides were identified. Among them, 20 were significantly up-regulated, and 73 were significantly down-regulated in the EGFR group. In addition, two EGFR phosphorylated peptides were significantly up-regulated, phosphorylated at S991 and Y1172, again confirming constitutive activation of the overexpressed full-length EGFR.

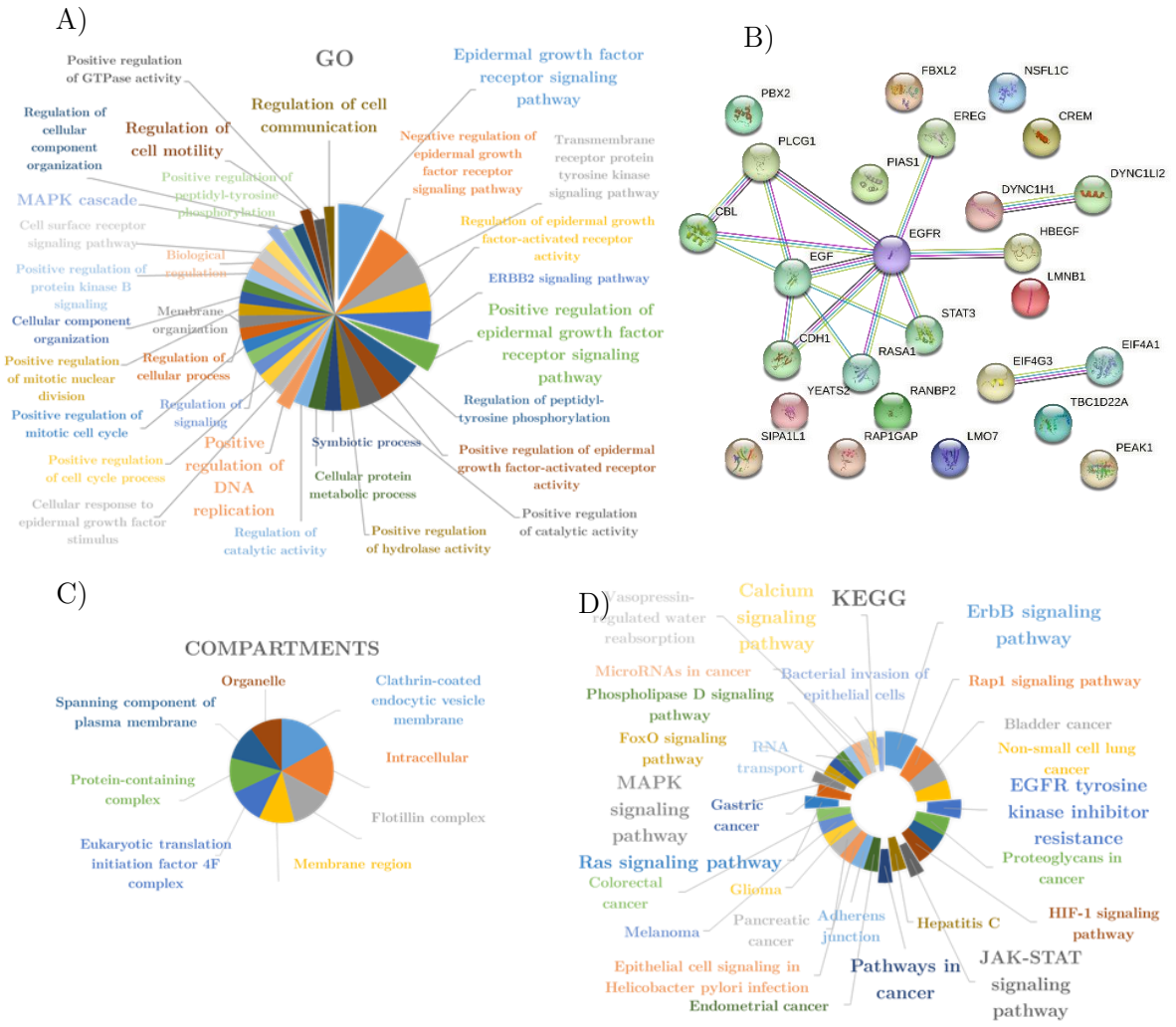


Figure 32: The String analysis of the phospho peptides significantly up-regulated in HeLa cells expressing EGFR when compared to the control group. A) The pie chart of the –log10FDR values for the enriched GeneOntology (GO) pathways; B) String network with the confidence of 0.9 and maximum 10 interactions; C) The pie chart of the –log10FDR values for the enriched cellular compartments; D) The pie chart of the –log10FDR values for the enriched KEGG pathways.

Again, the String was used to retrieve the network of the significantly upregulated phosphopeptides. Signalling pathways triggered by EGFR expression in the HeLa cell line, which were statistically relevant, were shown in (Figure 32). Enriched activated pathways were, as expected, ErbB signalling pathways and some downstream signalling pathways, such as MAPK, JAK-STAT, and calcium signalling pathways. Moreover, positive regulation of cell duplication, cell motility, and cell communication were among the enriched pathways. Resistance to the EGFR tyrosine kinase inhibitor was also enriched. The cellular compartments involved were as expected membrane, but also intracellular and clathrin-coated endocytic vesicle membrane.

Together with the Human Phospho Kinase Array, results from the phosphoproteomic analysis gave us great leads for further investigation.

4.5 Autophosphorylation of Truncated EGFR Shows to Be Insensitive to the Cetuximab

Monoclonal antibody Cetuximab binds to the EGFR with high affinity and blocks its interaction with the EGF ligand. Cetuximab binds to EGFR extracellular domain with a 10 times higher affinity than the EGF ligand. We investigated the effect of Cetuximab on phosphorylation of full-length and truncated EGFR. Cetuximab blocked ligand binding to the receptor in the HeLa cells line transfected with empty vector and cells overexpressing full-length EGFR. In correlation with previous results, truncated EGFR remained insensitive to EGF and Cetuximab (Figure 33).

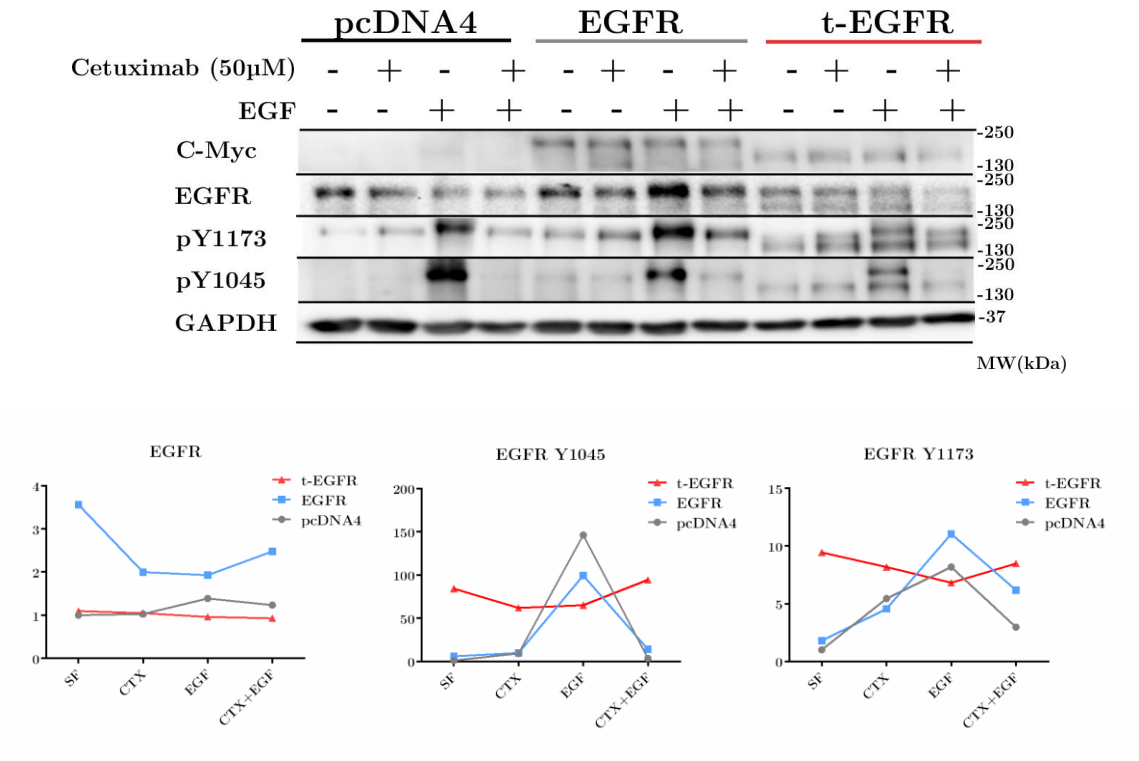


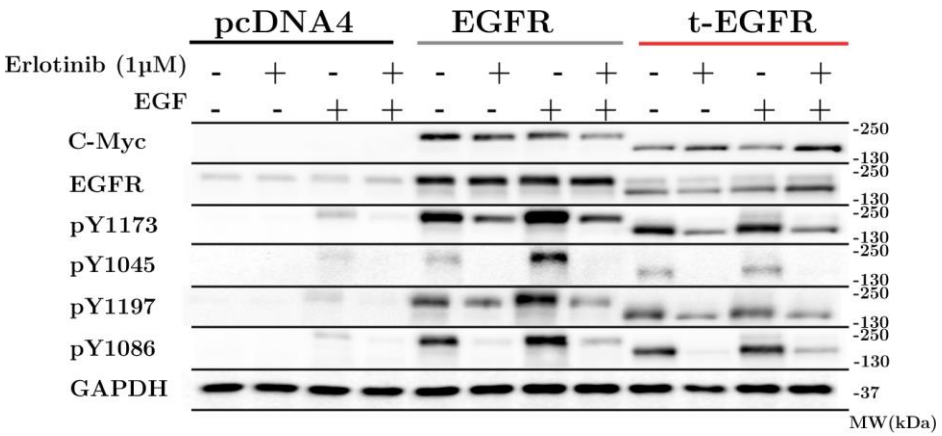
Figure 33: Western blot analysis of EGFR phosphorylation in HeLa cells expressing empty vector pcDNA4TM myc-His (pcDNA4), pcDNA4 EGFR (EGFR), and pcDNA4 t-EGFR (t-EGFR) after cetuximab treatment. Cells were incubated in serum-free or serum-free media with 50 μ g/ml of cetuximab for 48 hours. After incubation, cells were treated with 0 or 100 ng/ml of EGF for 8 min. Levels of EGFR and phosphorylated Y1173 and Y1045 EGFR residues were immunologically detected. The detection of GAPDH was used as a loading control. Bands were quantified, and after the GAPDH and EGFR normalization, band intensities were normalized to the pcDNA4 serum-free sample value. The detection of C-Myc was used as a transfection control to ensure the same level of transfected protein, band represents fusion protein C-Myc with EGFR/t-EGFR.

4.6 EGFR Shows Increased Resistance to the Tyrosine Kinase Inhibitor Erlotinib

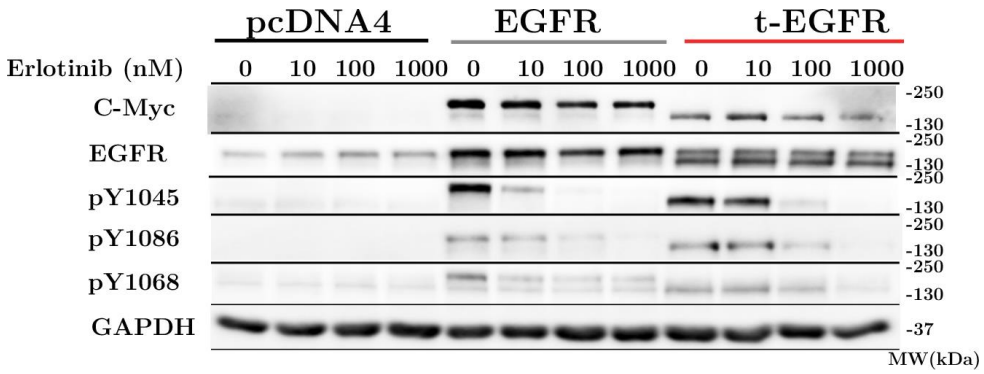
Contrary to cetuximab, erlotinib binds to the cytosolic part of the EGFR and blocks phosphorylation of truncated as well as full-length EGFR. However, truncated EGFR

showed much lower sensitivity to erlotinib since inhibition of the truncated EGFR phosphorylation required a much higher concentration of the inhibitor than inhibition of the full-length EGFR (Figure 34).

A)



B)



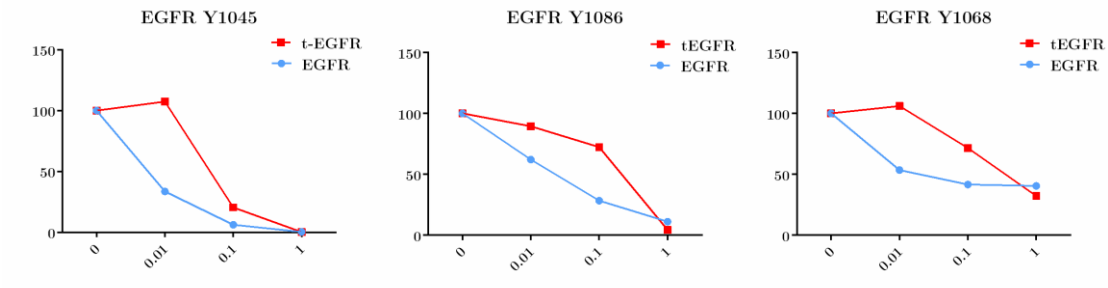
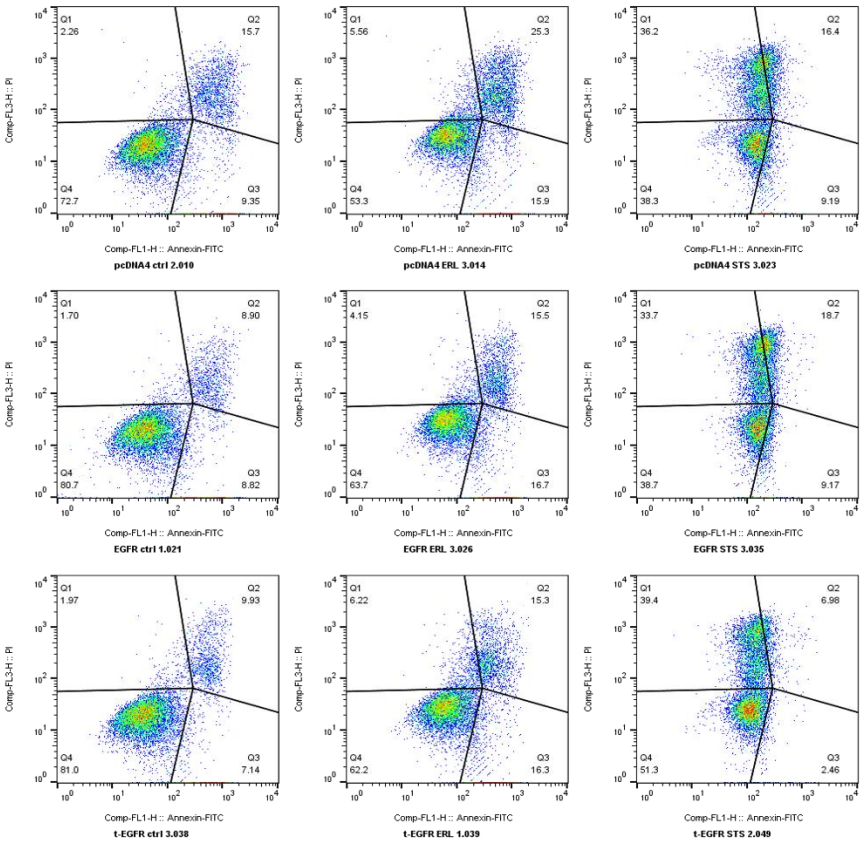


Figure 34: A) Western blot analysis of EGFR phosphorylation in HeLa cells expressing empty vector pcDNA4TM myc-His (pcDNA4), pcDNA4 EGFR (EGFR), and pcDNA4 t-EGFR (t-EGFR) after erlotinib treatment. Cells were incubated in serum-free media overnight. For the last hour, cells were incubated with 0 or 1 μ M erlotinib for 1 h. After incubation, cells were treated with 0 or 100 of EGF for 8 min. Levels of EGFR and phosphorylated Y1173, Y1045, Y1197 and Y1086 EGFR residues were immunologically detected. B) Western blot analysis of concentration dependence to erlotinib. After overnight incubation in serum-free media, HeLa cells expressing empty vector pcDNA4TM

myc-His (pcDNA4), pcDNA4 EGFR (EGFR), and pcDNA4 t-EGFR (t-EGFR) were incubated with 0, 10, 100 and 1000 nM of Erlotinib for 1 hour. Levels of EGFR and phosphorylated Y1045, Y1086 and Y1068 EGFR residues were immunologically detected. Bands were quantified, and after the GAPDH and EGFR normalization, band for EGFR with phosphorylated Y1045, Y1086 and Y1068 residues intensities were normalized to the 0 nM sample value and multiplied by 100. In both figures, the detection of C-Myc was used as a transfection control to ensure the same level of transfected protein, while the detection of GAPDH was used as a loading control.

4.7 Expression of Truncated EGFR Leads to Apoptosis Resistance

Resistance to tyrosine kinase inhibitors (TKIs) and apoptotic processes were among the enriched pathways after the STRING analysis. To validate these results, we tested the influence of the t-EGFR expression on the resistance to erlotinib or apoptosis inducer staurosporine (STS), using flow cytometry. After the treatment with STS, full-length and truncated EGFR expressing cells were stained with Annexin V and PI. The percentage of Annexin V and PI negative, or viable cells, increased in cells overexpressing full-length EGFR and t-EGFR. After the 48 h treatment with erlotinib, viability decreased in a similar trend as untreated cells. Furthermore, overnight treatment with apoptosis inducer staurosporine decreased the viability of the cells. However, the expression of the truncated EGFR caused a significant increase in cell viability and resistance to apoptosis compared to full-length EGFR (Figure 35).



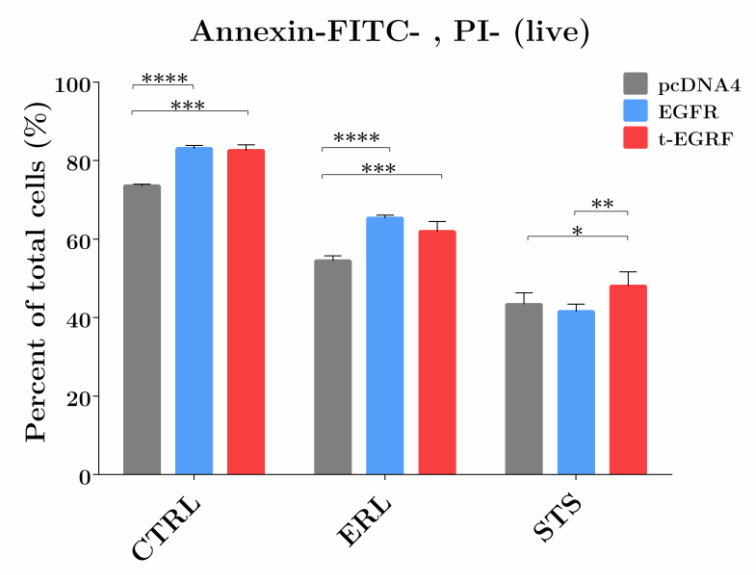


Figure 35: Cell viability analysis of HeLa cells expressing empty vector pcDNA4TM myc-His (pcDNA4), pcDNA4 EGFR (EGFR), and pcDNA4 t-EGFR (t-EGFR), after the treatment with TKI erlotinib (ERL) or apoptosis inducer staurosporine (STS) for 48h. The percentage of Annexin V and PI negative or viable cells were determined using flow cytometry. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; ****$p < 0.0001$, compared to negative control. Error bars show the standard deviations based on triplicate values of each dataset.

4.8 Expression of Truncated EGFR Decreases Cell Proliferation

Higher cell viability of HeLa cells overexpressing full-length and truncated EGFR could be caused by activation of survival signalling but also a consequence of some changes in cell proliferation. To test this, we used the BrdU proliferation assay (Figure 36) because BrdU is a synthetic thymidine analogue and incorporates into the newly synthesized DNA in proliferating cells. HeLa cells expressing empty vector pcDNA4^{TM} myc-His (pcDNA4), pcDNA4 EGFR (EGFR), and pcDNA4 t-EGFR (t-EGFR) were grown in serum-free or completed DMEM media for 40h. The growth of cells in cDMEM medina, when compared to the growth in serum-free media, increased the proliferation of all three cell types. However, Hela cells overexpressing truncated EGFR showed significantly lower proliferation compared to the cells expressing empty vector or full-length EGFR.

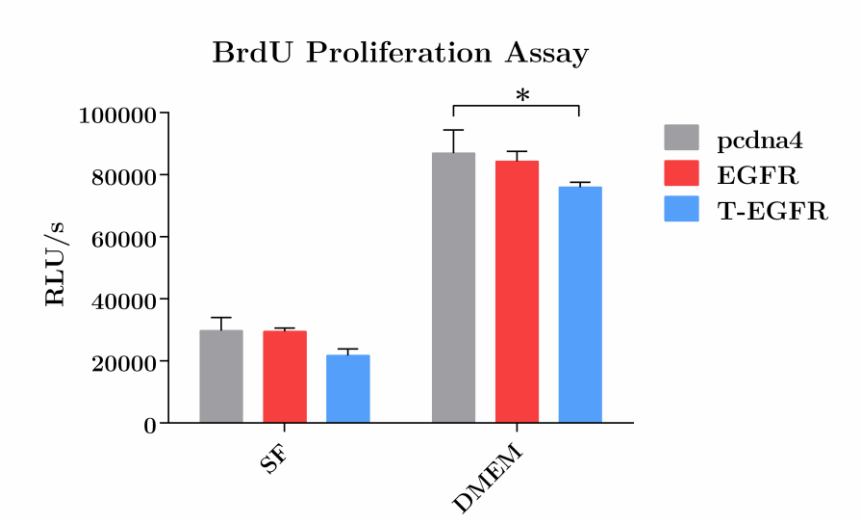


Figure 36: BrdU proliferation assay of HeLa cells expressing empty vector pcDNA4™ myc-His (pcDNA4), pcDNA4 EGFR (EGFR), and pcDNA4 t-EGFR (t-EGFR), grown in serum-free (SF) cDMEM or cDMEM (DMEM) for the 40 hours. ** $p < 0.01$, compared to negative control. Error bars show the standard deviations based on triplicate values of each dataset.

4.9 Nuclear Localization of the Truncated EGFR

Resistance to tyrosine kinase inhibitors (TKIs) and cetuximab was previously connected with nuclear localization and phosphorylation of EGFR and STAT3. Also, the phosphoproteomics results connected the expression of the t-EGFR with the nucleus as an enriched cellular compartment after the String analysis. To test the nuclear and cytosolic localization of these two proteins, transfected HeLa cells were fractionated using REAP protocol ([209]). EGFR endogenously present in HeLa cells was present in the cytosolic fraction but was not detected in the nuclear fraction (Figure 37). However, overexpressed full-length and truncated EGFR were present in the cytosolic and nuclear fractions. Even though the overall expression level of full-length EGFR and t-EGFR were similar in the whole cell lysate (w.c.l.), the relative nuclear localization of the t-EGFR was higher than full-length EGFR. Both full-length EGFR and t-EGFR showed constitutive phosphorylation in the cytosolic and nuclear fraction, whereas again, t-EGFR showed almost two times higher phosphorylation compared to full-length EGFR. However, compared to the whole cell lysate, only a small part of the phosphorylated cellular t-EGFR had nuclear localization.

Furthermore, STAT3 was present in all fractions, but a small part of the cellular level of the protein had nuclear localization. However, phosphorylation of the STAT3 at Y705 showed that almost all of the phosphorylated Y705 had nuclear localization. HeLa cells with overexpression of full-length EGFR and t-EGFR showed increased phosphorylation of the STAT3 at Y705. Moreover, Y705 STAT3 had much higher phosphorylation in cells expressing truncated EGFR compared to full-length EGFR. A small part of the cellular STAT3 phosphorylated at S727 was present in the nuclear fraction in HeLa cells expressing full-length EGFR and t-EGFR. At the same time, negative control showed no nuclear localization of this phosphorylation site. As expected, loading control GAPDH was detected in w.c.l. and cytosolic fraction, while Histone H3 was detected in w.c.l. and nuclear fraction.

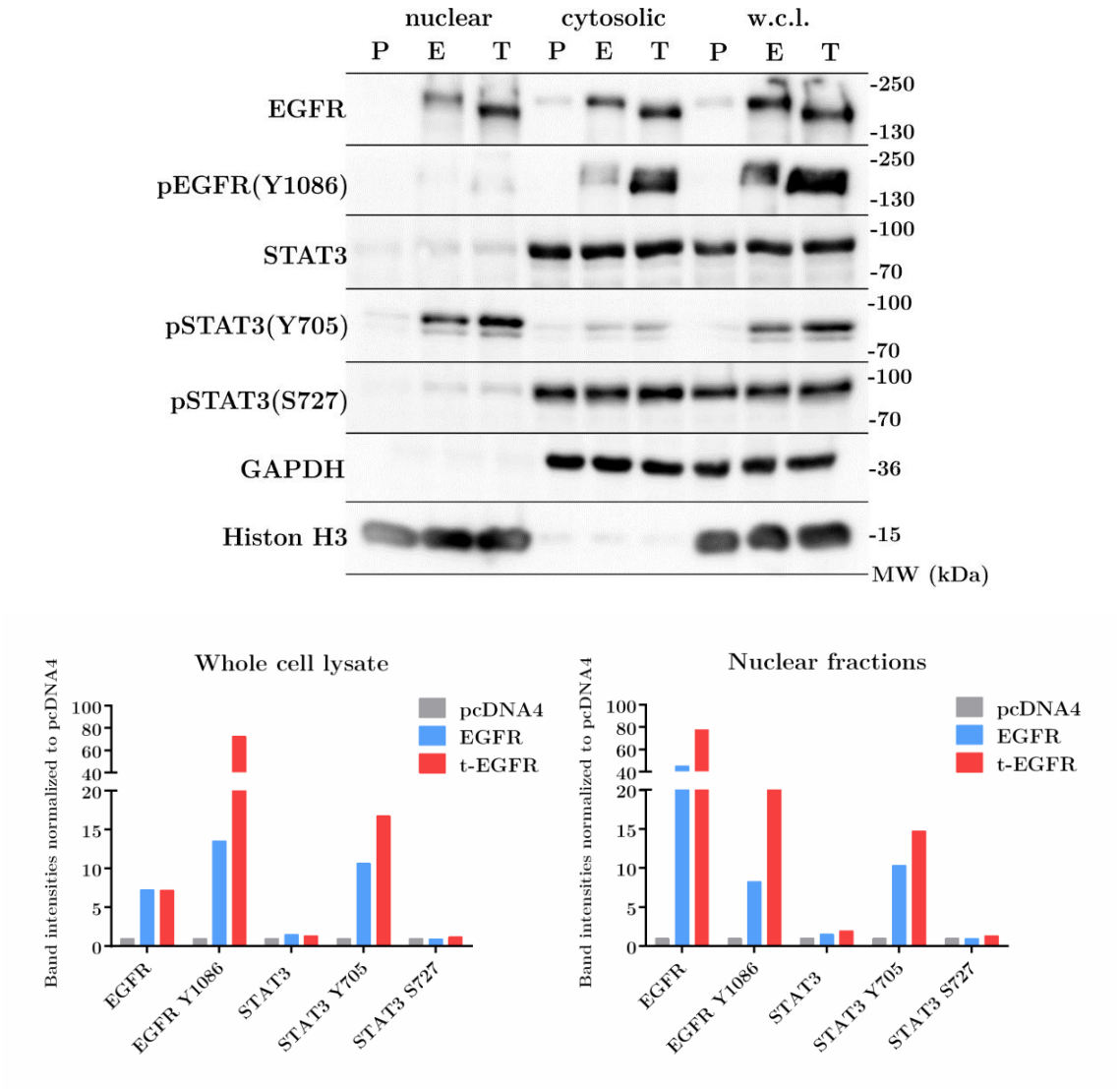
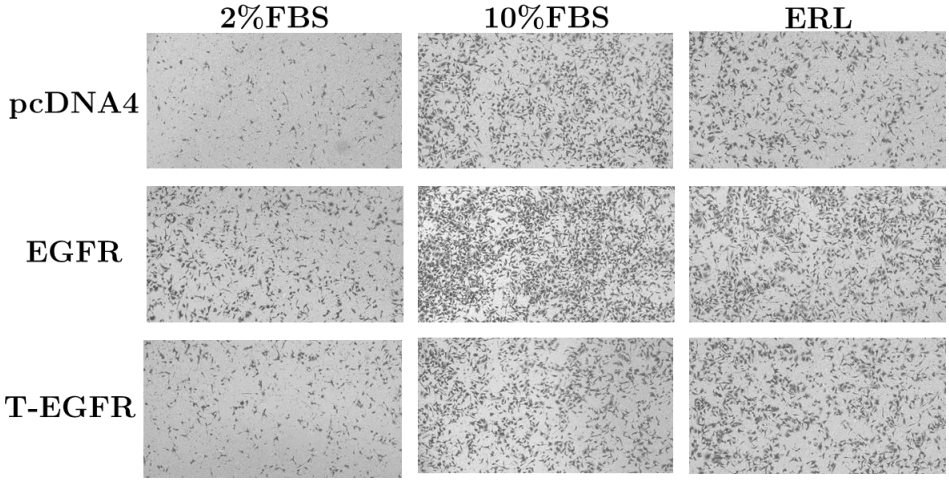


Figure 37: Western blot analysis of the HeLa cells expressing empty vector pcDNA4™ myc-His (P), pcDNA4 EGFR (E), and pcDNA4 t-EGFR (T) after fractionation according to the REAP protocol. Whole cell lysate (w.c.l.), cytosolic and nuclear fractions were analyzed for EGFR and STAT3 localization and phosphorylation. Levels of EGFR, EGFR phosphorylated Y1086 residue, STAT3, and STAT3 phosphorylated Y705, and S727 residues were immunologically detected. Antibodies against GAPDH and Histon H3 were used as loading controls. Bars represent the values of the band intensities normalized to the pcDNA4 value.

4.10 Migration

Cell migration is another EGFR-mediated cellular process. The regulation of cell motility was among the enriched pathways after the String analysis. To test whether overexpression of the full-length or truncated EGFR influences HeLa cells migration, we performed a migration assay using a 24-well insert transwell with an 8 µm porose membrane (Figure 13). We observed that more cells migrated towards the 10%FBS than the 2%FBS (Figure 38). Moreover, when compared with cells expressing empty vectors, cells expressing full-length EGFR showed a significantly higher number of migrated cells, but the expression

of truncated EGFR did not cause significant changes in cell migration. However, erlotinib treatment significantly decreased migration in cells expressing empty vector and full-length EGFR, but in cells expressing truncated EGFR, the decrease was not significant.



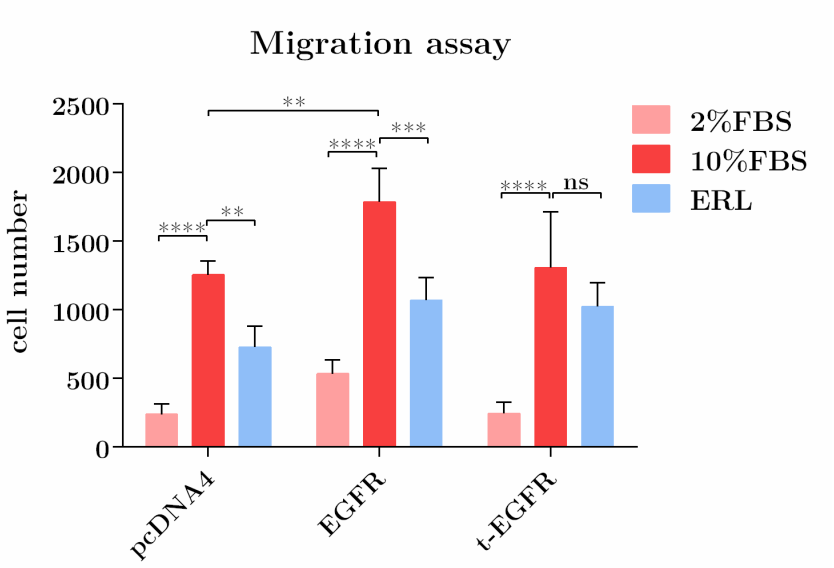


Figure 38: Migration of the HeLa cells expressing empty vector pcDNA4TM myc-His (pcDNA4), pcDNA4 EGFR (EGFR), and pcDNA4 t-EGFR (t-EGFR), through the transwell for 24 hours. Cell migrated from the transwell containing serum-free cDMEM, through the 8 μm porose membrane of the transwell, into the well containing 2% or 10% FBS cDMEM culture media. In the erlotinib-treated sample, cells migrated from the serum-free cDMEM culture media containing 0.01 μM erlotinib into the 10% FBS cDMEM culture media. The cells were fixed, stained using Crystal Violet, and observed under the microscope. The number of cells was counted in three different fields of view to acquire an average sum of the migrated cells. ** $p < 0.01$; *** $p < 0.001$; ****$p < 0.0001$. Error bars show the standard deviations based on triplicate values of each dataset.

Chapter 5

Discussion

Alteration in EGFR signalling was strongly connected with tumor promotion [105], [117]. EGFR contributes to cancer progression through receptor overexpression, cross-talk with other receptors, or through the signalling of several EGFR mutants. Also, EGFR has been intensively studied as a therapeutic target. Several Food and Drug Administration (FDA)-approved EGFR inhibitors have shown promising results in the treatment of several cancers, such as non-small cell lung cancer (NSCLC), breast cancer, and colorectal cancer [115]. However, patients who initially respond well to EGFR inhibitors almost inevitably develop resistance. Consequently, a better understanding of the mechanism underlying resistance has become essential.

EGFRvIII is an EGFR mutant present in various cancers, missing extracellular domain I and two-thirds of domain II [149], [160], [161], [163], [214]. EGFRvIII is constitutively active, and this activity has been explained by its asymmetric active conformation, facilitated by the absence of the N-terminal extracellular region.

EGFR has also been recognized as a substrate of several proteases. As in the case of genetic deletion, complete or partial removal of the EGFR ectodomain by proteolytic cleavage can also influence the activity and functionality of the receptor [164]. Cysteine cathepsins, proteases upregulated in various cancers, are secreted into the tumor microenvironment, where they cleave various membrane-bound proteins, such as EGFR [4]–[6], [97]. Cysteine cathepsins L and S, were found to cleave the extracellular domain of EGFR [97], suggesting their role as modifiers of EGFR activity in cancer.

In this thesis, we have further evaluated the functional relevance of EGFR cleavage by cathepsin L. Using the Fast Profiling of Protease Specificity (FPPS) method [215], we have determined the cathepsin L cleavage site on EGFR. Our results showed that cathepsin L cleaves EGFR at the extracellular domain II at the R224-G225 bond, creating truncated EGFR. Also, cathepsin L-mediated cleavage of EGFR leads to constitutive activation, and this phosphorylation was independent of ligand stimulation. Therefore, similar to EGFRvIII, we can suggest that cleavage of EGFR by extracellular cathepsin L causes a conformational change into the asymmetric active conformation, responsible for receptor constitutive activity.

Using physical simulation models, transmembrane protein cluster formation was predicted due to hydrophobic mismatch and liquid forces which transmembrane proteins create immersed in lipid bi-layers [216]. In the case of EGFRvIII signalling, dimerization-independent signaling was already described [217]. However, in a more recent study, high-precision localization microscopy demonstrated that EGFRvIII is organized in clusters, while signal-to-signal distance indicated EGFRvIII dimer formation [218]. Moreover, disruption of the kinase dimer interface using point mutations, also indicated the

importance of asymmetric kinase dimer formation in EGFRvIII signalling [219], all suggesting that the cathepsin L-truncated EGFR will also form a dimer

During ligand binding, the EGFR extracellular domains I, II, and III form a C shape, where EGF is located between domains I and III [220], [221]. In cathepsin L-truncated EGFR, the EGF binding region is cleaved off. According to our results, cathepsin L-mediated cleavage of EGFR results in decreased EGF binding and increased EGFR autophosphorylation. Our finding is supported by a previous study in which cathepsin L-knockout mouse keratinocytes were more responsive to EGF than wild-type keratinocytes [222]. The authors connected this enhanced responsiveness with increased recycling of the receptor and ligand. However, decreased ectodomain cleavage of the receptor is another potential explanation for improved responsiveness to EGF in cathepsin L-knockout keratinocytes.

We confirmed constitutive phosphorylation of the t-EGFR at five different phosphorylation sites. Similarly, Huang et al. analysed EGFRvIII phosphorylation in glioblastomas (GBM) and confirmed constitutive activation of the eight EGFRvIII phosphorylation sites, including Y1068, Y1086, and Y1173 [223]. A difference in t-EGFR and EGFRvIII phosphorylation is at the Y1045 phosphorylation site. The truncated EGFR version shows high constitutive phosphorylation at this site, while the EGFRvIII was shown to have impaired phosphorylation at this site. This decrease in EGFRvIII phosphorylation at Y1045 was connected with its impaired degradation, prolonged activity on the plasma membrane and higher oncogenicity [162]. On the other hand, Stec et al. recently suggested that phosphorylation of the EGFRvIII at Y1045 was insufficient to provoke degradation [224]. The reasons for the EGFRvIII's high stability on the plasma membrane remained controversial.

Furthermore, our results showed that overexpression of the full-length EGFR caused receptor autophosphorylation without ligand stimulation. This phosphorylation of the overexpressed full-length EGFR is ligand-dependent and increases with the EGF ligand concentration. Constitutive signalling of overexpressed EGFR has been described as noncanonical since this signalling does not lead to the activation of several canonical signalling pathways, such as the ERK and Akt signalling pathways, which is in correlation with our results [159].

Using two different methods, we also showed that activation of the truncated EGFR in HeLa cells causes alterations in the MAPK family, JAK-STAT, and PI3K-Akt signalling pathways, as well as in pathways that are involved in the regulation of cell death, cell growth, cell mobility and resistance to tyrosine kinase inhibitors. If we compare our t-EGFR with the EGFRvIII variant, constitutive activation of the EGFRvIII is found to similarly promote cell proliferation, invasion, and angiogenesis, and reduction in apoptosis [139], [147], [150]. According to these results, we can suggest that cathepsin L truncation of the EGFR has tumorigenic functions.

After the analysis, resistance to tyrosine kinase inhibitors was found among the signalling pathways significantly altered by constitutive activation of truncated EGFR in HeLa cells. Of the two inhibitors used, namely, erlotinib and cetuximab, the binding of cetuximab didn't change the autophosphorylation of the truncated EGFR. The EGFR binding place of the cetuximab is at the domain III of the extracellular domain, and consequently binding with truncated EGFR is possible [225]. Similary, cetuximab can bind the EGFRvIII and cause its internalization, but it fails to inhibit EGFRvIII activity [226]. However, in this case, EGFRvIII was translocated to the mitochondria, which is in contrast with full-length EGFR and truncated EGFR. Our results also suggest that cetuximab blocks the ligand-induced signalling of overexpressed full-length EGFR, while autophosphorylation of overexpressed full-length EGFR remains the same, consistent with the findings of Chakraborty et al. [159].

On the other hand, erlotinib blocks the autophosphorylation of both truncated and overexpressed full-length EGFR. Erlotinib blocks the activation of tyrosine kinases through binding to the ATP binding site at the tyrosine kinase domain. This led us to suggest that kinase activity is necessary for the autophosphorylation of t-EGFR. However, our results also showed that cleavage with cathepsin L generates truncated EGFR, which is more resistant to erlotinib than overexpressed full-length EGFR. This could be another way of explaining the general resistance of cancer cells to erlotinib. Moreover, such cathepsin-mediated truncation of EGFR could be explored as a biomarker of anti-EGFR therapeutic resistance [227], [228].

There are several studies describing EGFRvIII as a mediator of erlotinib resistance. The shRNA-mediated knockdown of EGFRvIII resensitized erlotinib-resistant cells [229]. EGFRvIII was also suggested as a potential target for immunological targeting in TKI-resistant tumors [230]. Also, in GBM-derived cell line BS153, erlotinib resistance was connected with an increase in EGFRvIII protein and upregulation of PI3K downstream signalling cascade [231]. According to our Human Phospho Kinase Array and phosphoproteomics results, PI3K-Akt signalling is one of the enriched signalling pathways These results could suggest one of the mechanisms of the t-EGFR influence on erlotinib resistance. However, this mechanism of resistance should be further investigated.

In another study, a pool of A-431 cells resistant to erlotinib was created. A-431 represents a cell line with full-length EGFR overexpression. In this study, erlotinib reduced phosphorylation in both resistant and non-resistant pools. However, the difference in cell viability after the treatment was obvious. The resistant poll showed higher viability after the erlotinib treatment [232]. Our results showed that HeLa cells overexpressing, either t-EGFR or full-length EGFR, had increased viability compared to the control. The increase in viability was not a consequence of the increased proliferation. Also, cells overexpressing t-EGFR or full-length EGFR seem more resistant to erlotinib. However, when treated with the apoptosis inducer staurosporine, t-EGFR-expressing cells showed higher viability than control and full-length EGFR-expressing cells. This could be another proof of t-EGFR involvement in cancer cell resistance to erlotinib and apoptosis. There are several published explanations. Impaired responses to apoptosis in glioblastomas were connected with the interaction of full-length EGFR and EGFRvIII with the proapoptotic protein PUMA [233]. These receptors block PUMA translocation to the mitochondria and thus block apoptosis. Also, one of the studies indicates that EGFRvIII regulates apoptosis resistance through BCL-XI up-regulation and suppression of caspase 3 activity [234].

Another possible mediator of resistance to erlotinib/cetuximab and apoptosis could be the interaction of EGFR with STAT3 and their nuclear localization. There are several proposed mechanisms for EGFR nuclear translocation reviewed by [171]. However, the exact mechanism of the EGFR nuclear translocation is still an open question.

One of the mechanisms of nuclear translocations is trough endosomal sorting. During the endosomal sorting, EGFR goes through clathrin-dependent or clathrin-independent endocytosis, further leading to receptor: (i) downregulation; (ii) degradation: (iii) recycling back to the plasma membrane; or (iv) translocation to the nucleus. Since our results showed constitutive activation of Y1086, Y1068 and Y1045 on the truncated EGFR, we suggested that truncated EGFR can be internalized by endocytosis, and possibly translocated to the nucleus. These tyrosine residues are shown to be involved in EGFR endocytosis, and present docking sites for adaptor proteins responsible for clathrin recruitment to the EGFR [235]–[237]. Moreover, analysis of the phosphoproteomics data, when t-EGFR sample was compared with the control, among the enriched cellular compartments also showed clathrin-coated endocytic vesicle membrane and nucleus.

Furthermore, we confirmed that the STAT3 Y705 and STAT1 Y701 phospho-sites had significantly increased phosphorylation in HeLa cells that express t-EGFR. Increasing

evidence shows that STAT3 signalling mediates and promotes resistance to EGFR therapeutics, as reviewed in [190]. Our results also show nuclear localization and phosphorylation of STAT3 and t-EGFR. The high level of nuclear EGFR was associated with resistance to gefitinib and cetuximab in several types of cancer cells [166], [173], [238], [239]. Moreover, published data suggest that the translocation of EGFR to the nucleus occurs independently of its kinase activity and that current anti-EGFR therapeutics cannot properly target EGFR in the nucleus [166], [240]. Although our results showed nuclear localization and phosphorylation of STAT3 and t-EGFR, the exact mechanisms of t-EGFR nuclear translocation and t-EGFR/STAT3 interaction remain unclear.

EGFR signalling has a major role in cancer cell migration, invasion and metastasis. Cells overexpressing EGFR showed elevated and sustained migration in various cancers [241], [242], which correlates with our results. However, erlotinib treatment significantly decreased migration in controls but not in HeLa cells expressing truncated EGFR. This result confirms the resistance of the t-EGFR-expressing cells to the TKI erlotinib. Our findings can be connected with EGFR/STAT3 activation and involvement in TKI resistance [243], [244]. Furthermore, combined inhibition of EGFR and STAT3 was proposed during the treatment to overcome resistance [182].

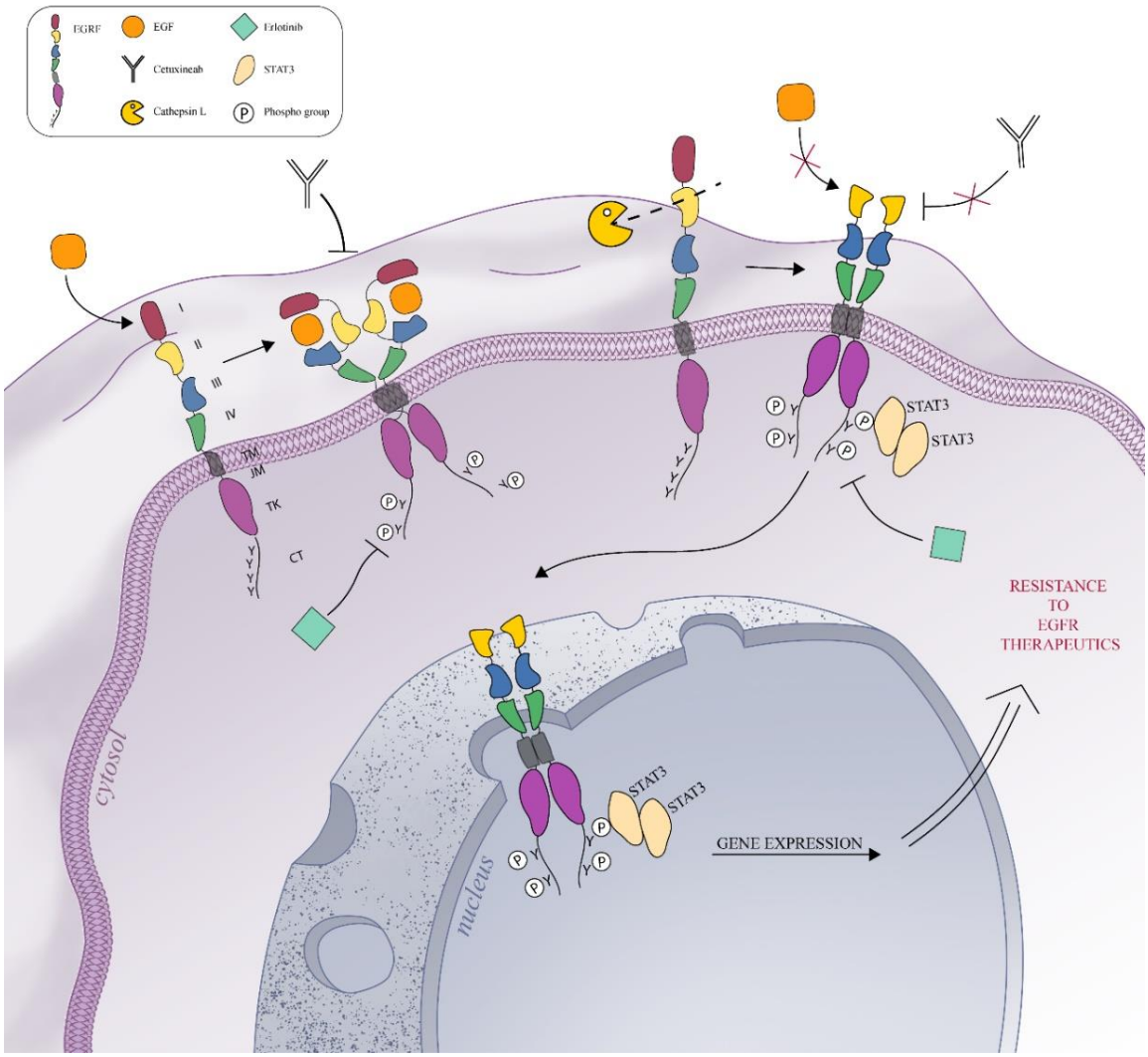


Figure 39: Molecular mechanism of EGFR signalling after the cathepsin L mediated ectodomain cleavage. EGFR signalling is activated by ligand binding to its extracellular

domain. Ligand binding causes receptor conformational change and activation. Activation of EGFR leads to the phosphorylation of tyrosine residues within its cytoplasmic tail and further activation of various intracellular signalling pathways. EGFR signalling can be inhibited by monoclonal antibodies such as cetuximab or tyrosine kinase inhibitors such as erlotinib. After EGFR ectodomain cleavage mediated by cathepsin L, newly generated truncated EGFR shows constitutive activity even in the absence of ligand binding. The lack of domain I inhibits the binding of EGF and cetuximab to t-EGFR. Erlotinib blocks t-EGFR activity, but cells expressing t-EGFR seem more resistant to erlotinib than cells overexpressing full-length EGFR. Activation of t-EGFR leads to STAT3 activation and nuclear translocation. High levels of nuclear t-EGFR and STAT3 lead to the resistance of cancer cells to anti-EGFR therapeutics.

Finally, we can propose a molecular mechanism of EGFR signalling after cathepsin L-mediated ectodomain cleavage (Figure 39). After EGFR ectodomain cleavage mediated by cathepsin L, newly generated truncated EGFR shows constitutive activity even in the absence of ligand binding. The lack of domain I inhibits the binding of EGF as well as cetuximab to t-EGFR. Erlotinib blocks t-EGFR activity, but cells expressing t-EGFR seem more resistant to erlotinib than cells overexpressing full-length EGFR. Activation of t-EGFR possibly leads to STAT3 activation and nuclear translocation. High levels of nuclear EGFR and STAT3 lead to the resistance of cancer cells to anti-EGFR therapeutics.

Chapter 6

Conclusions

This dissertation brings new knowledge about the important role of extracellular cysteine cathepsins in cancer progression. The findings confirmed cathepsin L cleavage of the EGFR ectodomain and revealed the physiological relevance of this cleavage. We used a mass spectrometry-based approach to determine the cathepsin L cleavage site on the EGFR.

Our findings confirmed our hypothesis that cathepsin L creates a stable-proteolytic product, truncated EGFR. Additionally, we demonstrated that truncated EGFR exhibits constitutive activation in the absence of ligand binding. As we stated in our hypothesis this autophosphorylation of the truncated EGFR is enhanced and distinct from the ligand-dependent autophosphorylation.

The constitutive activation of the truncated EGFR activates different intracellular signaling pathways, thereby confirming our hypothesis. Activated signaling pathways are involved in processes like cell growth, cell mobility, cell death and resistance to receptor tyrosine kinases. The phosphorylation profile of the cells expressing truncated EGFR revealed t-EGFR tumorigenic properties and changes in cellular physiology, as proposed in our hypothesis.

Our result also showed that truncated EGFR translocates to the cell nucleus after activation. Nuclear t-EGFR further leads to STAT3 nuclear localization and activity. Furthermore, expression of truncated EGFR leads to impaired response to apoptosis and increases cancer cell viability and resistance to EGFR therapeutics, confirming our hypothesis.

Finally, our findings could lead to a better understanding of EGFR signaling and possibly more effective anticancer treatment strategies in which EGFR inhibitors are combined with cathepsin inhibition.

References

[1] N. D. Rawlings, A. J. Barrett, and A. Bateman, "MEROPS: the peptidase database," *Nucleic Acids Res.*, vol. 38, no. suppl_1, pp. D227–D233, Jan. 2010, doi: 10.1093/nar/gkp971.

[2] B. Turk, "Targeting proteases: successes, failures and future prospects," *Nat. Rev. Drug Discov.*, vol. 5, no. 9, pp. 785–799, Sep. 2006, doi: 10.1038/nrd2092.

[3] A. J. Barrett, J. F. Woessner, and N. D. Rawlings, *Handbook of Proteolytic Enzymes, Volume 1*. Elsevier, 2012.

[4] M. Fonović and B. Turk, "Cysteine cathepsins and their potential in clinical therapy and biomarker discovery," *PROTEOMICS – Clin. Appl.*, vol. 8, no. 5–6, pp. 416–426, 2014, doi: 10.1002/prca.201300085.

[5] M. Biasizzo, U. Javoršek, E. Vidak, M. Zarić, and B. Turk, "Cysteine cathepsins: A long and winding road towards clinics," *Mol. Aspects Med.*, vol. 88, p. 101150, Dec. 2022, doi: 10.1016/j.mam.2022.101150.

[6] L. Kramer, D. Turk, and B. Turk, "The Future of Cysteine Cathepsins in Disease Management," *Trends Pharmacol. Sci.*, vol. 38, no. 10, pp. 873–898, Oct. 2017, doi: 10.1016/j.tips.2017.06.003.

[7] V. Turk *et al.*, "Cysteine cathepsins: From structure, function and regulation to new frontiers," *Biochim. Biophys. Acta BBA - Proteins Proteomics*, vol. 1824, no. 1, pp. 68–88, Jan. 2012, doi: 10.1016/j.bbapap.2011.10.002.

[8] A. Rossi, Q. Deveraux, B. Turk, and A. Sali, "Comprehensive search for cysteine cathepsins in the human genome," vol. 385, no. 5, pp. 363–372, May 2004, doi: 10.1515/BC.2004.040.

[9] M. Asagiri and H. Takayanagi, "The molecular understanding of osteoclast differentiation," *Bone*, vol. 40, no. 2, pp. 251–264, Feb. 2007, doi: 10.1016/j.bone.2006.09.023.

[10] Y. Yasuda, J. Kaleta, and D. Brömme, "The role of cathepsins in osteoporosis and arthritis: Rationale for the design of new therapeutics," *Adv. Drug Deliv. Rev.*, vol. 57, no. 7, pp. 973–993, May 2005, doi: 10.1016/j.addr.2004.12.013.

[11] D. Bromme, Z. Li, M. Barnes, and E. Mehler, "Human Cathepsin V Functional Expression, Tissue Distribution, Electrostatic Surface Potential, Enzymatic Characterization, and Chromosomal Localization," p. 9.

[12] C. Linnevers, S. P. Smeekens, and D. Brömme, "Human cathepsin W, a putative cysteine protease predominantly expressed in CD8+ T-lymphocytes," *FEBS Lett.*, vol. 405, no. 3, pp. 253–259, Apr. 1997, doi: 10.1016/S0014-5793(97)00118-X.

[13] L. C. Hsing and A. Y. Rudensky, "The lysosomal cysteine proteases in MHC class II antigen presentation," *Immunol. Rev.*, vol. 207, no. 1, pp. 229–241, 2005, doi: 10.1111/j.0105-2896.2005.00310.x.

[14] I. Klemenčič *et al.*, "Biochemical characterization of human cathepsin X revealed that the enzyme is an exopeptidase, acting as carboxymonopeptidase or carboxydipeptidase," *Eur. J. Biochem.*, vol. 267, no. 17, pp. 5404–5412, 2000, doi: 10.1046/j.1432-1327.2000.01592.x.

[15] D. Musil *et al.*, "The refined 2.15 A X-ray crystal structure of human liver cathepsin B: the structural basis for its specificity.," *EMBO J.*, vol. 10, no. 9, pp. 2321–2330, Sep. 1991, doi: 10.1002/j.1460-2075.1991.tb07771.x.

[16] G. Gunčar, M. Podobnik, J. Pungerčar, BorutŠtrukelj, V. Turk, and D. Turk, "Crystal structure of porcine cathepsin H determined at 2.1 å resolution: location of the mini-chain C-terminal carboxyl group defines cathepsin H aminopeptidase function," *Structure*, vol. 6, no. 1, pp. 51–61, Jan. 1998, doi: 10.1016/S0969-2126(98)00007-0.

[17] B. Turk *et al.*, "Regulation of the Activity of Lysosomal Cysteine Proteinases by pH-Induced Inactivation and/or Endogenous Protein Inhibitors, Cystatins," *Biol. Chem. Hoppe. Seyler*, vol. 376, no. 4, pp. 225–230, Apr. 1995, doi: 10.1515/bchm3.1995.376.4.225.

[18] V. Gocheva and J. A. Joyce, "Cysteine cathepsins and the cutting edge of cancer invasion," *Cell Cycle Georget. Tex*, vol. 6, no. 1, pp. 60–64, Jan. 2007, doi: 10.4161/cc.6.1.3669.

[19] S. Tedelind *et al.*, "Nuclear cysteine cathepsin variants in thyroid carcinoma cells," vol. 391, no. 8, pp. 923–935, Aug. 2010, doi: 10.1515/bc.2010.109.

[20] E. Vidak, U. Javoršek, M. Vizovišek, and B. Turk, "Cysteine Cathepsins and Their Extracellular Roles: Shaping the Microenvironment," *Cells*, vol. 8, no. 3, Art. no. 3, Mar. 2019, doi: 10.3390/cells8030264.

[21] M. Fonović and B. Turk, "Cysteine cathepsins and extracellular matrix degradation," *Biochim. Biophys. Acta BBA - Gen. Subj.*, vol. 1840, no. 8, pp. 2560–2570, Aug. 2014, doi: 10.1016/j.bbagen.2014.03.017.

[22] D. Turk, G. Gunčar, M. Podobnik, and B. Turk, "Revised Definition of Substrate Binding Sites of Papain-Like Cysteine Proteases," vol. 379, no. 2, pp. 137–148, Jan. 1998, doi: 10.1515/bchm.1998.379.2.137.

[23] I. Schechter, "Reprint of 'On the Size of the Active Site in Proteases. I. Papain,'" *Biochem. Biophys. Res. Commun.*, vol. 425, no. 3, pp. 497–502, Aug. 2012, doi: 10.1016/j.bbrc.2012.08.015.

[24] D. Turk, B. Turk, and V. Turk, "Papain-like lysosomal cysteine proteases and their inhibitors: drug discovery targets?," *Biochem. Soc. Symp.*, vol. 70, pp. 15–30, Sep. 2003, doi: 10.1042/bss0700015.

[25] M. Vizovišek *et al.*, "Fast profiling of protease specificity reveals similar substrate specificities for cathepsins K, L and S," *PROTEOMICS*, vol. 15, no. 14, pp. 2479–2490, 2015, doi: 10.1002/pmic.201400460.

[26] R. Vidmar, M. Vizovišek, D. Turk, B. Turk, and M. Fonović, "Protease cleavage site fingerprinting by label-free in-gel degradomics reveals pH-dependent specificity switch

of legumain," *EMBO J.*, vol. 36, no. 16, pp. 2455–2465, Aug. 2017, doi: 10.15252/embj.201796750.

[27] M. L. Biniossek, D. K. Nägler, C. Becker-Pauly, and O. Schilling, "Proteomic Identification of Protease Cleavage Sites Characterizes Prime and Non-prime Specificity of Cysteine Cathepsins B, L, and S," *J. Proteome Res.*, vol. 10, no. 12, pp. 5363–5373, Dec. 2011, doi: 10.1021/pr200621z.

[28] Y. Choe *et al.*, "Substrate Profiling of Cysteine Proteases Using a Combinatorial Peptide Library Identifies Functionally Unique Specificities *," *J. Biol. Chem.*, vol. 281, no. 18, pp. 12824–12832, May 2006, doi: 10.1074/jbc.M513331200.

[29] P. Saftig and J. Klumperman, "Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function," *Nat. Rev. Mol. Cell Biol.*, vol. 10, no. 9, Art. no. 9, Sep. 2009, doi: 10.1038/nrm2745.

[30] M. F. Coutinho, M. J. Prata, and S. Alves, "A shortcut to the lysosome: The mannose-6-phosphate-independent pathway," *Mol. Genet. Metab.*, vol. 107, no. 3, pp. 257–266, Nov. 2012, doi: 10.1016/j.ymgme.2012.07.012.

[31] B. Wiederanders, G. Kaulmann, and K. Schilling, "Functions of Propeptide Parts in Cysteine Proteases," *Curr. Protein Pept. Sci.*, vol. 4, no. 5, pp. 309–326, Oct. 2003, doi: 10.2174/1389203033487081.

[32] D. Brömme, F. S. Nallaseth, and B. Turk, "Production and activation of recombinant papain-like cysteine proteases," *Methods*, vol. 32, no. 2, pp. 199–206, Feb. 2004, doi: 10.1016/S1046-2023(03)00212-3.

[33] S. W. Dahl *et al.*, "Human Recombinant Pro-dipeptidyl Peptidase I (Cathepsin C) Can Be Activated by Cathepsins L and S but Not by Autocatalytic Processing," *Biochemistry*, vol. 40, no. 6, pp. 1671–1678, Feb. 2001, doi: 10.1021/bi001693z.

[34] O. Vasiljeva, M. Dolinar, J. R. Pungerčar, V. Turk, and B. Turk, "Recombinant human procathepsin S is capable of autocatalytic processing at neutral pH in the presence of glycosaminoglycans," *FEBS Lett.*, vol. 579, no. 5, pp. 1285–1290, Feb. 2005, doi: 10.1016/j.febslet.2004.12.093.

[35] D. Caglič, J. R. Pungerčar, G. Pejler, V. Turk, and B. Turk, "Glycosaminoglycans Facilitate Procathepsin B Activation through Disruption of Propeptide-Mature Enzyme Interactions*," *J. Biol. Chem.*, vol. 282, no. 45, pp. 33076–33085, Nov. 2007, doi: 10.1074/jbc.M705761200.

[36] M. Fairhead, S. M. Kelly, and C. F. van der Walle, "A heparin binding motif on the pro-domain of human procathepsin L mediates zymogen destabilization and activation," *Biochem. Biophys. Res. Commun.*, vol. 366, no. 3, pp. 862–867, Feb. 2008, doi: 10.1016/j.bbrc.2007.12.062.

[37] T. Yadati, T. Houben, A. Bitorina, and R. Shiri-Sverdlov, "The Ins and Outs of Cathepsins: Physiological Function and Role in Disease Management," *Cells*, vol. 9, no. 7, Art. no. 7, Jul. 2020, doi: 10.3390/cells9071679.

[38] D. L. Medina *et al.*, "Transcriptional Activation of Lysosomal Exocytosis Promotes Cellular Clearance," *Dev. Cell*, vol. 21, no. 3, pp. 421–430, Sep. 2011, doi: 10.1016/j.devcel.2011.07.016.

[39] M. A. Samie and H. Xu, "Lysosomal exocytosis and lipid storage disorders," *J. Lipid Res.*, vol. 55, no. 6, pp. 995–1009, Jun. 2014, doi: 10.1194/jlr.R046896.

[40] T. Castro-Gomes, M. Corrotte, C. Tam, and N. W. Andrews, "Plasma Membrane Repair Is Regulated Extracellularly by Proteases Released from Lysosomes," *PLOS ONE*, vol. 11, no. 3, p. e0152583, Mar. 2016, doi: 10.1371/journal.pone.0152583.

[41] L. J. Olson, O. Hindsgaul, N. M. Dahms, and J.-J. P. Kim, "Structural Insights into the Mechanism of pH-dependent Ligand Binding and Release by the Cation-dependent Mannose 6-Phosphate Receptor *," *J. Biol. Chem.*, vol. 283, no. 15, pp. 10124–10134, Apr. 2008, doi: 10.1074/jbc.M708994200.

[42] A. Uhlman, K. Folkers, J. Liston, H. Pancholi, and A. Hinton, "Effects of Vacuolar H+-ATPase Inhibition on Activation of Cathepsin B and Cathepsin L Secreted from MDA-MB231 Breast Cancer Cells," *Cancer Microenviron.*, vol. 10, no. 1, pp. 49–56, Dec. 2017, doi: 10.1007/s12307-017-0196-7.

[43] M. Vizovišek, M. Fonović, and B. Turk, "Cysteine cathepsins in extracellular matrix remodeling: Extracellular matrix degradation and beyond," *Matrix Biol.*, vol. 75–76, pp. 141–159, Jan. 2019, doi: 10.1016/j.matbio.2018.01.024.

[44] A. D. Theocharis, S. S. Skandalis, C. Gialeli, and N. K. Karamanos, "Extracellular matrix structure," *Adv. Drug Deliv. Rev.*, vol. 97, pp. 4–27, Feb. 2016, doi: 10.1016/j.addr.2015.11.001.

[45] C. Bonnans, J. Chou, and Z. Werb, "Remodelling the extracellular matrix in development and disease," *Nat. Rev. Mol. Cell Biol.*, vol. 15, no. 12, Art. no. 12, Dec. 2014, doi: 10.1038/nrm3904.

[46] A. E. Page, A. R. Hayman, L. M. B. Andersson, T. J. Chambers, and M. J. Warburton, "Degradation of bone matrix proteins by osteoclast cathepsins," *Int. J. Biochem.*, vol. 25, no. 4, pp. 545–550, Apr. 1993, doi: 10.1016/0020-711X(93)90662-X.

[47] B. R. Troen, "The role of cathepsin K in normal bone resorption," *Drug News Perspect.*, vol. 17, no. 1, pp. 19–28, Jan. 2004, doi: 10.1358/dnp.2004.17.1.829022.

[48] P. Garnero *et al.*, "The Collagenolytic Activity of Cathepsin K Is Unique among Mammalian Proteinases *," *J. Biol. Chem.*, vol. 273, no. 48, pp. 32347–32352, Nov. 1998, doi: 10.1074/jbc.273.48.32347.

[49] V. Sharma *et al.*, "Structural requirements for the collagenase and elastase activity of cathepsin K and its selective inhibition by an exosite inhibitor," *Biochem. J.*, vol. 465, no. 1, pp. 163–173, Dec. 2014, doi: 10.1042/BJ20140809.

[50] H. Büth *et al.*, "Cathepsin B is essential for regeneration of scratch-wounded normal human epidermal keratinocytes," *Eur. J. Cell Biol.*, vol. 86, no. 11, pp. 747–761, Dec. 2007, doi: 10.1016/j.ejcb.2007.03.009.

[51] H. Büth *et al.*, "HaCaT keratinocytes secrete lysosomal cysteine proteinases during migration," *Eur. J. Cell Biol.*, vol. 83, no. 11, pp. 781–795, Jan. 2004, doi: 10.1078/0171-9335-00428.

[52] B. Friedrichs *et al.*, "Thyroid functions of mouse cathepsins B, K, and L," *J. Clin. Invest.*, vol. 111, no. 11, pp. 1733–1745, Jun. 2003, doi: 10.1172/JCI15990.

[53] L. Funkelstein *et al.*, "Human cathepsin V protease participates in production of enkephalin and NPY neuropeptide neurotransmitters," *J. Biol. Chem.*, vol. 287, no. 19, pp. 15232–15241, May 2012, doi: 10.1074/jbc.M111.310607.

[54] S.-R. Hwang *et al.*, "Cathepsin L Expression Is Directed to Secretory Vesicles for Enkephalin Neuropeptide Biosynthesis and Secretion *," *J. Biol. Chem.*, vol. 282, no. 13, pp. 9556–9563, Mar. 2007, doi: 10.1074/jbc.M605510200.

[55] M. Vizovišek, E. Vidak, U. Javoršek, G. Mikhaylov, A. Bratovš, and B. Turk, "Cysteine cathepsins as therapeutic targets in inflammatory diseases," *Expert Opin. Ther. Targets*, vol. 24, no. 6, pp. 573–588, Jun. 2020, doi: 10.1080/14728222.2020.1746765.

[56] W. Kafienah, D. Brömme, D. J. Buttle, L. J. Croucher, and A. P. Hollander, "Human cathepsin K cleaves native type I and II collagens at the N-terminal end of the triple helix," *Biochem. J.*, vol. 331 (Pt 3), pp. 727–732, May 1998, doi: 10.1042/bj3310727.

[57] B. J. Votta *et al.*, "Peptide aldehyde inhibitors of cathepsin K inhibit bone resorption both in vitro and in vivo," *J. Bone Miner. Res. Off. J. Am. Soc. Bone Miner. Res.*, vol. 12, no. 9, pp. 1396–1406, Sep. 1997, doi: 10.1359/jbmr.1997.12.9.1396.

[58] T. Inui *et al.*, "Cathepsin K antisense oligodeoxynucleotide inhibits osteoclastic bone resorption," *J. Biol. Chem.*, vol. 272, no. 13, pp. 8109–8112, Mar. 1997, doi: 10.1074/jbc.272.13.8109.

[59] P. Saftig *et al.*, "Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin-K-deficient mice," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 23, pp. 13453–13458, Nov. 1998, doi: 10.1073/pnas.95.23.13453.

[60] D. T. Felson and T. Neogi, "Osteoarthritis: is it a disease of cartilage or of bone?," *Arthritis Rheum.*, vol. 50, no. 2, pp. 341–344, Feb. 2004, doi: 10.1002/art.20051.

[61] S. R. Goldring, "Pathogenesis of bone and cartilage destruction in rheumatoid arthritis," *Rheumatol. Oxf. Engl.*, vol. 42 Suppl 2, pp. ii11-16, May 2003, doi: 10.1093/rheumatology/keg327.

[62] Y. T. Konttinen *et al.*, "Acidic cysteine endoproteinase cathepsin K in the degeneration of the superficial articular hyaline cartilage in osteoarthritis," *Arthritis Rheum.*, vol. 46, no. 4, pp. 953–960, Apr. 2002, doi: 10.1002/art.10185.

[63] W. Rj and H. Ac, "Control of matrix synthesis in isolated bovine chondrocytes by extracellular and intracellular pH," *J. Cell. Physiol.*, vol. 164, no. 3, Sep. 1995, doi: 10.1002/jcp.1041640305.

[64] X. W. Cheng, G.-P. Shi, M. Kuzuya, T. Sasaki, K. Okumura, and T. Murohara, "Role for cysteine protease cathepsins in heart disease: focus on biology and mechanisms with clinical implication," *Circulation*, vol. 125, no. 12, pp. 1551–1562, Mar. 2012, doi: 10.1161/CIRCULATIONAHA.111.066712.

[65] X. W. Cheng *et al.*, "Elastolytic cathepsin induction/activation system exists in myocardium and is upregulated in hypertensive heart failure," *Hypertens. Dallas Tex 1979*, vol. 48, no. 5, pp. 979–987, Nov. 2006, doi: 10.1161/01.HYP.0000242331.99369.2f.

[66] J. Stypmann *et al.*, "Dilated cardiomyopathy in mice deficient for the lysosomal cysteine peptidase cathepsin L," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 9, pp. 6234–6239, Apr. 2002, doi: 10.1073/pnas.092637699.

[67] C.-L. Liu, J. Guo, X. Zhang, G. K. Sukhova, P. Libby, and G.-P. Shi, "Cysteine protease cathepsins in cardiovascular disease: from basic research to clinical trials," *Nat. Rev. Cardiol.*, vol. 15, no. 6, Art. no. 6, Jun. 2018, doi: 10.1038/s41569-018-0002-3.

[68] S. Helske *et al.*, "Increased Expression of Elastolytic Cathepsins S, K, and V and Their Inhibitor Cystatin C in Stenotic Aortic Valves," *Arterioscler. Thromb. Vasc. Biol.*, vol. 26, no. 8, pp. 1791–1798, Aug. 2006, doi: 10.1161/01.ATV.0000228824.01604.63.

[69] J.-L. Figueiredo *et al.*, "Selective Cathepsin S Inhibition Attenuates Atherosclerosis in Apolipoprotein E–Deficient Mice with Chronic Renal Disease," *Am. J. Pathol.*, vol. 185, no. 4, pp. 1156–1166, Apr. 2015, doi: 10.1016/j.ajpath.2014.11.026.

[70] W. Li, L. Kornmark, L. Jonasson, C. Forssell, and X.-M. Yuan, "Cathepsin L is significantly associated with apoptosis and plaque destabilization in human atherosclerosis," *Atherosclerosis*, vol. 202, no. 1, pp. 92–102, Jan. 2009, doi: 10.1016/j.atherosclerosis.2008.03.027.

[71] J. Liu, G. K. Sukhova, J.-S. Sun, W.-H. Xu, P. Libby, and G.-P. Shi, "Lysosomal Cysteine Proteases in Atherosclerosis," *Arterioscler. Thromb. Vasc. Biol.*, vol. 24, no. 8, pp. 1359–1366, Aug. 2004, doi: 10.1161/01.ATV.0000134530.27208.41.

[72] K. Menzel *et al.*, "Cathepsins B, L and D in inflammatory bowel disease macrophages and potential therapeutic effects of cathepsin inhibition in vivo," *Clin. Exp. Immunol.*, vol. 146, no. 1, pp. 169–180, Oct. 2006, doi: 10.1111/j.1365-2249.2006.03188.x.

[73] T. Hirai *et al.*, "Cathepsin K Is Involved in Development of Psoriasis-like Skin Lesions through TLR-Dependent Th17 Activation," *J. Immunol.*, vol. 190, no. 9, pp. 4805–4811, May 2013, doi: 10.4049/jimmunol.1200901.

[74] A. Schönefuß *et al.*, "Upregulation of cathepsin S in psoriatic keratinocytes," *Exp. Dermatol.*, vol. 19, no. 8, pp. e80–e88, 2010, doi: 10.1111/j.1600-0625.2009.00990.x.

[75] J. S. Ainscough *et al.*, "Cathepsin S is the major activator of the psoriasis-associated proinflammatory cytokine IL-36γ," *Proc. Natl. Acad. Sci.*, vol. 114, no. 13, pp. E2748–E2757, Mar. 2017, doi: 10.1073/pnas.1620954114.

[76] P. M. Grace, M. R. Hutchinson, S. F. Maier, and L. R. Watkins, "Pathological pain and the neuroimmune interface," *Nat. Rev. Immunol.*, vol. 14, no. 4, Art. no. 4, Apr. 2014, doi: 10.1038/nri3621.

[77] B. F. Sloane, J. R. Dunn, and K. V. Honn, "Lysosomal cathepsin B: correlation with metastatic potential," *Science*, vol. 212, no. 4499, pp. 1151–1153, Jun. 1981, doi: 10.1126/science.7233209.

[78] O. Vasiljeva, T. Reinheckel, C. Peters, D. Turk, V. Turk, and B. Turk, "Emerging roles of cysteine cathepsins in disease and their potential as drug targets," *Curr. Pharm. Des.*, vol. 13, no. 4, pp. 387–403, 2007, doi: 10.2174/138161207780162962.

[79] J. Reiser, B. Adair, and T. Reinheckel, "Specialized roles for cysteine cathepsins in health and disease," *J. Clin. Invest.*, vol. 120, no. 10, pp. 3421–3431, Oct. 2010, doi: 10.1172/JCI42918.

[80] L. Sevenich *et al.*, "Synergistic antitumor effects of combined cathepsin B and cathepsin Z deficiencies on breast cancer progression and metastasis in mice," *Proc. Natl. Acad. Sci.*, vol. 107, no. 6, pp. 2497–2502, Feb. 2010, doi: 10.1073/pnas.0907240107.

[81] V. Gocheva *et al.*, "Distinct roles for cysteine cathepsin genes in multistage tumorigenesis," *Genes Dev.*, vol. 20, no. 5, pp. 543–556, Jan. 2006, doi: 10.1101/gad.1407406.

[82] L. Akkari *et al.*, "Combined deletion of cathepsin protease family members reveals compensatory mechanisms in cancer," *Genes Dev.*, vol. 30, no. 2, pp. 220–232, Jan. 2016, doi: 10.1101/gad.270439.115.

[83] J. Dennemärker *et al.*, "Deficiency for the cysteine protease cathepsin L promotes tumor progression in mouse epidermis," *Oncogene*, vol. 29, no. 11, Art. no. 11, Mar. 2010, doi: 10.1038/onc.2009.466.

[84] J. S. Shim *et al.*, "Effect of Nitroxoline on Angiogenesis and Growth of Human Bladder Cancer," *JNCI J. Natl. Cancer Inst.*, vol. 102, no. 24, pp. 1855–1873, Dec. 2010, doi: 10.1093/jnci/djq457.

[85] A. Pogorzelska, B. Żołnowska, and R. Bartoszewski, "Cysteine cathepsins as a prospective target for anticancer therapies—current progress and prospects," *Biochimie*, vol. 151, pp. 85–106, Aug. 2018, doi: 10.1016/j.biochi.2018.05.023.

[86] C. Le Gall *et al.*, "A Cathepsin K Inhibitor Reduces Breast Cancer–Induced Osteolysis and Skeletal Tumor Burden," *Cancer Res.*, vol. 67, no. 20, pp. 9894–9902, Oct. 2007, doi: 10.1158/0008-5472.CAN-06-3940.

[87] B. Arneth, "Tumor Microenvironment," *Medicina (Mex.)*, vol. 56, no. 1, Art. no. 1, Jan. 2020, doi: 10.3390/medicina56010015.

[88] M. M. Mohamed and B. F. Sloane, "multifunctional enzymes in cancer," *Nat. Rev. Cancer*, vol. 6, no. 10, Art. no. 10, Oct. 2006, doi: 10.1038/nrc1949.

[89] O. C. Olson and J. A. Joyce, "Cysteine cathepsin proteases: regulators of cancer progression and therapeutic response," *Nat. Rev. Cancer*, vol. 15, no. 12, Art. no. 12, Dec. 2015, doi: 10.1038/nrc4027.

[90] D. Yan, H.-W. Wang, R. L. Bowman, and J. A. Joyce, "STAT3 and STAT6 Signaling Pathways Synergize to Promote Cathepsin Secretion from Macrophages via IRE1α Activation," *Cell Rep.*, vol. 16, no. 11, pp. 2914–2927, Sep. 2016, doi: 10.1016/j.celrep.2016.08.035.

[91] Y. Hashimoto, C. Kondo, and N. Katunuma, "An Active 32-kDa Cathepsin L Is Secreted Directly from HT 1080 Fibrosarcoma Cells and Not via Lysosomal Exocytosis," *PLOS ONE*, vol. 10, no. 12, p. e0145067, Dec. 2015, doi: 10.1371/journal.pone.0145067.

[92] M. R. Buck, D. G. Karustis, N. A. Day, K. V. Honn, and B. F. Sloane, "Degradation of extracellular-matrix proteins by human cathepsin B from normal and tumour tissues," *Biochem. J.*, vol. 282, no. 1, pp. 273–278, Feb. 1992, doi: 10.1042/bj2820273.

[93] N. Guinec, V. Dalet-Fumeron, and M. Pagano, "'In vitro' Study of Basement Membrane Degradation by the Cysteine Proteinases, Cathepsins B, B-Like and L. Digestion of Collagen IV, Laminin, Fibronectin, and Release of Gelatinase Activities front Basement Membrane Fibronectin," vol. 374, no. 7–12, pp. 1135–1146, Jan. 1993, doi: 10.1515/bchm3.1993.374.7-12.1135.

[94] K. Ishidoh and E. Kominami, "Procathepsin L Degrades Extracellular Matrix Proteins in the Presence of Glycosaminoglycans in Vitro," *Biochem. Biophys. Res. Commun.*, vol. 217, no. 2, pp. 624–631, Dec. 1995, doi: 10.1006/bbrc.1995.2820.

[95] A. Kudo, "Periostin in Bone Biology," in *Periostin*, A. Kudo, Ed., in Advances in Experimental Medicine and Biology. , Singapore: Springer, 2019, pp. 43–47. doi: 10.1007/978-981-13-6657-4_5.

[96] E. Gineyts *et al.*, "The C-Terminal Intact Forms of Periostin (iPTN) Are Surrogate Markers for Osteolytic Lesions in Experimental Breast Cancer Bone Metastasis," *Calcif. Tissue Int.*, vol. 103, no. 5, pp. 567–580, Nov. 2018, doi: 10.1007/s00223-018-0444-y.

[97] B. Sobotič *et al.*, "Proteomic Identification of Cysteine Cathepsin Substrates Shed from the Surface of Cancer Cells *," *Mol. Cell. Proteomics*, vol. 14, no. 8, pp. 2213–2228, Aug. 2015, doi: 10.1074/mcp.M114.044628.

[98] R. Roskoski, "The ErbB/HER family of protein-tyrosine kinases and cancer," *Pharmacol. Res.*, vol. 79, pp. 34–74, Jan. 2014, doi: 10.1016/j.phrs.2013.11.002.

[99] M. A. Lemmon and J. Schlessinger, "Cell Signaling by Receptor Tyrosine Kinases," *Cell*, vol. 141, no. 7, pp. 1117–1134, Jun. 2010, doi: 10.1016/j.cell.2010.06.011.

[100] R. Trenker and N. Jura, "Receptor tyrosine kinase activation: From the ligand perspective," *Curr. Opin. Cell Biol.*, vol. 63, pp. 174–185, Apr. 2020, doi: 10.1016/j.ceb.2020.01.016.

[101] S. Cohen, "Isolation of a Mouse Submaxillary Gland Protein Accelerating Incisor Eruption and Eyelid Opening in the New-born Animal," *J. Biol. Chem.*, vol. 237, no. 5, pp. 1555–1562, May 1962, doi: 10.1016/S0021-9258(19)83739-0.

[102] L. N. Klapper *et al.*, "The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors," *Proc. Natl. Acad. Sci.*, vol. 96, no. 9, pp. 4995–5000, Apr. 1999, doi: 10.1073/pnas.96.9.4995.

[103] F. Shi, S. E. Telesco, Y. Liu, R. Radhakrishnan, and M. A. Lemmon, "ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation," *Proc. Natl. Acad. Sci.*, vol. 107, no. 17, pp. 7692–7697, Apr. 2010, doi: 10.1073/pnas.1002753107.

[104] E. Kovacs, J. A. Zorn, Y. Huang, T. Barros, and J. Kuriyan, "A Structural Perspective on the Regulation of the EGF Receptor," *Annu. Rev. Biochem.*, vol. 84, pp. 739–764, Jun. 2015, doi: 10.1146/annurev-biochem-060614-034402.

[105] M. A. Lemmon, J. Schlessinger, and K. M. Ferguson, "The EGFR Family: Not So Prototypical Receptor Tyrosine Kinases," *Cold Spring Harb. Perspect. Biol.*, vol. 6, no. 4, p. a020768, Jan. 2014, doi: 10.1101/cshperspect.a020768.

[106] T. Amelia, R. E. Kartasasmita, T. Ohwada, and D. H. Tjahjono, "Structural Insight and Development of EGFR Tyrosine Kinase Inhibitors," *Molecules*, vol. 27, no. 3, Art. no. 3, Jan. 2022, doi: 10.3390/molecules27030819.

[107] L. He and K. Hristova, "Consequences of replacing EGFR juxtamembrane domain with an unstructured sequence," *Sci. Rep.*, vol. 2, no. 1, Art. no. 1, Nov. 2012, doi: 10.1038/srep00854.

[108] D.-H. Kim, H. M. Triet, and S. H. Ryu, "Regulation of EGFR activation and signaling by lipids on the plasma membrane," *Prog. Lipid Res.*, vol. 83, p. 101115, Jul. 2021, doi: 10.1016/j.plipres.2021.101115.

[109] J. Stamos, M. X. Sliwkowski, and C. Eigenbrot, "Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor," *J. Biol. Chem.*, vol. 277, no. 48, pp. 46265–46272, Nov. 2002, doi: 10.1074/jbc.M207135200.

[110] N. Jura, X. Zhang, N. F. Endres, M. A. Seeliger, T. Schindler, and J. Kuriyan, "Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms," *Mol. Cell*, vol. 42, no. 1, pp. 9–22, Apr. 2011, doi: 10.1016/j.molcel.2011.03.004.

[111] E. Kovacs *et al.*, "Analysis of the Role of the C-Terminal Tail in the Regulation of the Epidermal Growth Factor Receptor," *Mol. Cell. Biol.*, vol. 35, no. 17, pp. 3083–3102, Sep. 2015, doi: 10.1128/MCB.00248-15.

[112] M. Landau, S. J. Fleishman, and N. Ben-Tal, "A Putative Mechanism for Downregulation of the Catalytic Activity of the EGF Receptor via Direct Contact between Its Kinase and C-Terminal Domains," *Structure*, vol. 12, no. 12, pp. 2265–2275, Dec. 2004, doi: 10.1016/j.str.2004.10.006.

[113] R. C. Harris, E. Chung, and R. J. Coffey, "- EGF receptor ligands," in *The EGF Receptor Family*, G. Carpenter, Ed., Burlington: Academic Press, 2003, pp. 3–14. doi: 10.1016/B978-012160281-9/50002-5.

[114] M. R. Schneider and E. Wolf, "The epidermal growth factor receptor ligands at a glance," *J. Cell. Physiol.*, vol. 218, no. 3, pp. 460–466, 2009, doi: 10.1002/jcp.21635.

[115] R. A. Mitchell, R. B. Luwor, and A. W. Burgess, "Epidermal growth factor receptor: Structure-function informing the design of anticancer therapeutics," *Exp. Cell Res.*, vol. 371, no. 1, pp. 1–19, Oct. 2018, doi: 10.1016/j.yexcr.2018.08.009.

[116] H. E. Abud, W. H. Chan, and T. Jardé, "Source and Impact of the EGF Family of Ligands on Intestinal Stem Cells," *Front. Cell Dev. Biol.*, vol. 9, 2021, Accessed: Oct. 12, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fcell.2021.685665

[117] Z. Wang, "ErbB Receptors and Cancer," in *ErbB Receptor Signaling: Methods and Protocols*, Z. Wang, Ed., in Methods in Molecular Biology. , New York, NY: Springer, 2017, pp. 3–35. doi: 10.1007/978-1-4939-7219-7_1.

[118] C. P. Blobel, "ADAMs: key components in EGFR signalling and development," *Nat. Rev. Mol. Cell Biol.*, vol. 6, no. 1, Art. no. 1, Jan. 2005, doi: 10.1038/nrm1548.

[119] J. N. Higginbotham *et al.*, "Identification and characterization of EGF receptor in individual exosomes by fluorescence-activated vesicle sorting," *J. Extracell. Vesicles*, vol. 5, no. 1, p. 29254, Jan. 2016, doi: 10.3402/jev.v5.29254.

[120] B. Singh, G. Bogatcheva, M. K. Washington, and R. J. Coffey, "Transformation of polarized epithelial cells by apical mistrafficking of epiregulin," *Proc. Natl. Acad. Sci.*, vol. 110, no. 22, pp. 8960–8965, May 2013, doi: 10.1073/pnas.1305508110.

[121] A. Baker, A. Zlobin, and C. Osipo, "Notch-EGFR/HER2 bidirectional crosstalk in breast cancer," *Front. Oncol.*, vol. 4, no. NOV, 2014, doi: 10.3389/fonc.2014.00360.

[122] M. Köse, "GPCRs and EGFR – Cross-talk of membrane receptors in cancer," *Bioorg. Med. Chem. Lett.*, vol. 27, no. 16, pp. 3611–3620, Aug. 2017, doi: 10.1016/j.bmcl.2017.07.002.

[123] M. C. Franklin, K. D. Carey, F. F. Vajdos, D. J. Leahy, A. M. de Vos, and M. X. Sliwkowski, "Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex," *Cancer Cell*, vol. 5, no. 4, pp. 317–328, Apr. 2004, doi: 10.1016/S1535-6108(04)00083-2.

[124] N. Gotoh, A. Tojo, M. Hino, Y. Yazaki, and M. Shibuya, "A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming activity of human epidermal growth factor receptor," *Biochem. Biophys. Res. Commun.*, vol. 186, no. 2, pp. 768–774, Jul. 1992, doi: 10.1016/0006-291X(92)90812-Y.

[125] X. Zhang, J. Gureasko, K. Shen, P. A. Cole, and J. Kuriyan, "An Allosteric Mechanism for Activation of the Kinase Domain of Epidermal Growth Factor Receptor," *Cell*, vol. 125, no. 6, pp. 1137–1149, Jun. 2006, doi: 10.1016/j.cell.2006.05.013.

[126] A. W. Burgess *et al.*, "An Open-and-Shut Case? Recent Insights into the Activation of EGF/ErbB Receptors," *Mol. Cell*, vol. 12, no. 3, pp. 541–552, Sep. 2003, doi: 10.1016/S1097-2765(03)00350-2.

[127] R. N. Jorissen, F. Walker, N. Pouliot, T. P. J. Garrett, C. W. Ward, and A. W. Burgess, "- Epidermal growth factor receptor: Mechanisms of activation and signalling," in *The EGF Receptor Family*, G. Carpenter, Ed., Burlington: Academic Press, 2003, pp. 33–55. doi: 10.1016/B978-012160281-9/50004-9.

[128] Y. Yarden and M. X. Sliwkowski, "Untangling the ErbB signalling network," *Nat. Rev. Mol. Cell Biol.*, vol. 2, no. 2, pp. 127–137, 2001, doi: 10.1038/35052073.

[129] Y. D. Shaul and R. Seger, "The MEK/ERK cascade: from signaling specificity to diverse functions," *Biochim. Biophys. Acta*, vol. 1773, no. 8, pp. 1213–1226, Aug. 2007, doi: 10.1016/j.bbamcr.2006.10.005.

[130] M. Katz, I. Amit, and Y. Yarden, "Regulation of MAPKs by growth factors and receptor tyrosine kinases," *Biochim. Biophys. Acta*, vol. 1773, no. 8, pp. 1161–1176, Aug. 2007, doi: 10.1016/j.bbamcr.2007.01.002.

[131] N. Hay and N. Sonenberg, "Upstream and downstream of mTOR," *Genes Dev.*, vol. 18, no. 16, pp. 1926–1945, Aug. 2004, doi: 10.1101/gad.1212704.

[132] J. A. Engelman, J. Luo, and L. C. Cantley, "The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism," *Nat. Rev. Genet.*, vol. 7, no. 8, Art. no. 8, Aug. 2006, doi: 10.1038/nrg1879.

[133] M. Laplante and D. M. Sabatini, "mTOR Signaling in Growth Control and Disease," *Cell*, vol. 149, no. 2, pp. 274–293, Apr. 2012, doi: 10.1016/j.cell.2012.03.017.

[134] D. Mochly-Rosen, K. Das, and K. V. Grimes, "Protein kinase C, an elusive therapeutic target?," *Nat. Rev. Drug Discov.*, vol. 11, no. 12, Art. no. 12, Dec. 2012, doi: 10.1038/nrd3871.

[135] G. Kadamur and E. M. Ross, "Mammalian Phospholipase C," *Annu. Rev. Physiol.*, vol. 75, no. 1, pp. 127–154, Feb. 2013, doi: 10.1146/annurev-physiol-030212-183750.

[136] K. M. Quesnelle, A. L. Boehm, and J. R. Grandis, "STAT-mediated EGFR signaling in cancer," *J. Cell. Biochem.*, vol. 102, no. 2, pp. 311–319, 2007, doi: 10.1002/jcb.21475.

[137] F. Erdogan *et al.*, "JAK-STAT core cancer pathway: An integrative cancer interactome analysis," *J. Cell. Mol. Med.*, vol. 26, no. 7, pp. 2049–2062, 2022, doi: 10.1111/jcmm.17228.

[138] E. R. Purba, E. Saita, and I. N. Maruyama, "Activation of the EGF Receptor by Ligand Binding and Oncogenic Mutations: The 'Rotation Model,'" *Cells*, vol. 6, no. 2, Art. no. 2, Jun. 2017, doi: 10.3390/cells6020013.

[139] N. Prenzel, O. M. Fischer, S. Streit, S. Hart, and A. Ullrich, "The epidermal growth factor receptor family as a central element for cellular signal transduction and diversification," *Endocr. Relat. Cancer*, vol. 8, no. 1, pp. 11–31, Mar. 2001, doi: 10.1677/erc.0.0080011.

[140] R. Zandi, A. B. Larsen, P. Andersen, M.-T. Stockhausen, and H. S. Poulsen, "Mechanisms for oncogenic activation of the epidermal growth factor receptor," *Cell. Signal.*, vol. 19, no. 10, pp. 2013–2023, Oct. 2007, doi: 10.1016/j.cellsig.2007.06.023.

[141] A. E. G. Lenferink *et al.*, "Differential endocytic routing of homo- and hetero-dimeric ErbB tyrosine kinases confers signaling superiority to receptor heterodimers," *EMBO J.*, vol. 17, no. 12, pp. 3385–3397, 1998, doi: 10.1093/emboj/17.12.3385.

[142] T. P. J. Garrett *et al.*, "The Crystal Structure of a Truncated ErbB2 Ectodomain Reveals an Active Conformation, Poised to Interact with Other ErbB Receptors," *Mol. Cell*, vol. 11, no. 2, pp. 495–505, Feb. 2003, doi: 10.1016/S1097-2765(03)00048-0.

[143] M. A. Olayioye, R. M. Neve, H. A. Lane, and N. E. Hynes, "The ErbB signaling network: receptor heterodimerization in development and cancer," *EMBO J.*, vol. 19, no. 13, pp. 3159–3167, Jul. 2000, doi: 10.1093/emboj/19.13.3159.

[144] L. Moro *et al.*, "Integrins induce activation of EGF receptor: role in MAP kinase induction and adhesion-dependent cell survival," *EMBO J.*, vol. 17, no. 22, pp. 6622–6632, Nov. 1998, doi: 10.1093/emboj/17.22.6622.

[145] J. H. Ludes-Meyers *et al.*, "Transcriptional activation of the human epidermal growth factor receptor promoter by human p53," *Mol. Cell. Biol.*, vol. 16, no. 11, pp. 6009–6019, 1996, doi: 10.1128/MCB.16.11.6009.

[146] A. V. Vieira, C. Lamaze, and S. L. Schmid, "Control of EGF receptor signaling by clathrin-mediated endocytosis," *Science*, vol. 274, no. 5295, pp. 2086–2089, 1996, doi: 10.1126/science.274.5295.2086.

[147] L. P. Sousa, I. Lax, H. Shen, S. M. Ferguson, P. D. Camilli, and J. Schlessinger, "Suppression of EGFR endocytosis by dynamin depletion reveals that EGFR signaling occurs primarily at the plasma membrane," *Proc. Natl. Acad. Sci.*, vol. 109, no. 12, pp. 4419–4424, Mar. 2012, doi: 10.1073/pnas.1200164109.

[148] D. Johnston, H. Hall, T. P. DiLorenzo, and B. M. Steinberg, "Elevation of the epidermal growth factor receptor and dependent signaling in human papillomavirus-infected laryngeal papillomas," *Cancer Res.*, vol. 59, no. 4, pp. 968–974, Feb. 1999.

[149] A. J. Wong *et al.*, "Structural alterations of the epidermal growth factor receptor gene in human gliomas," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 89, no. 7, pp. 2965–2969, Apr. 1992, doi: 10.1073/pnas.89.7.2965.

[150] L. Frederick, X. Y. Wang, G. Eley, and C. D. James, "Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas," *Cancer Res.*, vol. 60, no. 5, pp. 1383–1387, Mar. 2000.

[151] A. Rutkowska, E. Stoczyńska-Fidelus, K. Janik, A. Włodarczyk, and P. Rieske, "$EGFR^{vIII}$: An Oncogene with Ambiguous Role," *J. Oncol.*, vol. 2019, p. e1092587, Dec. 2019, doi: 10.1155/2019/1092587.

[152] Z. An, O. Aksoy, T. Zheng, Q.-W. Fan, and W. A. Weiss, "Epidermal growth factor receptor and EGFRvIII in glioblastoma: signaling pathways and targeted therapies," *Oncogene*, vol. 37, no. 12, Art. no. 12, Mar. 2018, doi: 10.1038/s41388-017-0045-7.

[153] A. J. Ekstrand, N. Sugawa, C. D. James, and V. P. Collins, "Amplified and rearranged epidermal growth factor receptor genes in human glioblastomas reveal deletions of sequences encoding portions of the N- and/or C-terminal tails," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 89, no. 10, pp. 4309–4313, May 1992, doi: 10.1073/pnas.89.10.4309.

[154] H. Shigematsu *et al.*, "Clinical and Biological Features Associated With Epidermal Growth Factor Receptor Gene Mutations in Lung Cancers," *JNCI J. Natl. Cancer Inst.*, vol. 97, no. 5, pp. 339–346, Mar. 2005, doi: 10.1093/jnci/dji055.

[155] H. Shigematsu and A. F. Gazdar, "Somatic mutations of epidermal growth factor receptor signaling pathway in lung cancers," *Int. J. Cancer*, vol. 118, no. 2, pp. 257–262, 2006, doi: 10.1002/ijc.21496.

[156] A. Passaro, T. Mok, S. Peters, S. Popat, M.-J. Ahn, and F. de Marinis, "Recent Advances on the Role of EGFR Tyrosine Kinase Inhibitors in the Management of NSCLC With Uncommon, Non Exon 20 Insertions, EGFR Mutations," *J. Thorac. Oncol.*, vol. 16, no. 5, pp. 764–773, May 2021, doi: 10.1016/j.jtho.2020.12.002.

[157] N. F. Endres *et al.*, "Conformational Coupling across the Plasma Membrane in Activation of the EGF Receptor," *Cell*, vol. 152, no. 3, pp. 543–556, Jan. 2013, doi: 10.1016/j.cell.2012.12.032.

[158] D. B. Ramnarain *et al.*, "Differential Gene Expression Analysis Reveals Generation of an Autocrine Loop by a Mutant Epidermal Growth Factor Receptor in Glioma Cells," *Cancer Res.*, vol. 66, no. 2, pp. 867–874, Jan. 2006, doi: 10.1158/0008-5472.CAN-05-2753.

[159] S. Chakraborty *et al.*, "Constitutive and ligand-induced EGFR signalling triggers distinct and mutually exclusive downstream signalling networks," *Nat. Commun.*, vol. 5, no. 1, Art. no. 1, Dec. 2014, doi: 10.1038/ncomms6811.

[160] G. Guo, K. Gong, B. Wohlfeld, K. J. Hatanpaa, D. Zhao, and A. A. Habib, "Ligand-Independent EGFR Signaling," *Cancer Res.*, vol. 75, no. 17, pp. 3436–3441, Aug. 2015, doi: 10.1158/0008-5472.CAN-15-0989.

[161] H. S. Huang *et al.*, "The enhanced tumorigenic activity of a mutant epidermal growth factor receptor common in human cancers is mediated by threshold levels of constitutive tyrosine phosphorylation and unattenuated signaling," *J. Biol. Chem.*, vol. 272, no. 5, pp. 2927–2935, Jan. 1997, doi: 10.1074/jbc.272.5.2927.

[162] M. V. Grandal, R. Zandi, M. W. Pedersen, B. M. Willumsen, B. van Deurs, and H. S. Poulsen, "EGFRvIII escapes down-regulation due to impaired internalization and sorting to lysosomes," *Carcinogenesis*, vol. 28, no. 7, pp. 1408–1417, Jul. 2007, doi: 10.1093/carcin/bgm058.

[163] P. H. Huang, A. M. Xu, and F. M. White, "Oncogenic EGFR Signaling Networks in Glioma," *Sci. Signal.*, vol. 2, no. 87, pp. re6–re6, Sep. 2009, doi: 10.1126/scisignal.287re6.

[164] M. Chen, L.-M. Chen, C.-Y. Lin, and K. X. Chai, "The epidermal growth factor receptor (EGFR) is proteolytically modified by the Matriptase–Prostasin serine protease cascade in cultured epithelial cells," *Biochim. Biophys. Acta BBA - Mol. Cell Res.*, vol. 1783, no. 5, pp. 896–903, May 2008, doi: 10.1016/j.bbamcr.2007.10.019.

[165] C.-C. Huang, C.-C. Lee, H.-H. Lin, and J.-Y. Chang, "Cathepsin S attenuates endosomal EGFR signalling: A mechanical rationale for the combination of cathepsin S and EGFR tyrosine kinase inhibitors," *Sci. Rep.*, vol. 6, no. 1, Art. no. 1, Jul. 2016, doi: 10.1038/srep29256.

[166] T. M. Brand, M. Iida, N. Luthar, M. M. Starr, E. J. Huppert, and D. L. Wheeler, "Nuclear EGFR as a molecular target in cancer," *Radiother. Oncol.*, vol. 108, no. 3, pp. 370–377, Sep. 2013, doi: 10.1016/j.radonc.2013.06.010.

[167] T. Yamaoka, M. Ohba, and T. Ohmori, "Molecular-Targeted Therapies for Epidermal Growth Factor Receptor and Its Resistance Mechanisms," *Int. J. Mol. Sci.*, vol. 18, no. 11, Art. no. 11, Nov. 2017, doi: 10.3390/ijms18112420.

[168] Y.-N. Wang, H. Yamaguchi, J.-M. Hsu, and M.-C. Hung, "Nuclear trafficking of the epidermal growth factor receptor family membrane proteins," *Oncogene*, vol. 29, no. 28, Art. no. 28, Jul. 2010, doi: 10.1038/onc.2010.157.

[169] L. Xu and J. Massagué, "Nucleocytoplasmic shuttling of signal transducers," *Nat. Rev. Mol. Cell Biol.*, vol. 5, no. 3, pp. 209–219, 2004, doi: 10.1038/nrm1331.

[170] H.-W. Lo, M. Ali-Seyed, Y. Wu, G. Bartholomeusz, S.-C. Hsu, and M.-C. Hung, "Nuclear-cytoplasmic transport of EGFR involves receptor endocytosis, importin β1 and CRM1," *J. Cell. Biochem.*, vol. 98, no. 6, pp. 1570–1583, 2006, doi: 10.1002/jcb.20876.

[171] P. Shah, A. Chaumet, S. J. Royle, and F. A. Bard, "The NAE Pathway: Autobahn to the Nucleus for Cell Surface Receptors," *Cells*, vol. 8, no. 8, Art. no. 8, Aug. 2019, doi: 10.3390/cells8080915.

[172] A. Psyrri *et al.*, "Quantitative determination of nuclear and cytoplasmic epidermal growth factor receptor expression in oropharyngeal squamous cell cancer by using automated quantitative analysis," *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, vol. 11, no. 16, pp. 5856–5862, Aug. 2005, doi: 10.1158/1078-0432.CCR-05-0420.

[173] A. M. Traynor *et al.*, "Nuclear EGFR protein expression predicts poor survival in early stage non-small cell lung cancer," *Lung Cancer Amst. Neth.*, vol. 81, no. 1, pp. 138–141, Jul. 2013, doi: 10.1016/j.lungcan.2013.03.020.

[174] S.-C. Wang and M.-C. Hung, "Nuclear translocation of the epidermal growth factor receptor family membrane tyrosine kinase receptors," *Clin. Cancer Res.*, vol. 15, no. 21, pp. 6484–6489, 2009, doi: 10.1158/1078-0432.CCR-08-2813.

[175] H.-W. Lo *et al.*, "Nuclear interaction of EGFR and STAT3 in the activation of the iNOS/NO pathway," *Cancer Cell*, vol. 7, no. 6, pp. 575–589, Jun. 2005, doi: 10.1016/j.ccr.2005.05.007.

[176] H.-W. Lo, X. Cao, H. Zhu, and F. Ali-Osman, "Cyclooxygenase-2 is a novel transcriptional target of the nuclear EGFR-STAT3 and EGFRvIII-STAT3 signaling axes," *Mol. Cancer Res. MCR*, vol. 8, no. 2, pp. 232–245, Feb. 2010, doi: 10.1158/1541-7786.MCR-09-0391.

[177] M. L. Uribe, I. Marrocco, and Y. Yarden, "EGFR in Cancer: Signaling Mechanisms, Drugs, and Acquired Resistance," *Cancers*, vol. 13, no. 11, Art. no. 11, Jan. 2021, doi: 10.3390/cancers13112748.

[178] D. Sabbah, R. Hajjo, and K. Sweidan, "Review on Epidermal Growth Factor Receptor (EGFR) Structure, Signaling Pathways, Interactions, and Recent Updates of

EGFR Inhibitors," *Curr. Top. Med. Chem.*, vol. 20, Mar. 2020, doi: 10.2174/1568026620666200303123102.

[179] T. M. Brand, M. Iida, and D. L. Wheeler, "Molecular mechanisms of resistance to the EGFR monoclonal antibody cetuximab," *Cancer Biol. Ther.*, vol. 11, no. 9, pp. 777–792, May 2011, doi: 10.4161/cbt.11.9.15050.

[180] G. Fornasier, S. Francescon, and P. Baldo, "An Update of Efficacy and Safety of Cetuximab in Metastatic Colorectal Cancer: A Narrative Review," *Adv. Ther.*, vol. 35, no. 10, pp. 1497–1509, Oct. 2018, doi: 10.1007/s12325-018-0791-0.

[181] M. Steins, M. Thomas, and M. Geißler, "Erlotinib," in *Small Molecules in Oncology*, U. M. Martens, Ed., in Recent Results in Cancer Research. , Cham: Springer International Publishing, 2018, pp. 1–17. doi: 10.1007/978-3-319-91442-8_1.

[182] A. Ayati, S. Moghimi, S. Salarinejad, M. Safavi, B. Pouramiri, and A. Foroumadi, "A review on progression of epidermal growth factor receptor (EGFR) inhibitors as an efficient approach in cancer targeted therapy," *Bioorganic Chem.*, vol. 99, p. 103811, Jun. 2020, doi: 10.1016/j.bioorg.2020.103811.

[183] W.-Q. Cai *et al.*, "The Latest Battles Between EGFR Monoclonal Antibodies and Resistant Tumor Cells," *Front. Oncol.*, vol. 10, 2020, Accessed: Nov. 07, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fonc.2020.01249

[184] S. Misale *et al.*, "Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer," *Nature*, vol. 486, no. 7404, Art. no. 7404, Jun. 2012, doi: 10.1038/nature11156.

[185] N. Normanno, S. Tejpar, F. Morgillo, A. De Luca, E. Van Cutsem, and F. Ciardiello, "Implications for KRAS status and EGFR-targeted therapies in metastatic CRC," *Nat. Rev. Clin. Oncol.*, vol. 6, no. 9, Art. no. 9, Sep. 2009, doi: 10.1038/nrclinonc.2009.111.

[186] A. Bardelli *et al.*, "Amplification of the MET Receptor Drives Resistance to Anti-EGFR Therapies in Colorectal Cancer," *Cancer Discov.*, vol. 3, no. 6, pp. 658–673, Jun. 2013, doi: 10.1158/2159-8290.CD-12-0558.

[187] D. L. Wheeler *et al.*, "Mechanisms of acquired resistance to cetuximab: role of HER (ErbB) family members," *Oncogene*, vol. 27, no. 28, Art. no. 28, Jun. 2008, doi: 10.1038/onc.2008.19.

[188] K. Yonesaka *et al.*, "Activation of ERBB2 Signaling Causes Resistance to the EGFR-Directed Therapeutic Antibody Cetuximab," *Sci. Transl. Med.*, vol. 3, no. 99, pp. 99ra86-99ra86, Sep. 2011, doi: 10.1126/scitranslmed.3002442.

[189] J. C. Sok *et al.*, "Mutant Epidermal Growth Factor Receptor (EGFRvIII) Contributes to Head and Neck Cancer Growth and Resistance to EGFR Targeting," *Clin. Cancer Res.*, vol. 12, no. 17, pp. 5064–5073, Sep. 2006, doi: 10.1158/1078-0432.CCR-06-0913.

[190] A. A. Zulkifli, F. H. Tan, T. L. Putoczki, S. S. Stylli, and R. B. Luwor, "STAT3 signaling mediates tumour resistance to EGFR targeted therapeutics," *Mol. Cell. Endocrinol.*, vol. 451, pp. 15–23, Aug. 2017, doi: 10.1016/j.mce.2017.01.010.

[191] W. Pao *et al.*, "Acquired Resistance of Lung Adenocarcinomas to Gefitinib or Erlotinib Is Associated with a Second Mutation in the EGFR Kinase Domain," *PLOS Med.*, vol. 2, no. 3, p. e73, Feb. 2005, doi: 10.1371/journal.pmed.0020073.

[192] S. Kobayashi *et al.*, "EGFR Mutation and Resistance of Non–Small-Cell Lung Cancer to Gefitinib," *N. Engl. J. Med.*, vol. 352, no. 8, pp. 786–792, Feb. 2005, doi: 10.1056/NEJMoa044238.

[193] C.-H. Yun *et al.*, "The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP," *Proc. Natl. Acad. Sci.*, vol. 105, no. 6, pp. 2070–2075, Feb. 2008, doi: 10.1073/pnas.0709662105.

[194] B. Ko, D. Paucar, and B. Halmos, "EGFR T790M: revealing the secrets of a gatekeeper," *Lung Cancer Targets Ther.*, vol. 8, pp. 147–159, Dec. 2017, doi: 10.2147/LCTT.S117944.

[195] T. J. Lynch *et al.*, "Activating Mutations in the Epidermal Growth Factor Receptor Underlying Responsiveness of Non–Small-Cell Lung Cancer to Gefitinib," *N. Engl. J. Med.*, vol. 350, no. 21, pp. 2129–2139, May 2004, doi: 10.1056/NEJMoa040938.

[196] J. A. Engelman *et al.*, "MET Amplification Leads to Gefitinib Resistance in Lung Cancer by Activating ERBB3 Signaling," *Science*, vol. 316, no. 5827, pp. 1039–1043, May 2007, doi: 10.1126/science.1141478.

[197] L. V. Sequist *et al.*, "Randomized Phase II Study of Erlotinib Plus Tivantinib Versus Erlotinib Plus Placebo in Previously Treated Non–Small-Cell Lung Cancer," *J. Clin. Oncol.*, vol. 29, no. 24, pp. 3307–3315, Aug. 2011, doi: 10.1200/JCO.2010.34.0570.

[198] K. Takezawa *et al.*, "HER2 Amplification: A Potential Mechanism of Acquired Resistance to EGFR Inhibition in EGFR-Mutant Lung Cancers That Lack the Second-Site EGFRT790M Mutation," *Cancer Discov.*, vol. 2, no. 10, pp. 922–933, Oct. 2012, doi: 10.1158/2159-8290.CD-12-0108.

[199] T. Trowe *et al.*, "EXEL-7647 Inhibits Mutant Forms of ErbB2 Associated with Lapatinib Resistance and Neoplastic Transformation," *Clin. Cancer Res.*, vol. 14, no. 8, pp. 2465–2475, Apr. 2008, doi: 10.1158/1078-0432.CCR-07-4367.

[200] N. V. Sergina *et al.*, "Escape from HER-family tyrosine kinase inhibitor therapy by the kinase-inactive HER3," *Nature*, vol. 445, no. 7126, Art. no. 7126, Jan. 2007, doi: 10.1038/nature05474.

[201] K. E. Ware *et al.*, "Rapidly Acquired Resistance to EGFR Tyrosine Kinase Inhibitors in NSCLC Cell Lines through De-Repression of FGFR2 and FGFR3 Expression," *PLOS ONE*, vol. 5, no. 11, p. e14117, Nov. 2010, doi: 10.1371/journal.pone.0014117.

[202] Z. Zhang *et al.*, "Activation of the AXL kinase causes resistance to EGFR-targeted therapy in lung cancer," *Nat. Genet.*, vol. 44, no. 8, Art. no. 8, Aug. 2012, doi: 10.1038/ng.2330.

[203] F. Morgillo, J. K. Woo, E. S. Kim, W. K. Hong, and H.-Y. Lee, "Heterodimerization of Insulin-like Growth Factor Receptor/Epidermal Growth Factor Receptor and Induction of Survivin Expression Counteract the Antitumor Action of Erlotinib," *Cancer Res.*, vol. 66, no. 20, pp. 10100–10111, Oct. 2006, doi: 10.1158/0008-5472.CAN-06-1684.

[204] D. B. Costa *et al.*, "BIM Mediates EGFR Tyrosine Kinase Inhibitor-Induced Apoptosis in Lung Cancers with Oncogenic EGFR Mutations," *PLOS Med.*, vol. 4, no. 10, p. e315, okt 2007, doi: 10.1371/journal.pmed.0040315.

[205] J. Deng *et al.*, "Proapoptotic BH3-Only BCL-2 Family Protein BIM Connects Death Signaling from Epidermal Growth Factor Receptor Inhibition to the Mitochondrion," *Cancer Res.*, vol. 67, no. 24, pp. 11867–11875, Dec. 2007, doi: 10.1158/0008-5472.CAN-07-1961.

[206] A. C. Faber *et al.*, "BIM Expression in Treatment-Naïve Cancers Predicts Responsiveness to Kinase Inhibitors," *Cancer Discov.*, vol. 1, no. 4, pp. 352–365, Sep. 2011, doi: 10.1158/2159-8290.CD-11-0106.

[207] A. Staes *et al.*, "Selecting protein N-terminal peptides by combined fractional diagonal chromatography," *Nat. Protoc.*, vol. 6, no. 8, Art. no. 8, Aug. 2011, doi: 10.1038/nprot.2011.355.

[208] M. Mihelič, A. Doberšek, G. Gunčar, and D. Turk, "Inhibitory Fragment from the p41 Form of Invariant Chain Can Regulate Activity of Cysteine Cathepsins in Antigen Presentation," *J. Biol. Chem.*, vol. 283, no. 21, pp. 14453–14460, May 2008, doi: 10.1074/jbc.M801283200.

[209] K. Suzuki, P. Bose, R. Y. Leong-Quong, D. J. Fujita, and K. Riabowol, "REAP: A two minute cell fractionation method," *BMC Res. Notes*, vol. 3, no. 1, p. 294, Nov. 2010, doi: 10.1186/1756-0500-3-294.

[210] L. C. Crowley, B. J. Marfell, A. P. Scott, and N. J. Waterhouse, "Quantitation of Apoptosis and Necrosis by Annexin V Binding, Propidium Iodide Uptake, and Flow Cytometry," *Cold Spring Harb. Protoc.*, vol. 2016, no. 11, p. pdb.prot087288, Jan. 2016, doi: 10.1101/pdb.prot087288.

[211] H.-C. Chen, "Boyden Chamber Assay," in *Cell Migration: Developmental Methods and Protocols*, J.-L. Guan, Ed., in Methods in Molecular Biology™. , Totowa, NJ: Humana Press, 2005, pp. 15–22. doi: 10.1385/1-59259-860-9:015.

[212] K. D. Smith, M. J. Davies, D. Bailey, D. V. Renouf, and E. F. Hounsell, "Analysis of the Glycosylation Patterns of the Extracellular Domain of the Epidermal Growth Factor Receptor Expressed in Chinese Hamster Ovary Fibroblasts," *Growth Factors*, vol. 13, no. 1–2, pp. 121–132, Jan. 1996, doi: 10.3109/08977199609034572.

[213] Y. Zhen, R. M. Caprioli, and J. V. Staros, "Characterization of Glycosylation Sites of the Epidermal Growth Factor Receptor," *Biochemistry*, vol. 42, no. 18, pp. 5478–5492, May 2003, doi: 10.1021/bi027101p.

[214] R. Nishikawa *et al.*, "A mutant epidermal growth factor receptor common in human glioma confers enhanced tumorigenicity.," *Proc. Natl. Acad. Sci.*, vol. 91, no. 16, pp. 7727–7731, Aug. 1994, doi: 10.1073/pnas.91.16.7727.

[215] M. Vizovišek, R. Vidmar, and M. Fonović, "FPPS: Fast Profiling of Protease Specificity," in *Protein Terminal Profiling*, vol. 1574, O. Schilling, Ed., in Methods in Molecular Biology, vol. 1574. , New York, NY: Springer New York, 2017, pp. 183–195. doi: 10.1007/978-1-4939-6850-3_13.

[216] U. Schmidt, G. Guigas, and M. Weiss, "Cluster formation of transmembrane proteins due to hydrophobic mismatching," *Phys. Rev. Lett.*, vol. 101, no. 12, p. 128104, Sep. 2008, doi: 10.1103/PhysRevLett.101.128104.

[217] C. T. Chu, K. D. Everiss, C. J. Wikstrand, S. K. Batra, H. J. Kung, and D. D. Bigner, "Receptor dimerization is not a factor in the signalling activity of a

transforming variant epidermal growth factor receptor (EGFRvIII)," *Biochem. J.*, vol. 324 (Pt 3), no. Pt 3, pp. 855–861, Jun. 1997, doi: 10.1042/bj3240855.

[218] P. S. Boyd *et al.*, "Clustered localization of EGFRvIII in glioblastoma cells as detected by high precision localization microscopy," *Nanoscale*, vol. 8, no. 48, pp. 20037–20047, 2016, doi: 10.1039/C6NR05880A.

[219] R. K. Kancha, N. von Bubnoff, and J. Duyster, "Asymmetric kinase dimer formation is crucial for the activation of oncogenic EGFRvIII but not for ERBB3 phosphorylation," *Cell Commun. Signal.*, vol. 11, no. 1, p. 39, Jun. 2013, doi: 10.1186/1478-811X-11-39.

[220] K. M. Ferguson, M. B. Berger, J. M. Mendrola, H.-S. Cho, D. J. Leahy, and M. A. Lemmon, "EGF Activates Its Receptor by Removing Interactions that Autoinhibit Ectodomain Dimerization," *Mol. Cell*, vol. 11, no. 2, pp. 507–517, Feb. 2003, doi: 10.1016/S1097-2765(03)00047-9.

[221] I. Lax, F. Bellot, R. Howk, A. Ullrich, D. Givol, and J. Schlessinger, "Functional analysis of the ligand binding site of EGF-receptor utilizing chimeric chicken/human receptor molecules.," *EMBO J.*, vol. 8, no. 2, pp. 421–427, Feb. 1989, doi: 10.1002/j.1460-2075.1989.tb03393.x.

[222] T. Reinheckel *et al.*, "The lysosomal cysteine protease cathepsin L regulates keratinocyte proliferation by control of growth factor recycling," *J. Cell Sci.*, vol. 118, no. 15, pp. 3387–3395, Aug. 2005, doi: 10.1242/jcs.02469.

[223] P. H. Huang *et al.*, "Quantitative analysis of EGFRvIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma," *Proc. Natl. Acad. Sci.*, vol. 104, no. 31, pp. 12867–12872, Jul. 2007, doi: 10.1073/pnas.0705158104.

[224] W. Stec *et al.*, "Cyclic trans-phosphorylation in a homodimer as the predominant mechanism of EGFRvIII action and regulation," *Oncotarget*, vol. 9, no. 9, pp. 8560–8572, Jan. 2018, doi: 10.18632/oncotarget.24058.

[225] S. Li, K. R. Schmitz, P. D. Jeffrey, J. J. W. Wiltzius, P. Kussie, and K. M. Ferguson, "Structural basis for inhibition of the epidermal growth factor receptor by cetuximab," *Cancer Cell*, vol. 7, no. 4, pp. 301–311, Apr. 2005, doi: 10.1016/j.ccr.2005.03.003.

[226] A. Dreier, S. Barth, A. Goswami, and J. Weis, "Cetuximab induces mitochondrial translocalization of EGFRvIII, but not EGFR: involvement of mitochondria in tumor drug resistance?," *Tumor Biol.*, vol. 33, no. 1, pp. 85–94, Feb. 2012, doi: 10.1007/s13277-011-0248-4.

[227] M. Grozdanić, R. Vidmar, M. Vizovišek, and M. Fonović, "Degradomics in Biomarker Discovery," *PROTEOMICS – Clin. Appl.*, vol. 13, no. 6, p. 1800138, 2019, doi: 10.1002/prca.201800138.

[228] D. Ciardiello *et al.*, "Biomarker-Guided Anti-EGFR Rechallenge Therapy in Metastatic Colorectal Cancer," *Cancers*, vol. 13, no. 8, Art. no. 8, Jan. 2021, doi: 10.3390/cancers13081941.

[229] D. A. Nathanson *et al.*, "Targeted Therapy Resistance Mediated by Dynamic Regulation of Extrachromosomal Mutant EGFR DNA," *Science*, vol. 343, no. 6166, pp. 72–76, Jan. 2014, doi: 10.1126/science.1241328.

[230] J. H. Sampson *et al.*, "An epidermal growth factor receptor variant III–targeted vaccine is safe and immunogenic in patients with glioblastoma multiforme," *Mol.*

Cancer Ther., vol. 8, no. 10, pp. 2773–2779, Oct. 2009, doi: 10.1158/1535-7163.MCT-09-0124.

[231] A. Schulte *et al.*, "Erlotinib resistance in EGFR-amplified glioblastoma cells is associated with upregulation of EGFRvIII and PI3Kp110δ," *Neuro-Oncol.*, vol. 15, no. 10, pp. 1289–1301, Oct. 2013, doi: 10.1093/neuonc/not093.

[232] F. Yamasaki *et al.*, "Acquired Resistance to Erlotinib in A-431 Epidermoid Cancer Cells Requires Down-regulation of MMAC1/PTEN and Up-regulation of Phosphorylated Akt," *Cancer Res.*, vol. 67, no. 12, pp. 5779–5788, Jun. 2007, doi: 10.1158/0008-5472.CAN-06-3020.

[233] H. Zhu, X. Cao, F. Ali-Osman, S. Keir, and H.-W. Lo, "EGFR and EGFRvIII interact with PUMA to inhibit mitochondrial translocalization of PUMA and PUMA-mediated apoptosis independent of EGFR kinase activity," *Cancer Lett.*, vol. 294, no. 1, pp. 101–110, Aug. 2010, doi: 10.1016/j.canlet.2010.01.028.

[234] M. Nagane, A. Levitzki, A. Gazit, W. K. Cavenee, and H.-J. S. Huang, "Drug resistance of human glioblastoma cells conferred by a tumor-specific mutant epidermal growth factor receptor through modulation of Bcl-XL and caspase-3-like proteases," *Proc. Natl. Acad. Sci.*, vol. 95, no. 10, pp. 5724–5729, May 1998, doi: 10.1073/pnas.95.10.5724.

[235] X. Jiang, F. Huang, A. Marusyk, and A. Sorkin, "Grb2 Regulates Internalization of EGF Receptors through Clathrin-coated Pits," *Mol. Biol. Cell*, vol. 14, no. 3, pp. 858–870, Mar. 2003, doi: 10.1091/mbc.e02-08-0532.

[236] G. Levkowitz *et al.*, "Ubiquitin Ligase Activity and Tyrosine Phosphorylation Underlie Suppression of Growth Factor Signaling by c-Cbl/Sli-1," *Mol. Cell*, vol. 4, no. 6, pp. 1029–1040, Dec. 1999, doi: 10.1016/S1097-2765(00)80231-2.

[237] A. C. De Angelis Campos, M. A. Rodrigues, C. de Andrade, A. M. de Goes, M. H. Nathanson, and D. A. Gomes, "Epidermal growth factor receptors destined for the nucleus are internalized via a clathrin-dependent pathway," *Biochem. Biophys. Res. Commun.*, vol. 412, no. 2, pp. 341–346, Aug. 2011, doi: 10.1016/j.bbrc.2011.07.100.

[238] W.-C. Huang *et al.*, "Nuclear Translocation of Epidermal Growth Factor Receptor by Akt-dependent Phosphorylation Enhances Breast Cancer-resistant Protein Expression in Gefitinib-resistant Cells *," *J. Biol. Chem.*, vol. 286, no. 23, pp. 20558–20568, Jun. 2011, doi: 10.1074/jbc.M111.240796.

[239] C. Li, M. Iida, E. F. Dunn, A. J. Ghia, and D. L. Wheeler, "Nuclear EGFR contributes to acquired resistance to cetuximab," *Oncogene*, vol. 28, no. 43, Art. no. 43, Oct. 2009, doi: 10.1038/onc.2009.234.

[240] Z. Weihua *et al.*, "Survival of Cancer Cells Is Maintained by EGFR Independent of Its Kinase Activity," *Cancer Cell*, vol. 13, no. 5, pp. 385–393, May 2008, doi: 10.1016/j.ccr.2008.03.015.

[241] J. S. Kruger and K. B. Reddy, "Distinct Mechanisms Mediate the Initial and Sustained Phases of Cell Migration in Epidermal Growth Factor Receptor-Overexpressing Cells1," *Mol. Cancer Res.*, vol. 1, no. 11, pp. 801–809, Sep. 2003.

[242] F. Nie *et al.*, "Involvement of epidermal growth factor receptor overexpression in the promotion of breast cancer brain metastasis," *Cancer*, vol. 118, no. 21, pp. 5198–5209, 2012, doi: 10.1002/cncr.27553.

[243] X. L. Ji and M. He, "Sodium cantharidate targets STAT3 and abrogates EGFR inhibitor resistance in osteosarcoma," *Aging*, vol. 11, no. 15, pp. 5848–5863, Aug. 2019, doi: 10.18632/aging.102193.

[244] N. S. Nagaraj, M. K. Washington, and N. B. Merchant, "Combined Blockade of Src Kinase and Epidermal Growth Factor Receptor with Gemcitabine Overcomes STAT3-Mediated Resistance of Inhibition of Pancreatic Tumor Growth," *Clin. Cancer Res.*, vol. 17, no. 3, pp. 483–493, Feb. 2011, doi: 10.1158/1078-0432.CCR-10-1670.

Bibliography

Publications Related to the Thesis

Journal Article

Grozdanić, M., Sobotič, B., Biasizzo, M., Sever, T., Vidmar, R., Vizovišek, M., Turk, B., Fonović, M. (2023). Cathepsin L-mediated EGFR cleavage affects intracellular signalling pathways in cancer. Biological chemistry. [in press], 15 str. ISSN 1431-6730. DOI: 10.1515/hsz-2023-0213/html. [COBISS.SI-ID 171358723], [JCR, SNIP]

Review Article

Grozdanić, M., Vidmar, R., Vizovišek, M., Fonović, M. (2019). Degradomics in biomarker discovery. Proteomics. Clinical applications. vol .13, iss. 6, 1800138-1-1800138-11, ilustr. ISSN 1862-8346. DOI: 10.1002/prca.201800138. [COBISS.SI-ID 32877607]

Conference Papers

KOLARIČ, Matej, VIDMAR, Robert, SEVER, Tilen, GROZDANIĆ, Marija, IVANOVSKI, Sara, TURK, Boris, FONOVIĆ, Marko. PeptideVisualizer : a novel software solution for PROTOMAP-based determination of protease substrates. In: SEPČIĆ, Kristina (ed.), PAVŠIČ, Miha (ed.). 15th Meeting of the Slovenian Biochemical Society with International Participation : book of abstracts : Portorož, 20-23 September 2023]. Ljubljana: Slovenian Biochemical Society, 2023. str. 83. ISBN 978-961-95941-1-7. https://portoroz2023.sbd.si/upload/docs/SBD2023_Portoroz-Book_of_Abstracts.pdf. [COBISS.SI-ID 166424579]

KOLARIČ, Matej, VIDMAR, Robert, SEVER, Tilen, GROZDANIĆ, Marija, TURK, Boris, FONOVIĆ, Marko. Peptide visualizer : a PROTOMAP-based tool for downstream analysis of MaxQuant data. In: KOZAK, Andreja (ed.), et al. Proteolysis at the interface between health and disease : FEBS ICGEB 2022 Advanced Course : 18-21 September, Bled, Slovenia : book of abstracts. Ljubljana: Jožef Stefan Institute, 2022. str. 52. ISBN 978-961-264-229-7. https://proteolysis2022.febsevents.org/call-for-abstracts, https://www-b1.ijs.si/wp-content/uploads/2022/09/Abstract-Book-Bled2022-Final.pdf. [COBISS.SI-ID 130674691]

GROZDANIĆ, Marija, SOBOTIČ, Barbara, SEVER, Tilen, VIDMAR, Robert, VIZOVIŠEK, Matej, TURK, Boris, FONOVIĆ, Marko. Cathepsin cleavage of EGFR

affects intracellular signaling pathway. In: KOZAK, Andreja (ed.), et al. Proteolysis at the interface between health and disease : FEBS ICGEB 2022 Advanced Course : 18-21 September, Bled, Slovenia : book of abstracts. Ljubljana: Jožef Stefan Institute, 2022. Str. 50. ISBN 978-961-264-229-7. [COBISS.SI-ID 171354627]

GROZDANIĆ, Marija, SOBOTIČ, Barbara, SEVER, Tilen, VIZOVIŠEK, Matej, TURK, Boris, FONOVIĆ, Marko. Ectodomain shedding of epidermal growth factor receptor by cysteine cathepsins. FEBS open bio. 2021, vol. 11, iss. s1, str. 165. ISSN 2211-5463. DOI: 10.1002/2211-5463.13205. [COBISS.SI-ID 99753475]

GROZDANIĆ, Marija, SOBOTIČ, Barbara, TURK, Boris, FONOVIĆ, Marko. Ectodomain shedding of epidermal growth factor receptor by cysteine cathepsins. In: TODOROVIĆ, Tamara (ed.). Book of abstracts. Seventh Conference of the Young Chemists of Serbia, 2nd November 2019, Belgrade, Serbia. Belgrade: Serbian Chemical Society, 2019. str. 11. ISBN 978-86-7132-067-2. [COBISS.SI-ID 33124903]

GROZDANIĆ, Marija, SOBOTIČ, Barbara, VIDMAR, Robert, VIZOVIŠEK, Matej, TURK, Boris, FONOVIĆ, Marko. Ectodomain shedding of epidermal growth factor receptor by cysteine cathepsins : [presented at 44th FEBS Congress: From Molecules to Living Systems, Krakow, Poland, July 6-11, 2019]. FEBS open bio. 2019, vol. 9, suppl. 1, str. 341. ISSN 2211-5463. https://febs.onlinelibrary.wiley.com/doi/epdf/10.1002/2211-5463.12675. [COBISS.SI-ID 32738855]

GROZDANIĆ, Marija, SOBOTIČ, Barbara, VIZOVIŠEK, Matej, TURK, Boris, FONOVIĆ, Marko. Ectodomain shedding of epidermal growth factor receptor by cysteine cathepsins. In: TOPOLE, Martin (ed.), et al. Book of abstracts. 11th Jožef Stefan International Postgraduate School Students' Conference and 13th Young Researchers' Day, 15th and 16th May 2019, Planica, Slovenia. Ljubljana: Mednarodna podiplomska šola Jožefa Stefana: = Jožef Stefan International Postgraduate School: Inštitut Jožef Stefan: = Jožef Stefan Institute, 2019. str. 25, ilustr. [COBISS.SI-ID 33128231]

GROZDANIĆ, Marija, SOBOTIČ, Barbara, VIZOVIŠEK, Matej, TURK, Boris, FONOVIĆ, Marko. Ectodomain shedding of epidermal growth factor receptor by cysteine cathepsins. In: Program & abstract book. XXXVIth Winter School on Proteinases and Their Inhibitors, Tiers am Rosengarten, March 13th - 17th, 2019. [S. l.: s. n., 2019]. str. 59, ilustr. [COBISS.SI-ID 33127975]

GROZDANIĆ, Marija, SOBOTIČ, Barbara, VIZOVIŠEK, Matej, KAVČIČ, Nežka, TURK, Boris, FONOVIĆ, Marko. Ectodomain shedding of epidermal growth factor receptor by cysteine cathepsins. In: TURK, Boris (ed.). Book of abstracts. FEBS Workshop on Proteases, Inhibitors and Biological Control [within] 16. International Symposium on Proteases, Inhibitors and Biological Control, Portorož, Slovenia, 08. 09.-12. 09. 2018. Ljubljana: Jožef Stefan Institute, 2018. str. 67. ISBN 978-961-264-131-3. [COBISS.SI-ID 32097831]

Biography

Marija Grozdanić continued her education in the nanosciences and nanotechnologies programme at the Jožef Stefan International Postgraduate School in Ljubljana after finishing her Bachelor's and Master's degrees at the Faculty of Physical Chemistry, University of Belgrade. Since 2015, she has been working as a young researcher on her Ph.D. thesis at the Department of Biochemistry, Molecular and Structural Biology at the Jožef Stefan Institute under the supervision of Prof. Dr. Marko Fonović.