Marija Grozdanič, Barbara Sobotič, Monika Biasizzo, Tilen Sever, Robert Vidmar, Matej Vizovišek, Boris Turk* and Marko Fonovič*

**Cathepsin L-mediated EGFR cleavage affects intracellular signalling pathways in cancer**

https://doi.org/10.1515/hsz-2023-0213
Received May 12, 2023; accepted October 12, 2023; published online October 30, 2023

**Abstract:** Proteolytic activity in the tumour microenvironment is an important factor in cancer development since it can also affect intracellular signalling pathways via positive feedback loops that result in either increased tumour growth or resistance to anticancer mechanisms. In this study, we demonstrated 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, tyrosine kinase inhibitor (TKI) erlotinib and monoclonal antibody (mAb) 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.

**Keywords:** cancer; cysteine cathepsin; cathepsin L; EGFR; extracellular cleavage; resistance to TKIs

1 Introduction

Epidermal growth factor receptor (EGFR) is one of the most important regulators of cellular signalling, and it is involved in numerous physiological processes, such as cell proliferation, migration, adhesion and survival (Wang 2017). Structurally, EGFR belongs to the large group of membrane receptors known as receptor tyrosine kinases (RTKs). Within this group, the epidermal growth receptor family of RTKs consists of four members: EGFR (ErbB1), ErbB2 (HER2), ErbB3 (HER3), and ErbB4 (HER4) (Wieduwilt and Moasser 2008). EGFR is a type I transmembrane protein that is composed of an extracellular ligand-binding domain, a single transmembrane spanning domain, and a cytosolic tyrosine kinase domain. After binding to its extracellular ligands, such as epidermal growth factor (EGF), EGFR forms a dimer that triggers the trans-autophosphorylation of both cytosolic domains and leads to further activation of the downstream signalling pathways (Bogdan and Klämbt 2001). Alterations in EGFR signalling have been demonstrated to occur in many cancers, such as lung cancer, breast cancer, colorectal cancer and gliomas (Lemmon et al. 2014; Wang 2017). Moreover, two therapeutic strategies have been used to target EGFR in various human cancers, namely, monoclonal antibodies (mAbs) and small molecular weight tyrosine kinase inhibitors (TKIs) (Mitchell et al. 2018). Several Food and Drug Administration (FDA)-approved EGFR inhibitors have shown good results in the treatment of several cancers, such as non-small cell lung cancer (NSCLC), breast cancer, and colorectal cancer. 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.

The autophosphorylation of EGFR can also occur without ligand binding. Ligand-independent or constitutive activation is commonly observed in cancer cells due to EGFR overexpression, which causes constitutive activation and triggers an alternative signalling pathway by activating the transcription factor IRF3 (Chakraborty et al. 2014). Furthermore,
several mutations in the extracellular domain of EGFR cause constitutive activation of the EGFR receptor. Most commonly, these mutations remove a part of the extracellular domain that inhibits ligand-independent dimerization of the receptor (Orellana et al. 2019). A well-known example of such a mutant is the EGFRvIII variant, which is expressed in various human cancers, such as glioblastoma multiforme, breast cancer, and lung cancer; in these cancer cells, the EGFRvIII variant stimulates tumorigenic processes via different signalling pathways. EGFRvIII is a spliced form that lacks exons 2-7 of the EGFR gene. This leads to the deletion of amino acid residues 6-273, which form the dimerization arm and part of the ligand-binding pocket. As a consequence, EGFRvIII exhibits ligand-independent signalling activity and impaired endocytosis, thereby prolonging its presence on the cell surface (reviewed in Gan et al. 2013).

As in the case of genetic deletion, complete or partial removal of the EGFR ectodomain by proteolytic cleavage can also affect the activity and functionality of the receptor. Chen and colleagues showed that the matriptase-prostasin cascade of serine proteases proteolytically cleaves the N-terminal region of EGFR and triggers its constitutive activation (Chen et al. 2008). In addition, we have shown that cysteine cathepsins L and S can shed the extracellular domain of EGFR (Sobotič et al. 2015). Moreover, cathepsin S was also found to regulate EGFR signal transduction by mediating its endosomal degradation (Huang et al. 2016). These results suggest that cysteine cathepsins may be important modulators of EGFR activity. Their role could be particularly important under pathological conditions, such as carcinogenesis; cathepsins are overexpressed and often secreted into the extracellular milieu under such conditions (Kramer et al. 2017). In this work, we explored the downstream effects of cathepsin-mediated EGFR processing on receptor signalling and its possible implications in the molecular mechanism underlying the resistance of cancer to EGFR inhibitors.

2 Results

2.1 Determination of cathepsin cleavage site

Initially, we treated intact MDA-MB-231 cancer cells with recombinant cathepsin L and collected the supernatants, which contained proteins that were released from the cell surface. The proteins in the supernatants were N-terminally labelled with trideuteroaetylactylation, digested and identified by mass spectrometry (Figure 1A). MS analysis of the supernatants showed that EGFR was only present in the cathepsin-treated samples and that all the identified peptides were located in the extracellular region of EGFR, mainly within domains I and II (Figure 1B). Among the identified peptides, we observed two deuteroacteyaletlated N-termini at amino acid residues L25 and G225 that were not generated by trypsin. While cleavage at L25 corresponds to the signal peptide cleavage site, cleavage at Arg224-Gly225 was generated by cathepsin L. Furthermore, an EGFR peptide with a nontryptic C-terminus at H585 was also identified, suggesting another putative cathepsin cleavage site (Figure 1D and E). Immunological detection of the cleaved EGFR ectodomain in the supernatants of cathepsin L-treated cancer cells revealed the formation of two stable EGFR fragments (Figure 1C). The main fragment was observed at 40 kDa, and the cleavage product that was present in lower amounts was observed at 130 kDa. The molecular masses of both cleavage products as observed by Western blotting corresponded to the putative cleavage sites that were identified, considering the determined glycosylation pattern of the EGFR ectodomain (Smith et al. 1996; Zhen et al. 2003).

2.2 Autophosphorylation of the truncated EGFR protein

To further evaluate the mechanistic properties of cathepsin L cleavage and its possible physiological implications, we generated an N-terminally truncated form of EGFR, which corresponded to the cathepsin cleavage product starting at G225 (delta 1-224 EGFR). To determine its autophosphorylation potential, we overexpressed this EGFR variant (t-EGFR) in HeLa cells (Figure 2A). HeLa cells were used since this cell line expresses much lower levels of endogenous EGFR than MDA-MB-231 cells and exhibits no baseline EGFR phosphorylation (Figure 2B). Immunological analysis of several known EGFR phosphorylation sites showed that t-EGFR causes the constitutive phosphorylation of tyrosine residues Y1173, Y1045, Y1197, Y1086 and Y1068 (Figure 2A). We further showed that autophosphorylation of the truncated EGFR protein occurred independently of EGF treatment. As expected, in cells that expressed the empty vector (pcDNA4), endogenous EGFR showed no tyrosine phosphorylation at baseline, while overexpression of the full-length EGFR caused some degree of constitutive phosphorylation. In addition, the phosphorylation of five different phosphorylated tyrosine residues in the endogenous and overexpressed EGFR proteins was increased in an EGF concentration-dependent manner. Based on quantitative analysis of the WB results, truncated EGFR had at least two times higher phosphorylation at the Y1086 residue compared to the overexpressed full-length protein (Figure 2C) at the baseline level.
Figure 1: Cathepsin L cleavage site determination. A, Experimental workflow for the determination of the cathepsin L cleavage site in EGFR. B, Graphical representation of peptides (red and blue lines) located in the EGFR ectodomain segments that were released into the supernatant after cathepsin L treatment. Red lines represent peptides created by cathepsin L cleavage. C, Immunological detection of EGFR ectodomain fragments in the supernatants of MDA-MB-231 cells treated with cathepsin L. Supernatants were collected at 5, 15, 30 and 60 min, and cells treated with inhibited cathepsin L were used.
2.3 Phosphorylation profile of HeLa cells overexpressing truncated EGFR

To determine possible differences in downstream phosphorylation events that occurred in cells that expressed full-length and truncated EGFR, we next used the Human Phospho-Kinase array (R&D System, Minneapolis, MN, USA). The array enabled the identification of 37 kinase phosphorylation sites (Figure 3A). After quantification of the dot blot densities, some substantial differences in the activated signalling pathways were observed. Cells that expressed truncated EGFR had approximately twofold higher phosphorylation of Y1086 than cells that overexpressed full-length EGFR (Figure 3B). In addition to Y1086 of EGFR, constitutive activation of truncated EGFR significantly increased the phosphorylation of 10 phosphosites in the array compared to the empty vector (Figure 3B and C). Expression of truncated EGFR significantly increased the phosphorylation of ERK1/2, MSK1/2, p38α, c-Jun, STAT1, STAT3(S705), RSK1/2/3, Src and p70 S6 kinase. Together with EGFR, the interaction of significant phosphosites was analysed using STRING (Figure 3D). To determine the pathways that were activated by the constitutive activation of truncated EGFR, using an FDR value of less than 0.05, 256 pathways were significantly enriched using Gene Ontology (GO), 108 were significantly enriched using KEGG, and 82 were significantly enriched using the Reactome database (Table S1). The signalling pathways that caught our attention are shown in Figure 3E for GO, Figure 3F for KEGG and Figure 3G for Reactome enrichment. The enriched pathways included the EGFR signalling pathway and signalling pathways, such as the MAPK family signalling pathways and JAK-STAT and PI3K-Akt signalling pathways. However, we also observed differences in pathways that are involved in cancer and diseases, particularly in pathways that are involved resistance to tyrosine kinase inhibitors and the regulation of cell death and apoptosis, which prompted us to investigate the relevant effects.
Figure 3: HeLa cells expressing t-EGFR exhibited altered phosphorylation profiles. A. Human phospho-kinase array antibody blots showing changes in the phosphorylation profile of HeLa cells expressing empty vector as a negative control, empty vector with EGF stimulation, EGFR, or t-EGFR. Compared to the negative control, phosphorylated sites with significantly increased phosphorylation are numbered. B. Quantification of phosphorylation of the phospho-sites with significantly increased phosphorylation (in t-EGFR expressing HeLa cells compared to negative control). All the sites were normalized.
2.4 Autophosphorylation of truncated EGFR causes resistance to cetuximab and decreased sensitivity to erlotinib

The KEGG-enriched pathways also included pathways that could enhance resistance to tyrosine kinase inhibitors. To further investigate these results, we used two EGFR inhibitors with different mechanisms of inhibition, namely, mAb cetuximab and TKI erlotinib. Cetuximab (Erbitux™) is a human-mouse chimeric monoclonal antibody that binds to the EGFR domain III with 10-fold higher affinity than the ligand EGF. The binding of cetuximab to EGFR prevents conformational changes in the receptor and, consequently, dimerization of the receptor (Blick and Scott 2007; Mazzarella et al. 2018; Sabbah et al. 2020). Our results showed that cetuximab blocked the binding of EGF to the endogenous receptor in HeLa cells that expressed the empty vector and in cells that overexpressed EGFR, while cells that expressed truncated EGFR remained resistant to EGF and cetuximab (Figure 4A, with quantitative analysis in Supplementary Figure S1). On the other hand, erlotinib (Tarceva™), which is a first-generation TKI, selectively and reversibly inhibits EGFR activity by competing with ATP for the ATP binding pocket in the EGFR kinase domain. After the binding of erlotinib to EGFR, the cytoplasmic tyrosine residues of EGFR could not be phosphorylated, and intracellular signalling was not initiated. Our experiment showed that, unlike cetuximab, erlotinib binds to and blocks the phosphorylation of truncated EGFR as well as overexpressed full-length EGFR (Figure 4B). However, truncated EGFR showed much lower sensitivity to erlotinib since inhibition of truncated EGFR phosphorylation required a much higher inhibitor concentration than inhibition of full-length EGFR phosphorylation (Figure 4C, with quantitative analysis in Supplementary Figure S1).

2.5 Expression of truncated EGFR causes resistance to apoptosis

Resistance to tyrosine kinase inhibitors (TKIs) and regulation of apoptotic processes were among the enriched pathways identified in the STRING analysis. To validate these results, we investigated the effect of t-EGFR expression on resistance to erlotinib or the apoptosis inducer staurosporine (STS) by using flow cytometry. After treatment, cells expressing empty vector, full-length EGFR or truncated EGFR were stained with Annexin V and PI. The percentage of Annexin V- and PI-negative cells (viable cells) was increased among the cells overexpressing EGFR and t-EGFR. After 48 h of treatment with erlotinib, the viability decreased in a similar trend as untreated cells. Furthermore, overnight treatment with the apoptosis inducer staurosporine decreased the viability of the cells. However, expression of truncated EGFR significantly increased cell viability and resistance to apoptosis compared to expression of EGFR (Figure 5A).

2.6 Nuclear localization of truncated EGFR

Resistance to tyrosine kinase inhibitors (TKIs) and cetuximab was previously associated with the nuclear localization and phosphorylation of EGFR and signal transducer and activator of transcription 3 (STAT3) (Zulkifi et al. 2017). To investigate the nuclear and cytosolic localization of these two proteins, transfected HeLa cells were fractionated using the REAP protocol (Figure 5B with quantitative analysis in Supplementary Figure S2). Endogenous EGFR was present in the cytosolic fraction but not in the nuclear fraction of HeLa cells. However, overexpressed full-length and truncated EGFR were both present in the cytosolic and nuclear fractions. Although the overall levels of EGFR and t-EGFR were similar in the whole cell lysate (w.c.l.), the relative nuclear localization of t-EGFR was higher than that of overexpressed EGFR. Both EGFR and t-EGFR were constitutively phosphorylated in the cytosolic and nuclear fractions, and again, t-EGFR had almost two times higher phosphorylation than EGFR. However, compared to the whole cell lysate, only a small portion of the phosphorylated cellular t-EGFR was localized to the nucleus.

Furthermore, STAT3 was present in all fractions, but a small portion of the cellular STAT3 levels were localized to the nucleus. However, when the phosphorylation of STAT3 at Y705 was analysed, the results showed that almost all phosphorylated STAT3 was localized to the nucleus. HeLa cells overexpressing EGFR or t-EGFR exhibited increased phosphorylation of STAT3 at Y705. Moreover, the Y705 site of STAT3 was more highly phosphorylated in cells expressing
truncated EGFR than in cells overexpressing EGFR. Low levels of cellular STAT3 phosphorylated at S727 were present in the nuclear fraction of HeLa cells expressing EGFR or t-EGFR. In contrast, cells expressing the empty vector showed no nuclear localization of STAT3 phosphorylated at S727. As expected, the loading control GAPDH was detected in the w.c.l. and cytosolic fraction, and histone H3 was detected in the w.c.l. and nuclear fraction.

**3 Discussion**

Cysteine cathepsins are highly upregulated in various types of cancer, such as lung cancer, breast cancer, pancreatic cancer, brain cancer and melanoma (Fonović and Turk 2014; Kramer et al. 2017). Genetic ablation of several cathepsins (cathepsin B, C, L, S and X) in mouse cancer models, including pancreatic islet cancer and mammary gland tumour models, caused a decrease in cancer growth and metastasis, confirming the important role of cathepsins in tumour growth and invasion (Biasizzo et al. 2022; Gocheva et al. 2006; Sevenich et al. 2010; Vidak et al. 2019). Cathepsins are secreted into the tumour extracellular microenvironment by several different cells, including cancer cells, endothelial cells, fibroblasts, and tumour-associated macrophages (TAMs) (Biasizzo et al. 2022; Fonović and Turk 2014; Kramer et al. 2017; Mason and Joyce 2011; Vidak et al. 2019; Vizovišek et al. 2019). After secretion, cysteine cathepsins S and L...
cleave the ectodomains of several membrane-bound proteins, as shown in an in vivo mouse model of pancreatic cancer. The membrane-bound proteins shown to be cleaved by these cathepsins include cell adhesion proteins such as ALCAM and transmembrane receptors such as plexin B2 and EGFR (Sobotič et al. 2015).

In this study, we evaluated the functional relevance of EGFR cleavage by cathepsin L. Upon binding of a ligand, EGFR undergoes a conformational change from a closed to an open conformation. Extracellular domains I, II and III form a C shape, where EGF is located between domains I and III (Ferguson et al. 2003; Lax et al. 1989). Cathepsin

Figure 5: Expression of truncated EGFR caused significant differences in cell viability and resistance to apoptosis. A. Cell viability analysis of HeLa cells expressing empty vector pcDNA4™ myc-His (pcDNA4), pcDNA4 EGFR (EGFR), or pcDNA4 t-EGFR (t-EGFR) after treatment with the TKI erlotinib (ERL) or apoptosis inducer staurosporine (STS) for 48 h. The percentage of Annexin- and PI-negative or viable cells was determined using flow cytometry. *p < 0.05; **p < 0.01; ***p < 0.001, compared to negative control. Error bars show the standard deviations based on triplicate values of each dataset. B. Western blotting analysis of HeLa cells expressing empty vector pcDNA4™ myc-His (P), pcDNA4 EGFR (E), or pcDNA4 t-EGFR (T) after fractionation according to the REAP protocol. Whole cell lysate (w.c.l.), cytosolic fractions and nuclear fractions were analysed to determine EGFR and STAT3 localization and phosphorylation. The levels of EGFR, EGFR phosphorylated at the Y1086 residue, STAT3, and STAT3 phosphorylated at the Y705 and S727 residues were immunologically detected. Antibodies against GAPDH and histone H3 were used as loading controls.
L-mediated cleavage of EGFR results in the removal of the EGF binding region. Cathepsin L cleaves EGFR within the extracellular domain II at the R224-G225 bond, resulting in a decreased binding of EGF and a consequent increase in EGFR autophosphorylation. Our finding is supported by a previous study in which cathepsin L-knockout mouse keratinocytes were more responsive to EGF than the wild-type keratinocytes (Reinheckel et al. 2005). On the other hand, our truncated receptor exhibited constitutive autophosphorylation. The constitutive activity of the EGFRvIII splice variant has been explained by its asymmetric active conformation, which is facilitated by the absence of the N-terminal extracellular region (Guo et al. 2015; Huang et al. 2009; Nagane et al. 1996; Nishikawa et al. 1994; Wong et al. 1992). Therefore, we suggest that also the cleavage of EGFR by extracellular cathepsin L results in the asymmetric active conformation. Moreover, high-precision localization microscopy demonstrated that EGFRvIII creates a dimer (Boyd et al. 2016), suggesting that the truncated EGFR will also form a dimer. Such activation of the truncated EGFR caused alterations in the MAPK family, JAK-STAT, and PI3K-Akt signalling pathways as well as in pathways that are involved in cancer progression, regulation of cell death, and resistance to tyrosine kinase inhibitors. Furthermore, our results showed that overexpression of the full-length EGFR caused receptor autophosphorylation in the absence of ligand stimulation. Such constitutive signalling of overexpressed EGFR has been described as noncanonical since this signalling ‘does not lead to activation of several canonical signalling pathways, such as the ERK and Akt signalling pathways. However, after stimulation with a ligand, overexpressed EGFR did regulate canonical signalling pathways (Chakraborty et al. 2014). This suggested that the constitutively active truncated t-EGFR protein could also regulate different signalling pathways. Since expression of t-EGFR in HeLa cells affected receptor signalling, we explored the potential resistance to EGFR inhibitors, which is a well-known problem in cancer treatment. Our results suggest that the mAb cetuximab blocked the ligand-induced signalling of overexpressed EGFR, while autophosphorylation of truncated EGFR remained the same. This is interesting as Cetuximab binds exclusively to domain III of EGFR, which remains intact after cathepsin L cleavage, thereby suggesting that a conformational change may follow the cleavage. In support of our results, Cetuximab was also shown to bind the EGFRvIII and cause its internalization, although it failed to inhibit EGFR VIII activity (Dreier et al. 2012). However, in the latter case EGFRvIII was localized to mitochondria and not nucleus, in contrast to normal EGFR and our t-EGFR, although the mechanism of mitochondrial translocation is not clear. On the other hand, TKI erlotinib blocks the autophosphorylation of both truncated and overexpressed full-length EGFR. 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 a truncated EGFR, which is more resistant to erlotinib than the overexpressed full-length EGFR. This could be another way to explain the general resistance of cancer cells to erlotinib. Moreover, such cathepsin-mediated truncated form of EGFR could be explored as a biomarker for anti-EGFR therapeutic resistance (Grozdanić et al. 2019).

One of the possible mediators of resistance to erlotinib and cetuximab and apoptosis could be the interaction of EGFR with STAT3 and their nuclear localization. EGFR is known to translocate to the cell nucleus (Brand et al. 2013; Roskoski 2014; Yamaoka et al. 2017), although the exact mechanisms of EGFR nuclear translocation is unclear (Shah et al. 2019). In the nucleus, EGFR interacts with the SH2 domain of STAT3, leading to STAT3 phosphorylation at the Y705 and S727 residues and its activation (Quesnelle et al. 2007). In addition, activated STAT3 can homodimerize or heterodimerize with STAT1 and further translocate to the cell nucleus, where it regulates the transcription of various genes that are involved in cellular proliferation, resistance to apoptosis and angiogenesis. Indeed, we confirmed that the STAT3 Y705 and STAT1 Y701 phospho-sites had significantly increased phosphorylation in HeLa cells that express t-EGFR. There is increasing evidence shows that STAT3 signalling mediates and promotes resistance to EGFR therapeutics, as reviewed in (Zulki If et al. 2017). 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 (Brand et al. 2013; Huang et al. 2011; Li et al. 2009). 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 into the nucleus (Brand et al. 2013; Weihua et al. 2008). 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.

4 Conclusions

In summary, the findings of this study confirm cathepsin L cleavage of the ectodomain of EGFR and reveal the
physiological relevance of this cleavage (Figure 6). Herein, we have demonstrated that truncated EGFR exhibits constitutive activation in the absence of ligand binding. Truncated EGFR activates different intracellular signalling pathways and translocates to the cell nucleus. Moreover, the expression of truncated EGFR and its nuclear translocation possibly leads to STAT3 nuclear localization and activity. Altogether, the expression of truncated EGFR leads to an increase in cancer cell viability and resistance to EGFR therapeutics. Finally, our findings could lead to a better understanding of EGFR signalling and possibly more effective anticancer treatment strategies in which EGFR inhibitors are combined with cathepsin inhibition.

5 Materials and methods

5.1 Cathepsins

Recombinant human cathepsin L was expressed in the Pichia pastoris methylotrophic yeast expression system and purified using the protocol that was previously described in (Mihelić et al. 2008).
5.2 Cell culture

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% foetal bovine serum FBS, 1% penicillin/streptomycin, and 1% glutamine (all from Sigma-Aldrich, Saint Louis, MO, USA). The cell lines were grown in a humidified incubator at 37 °C and 5% CO2. These cell lines were chosen based on their endogenous epidermal growth factor receptor expression. HeLa cells are easy to transfect and have low endogenous expression levels of EGFR, while the MDA-MB-231 cell line expresses high levels of endogenous EGFR. The HeLa cell line was treated with the EGFR ligand EGF (Sigma–Aldrich). Before treatment with EGF ligand, HeLa cells were starved overnight in serum-free medium. After starvation, the cells were incubated with EGF (10 or 100 ng/ml) for 8 min.

5.3 Construction of plasmids and transfection methods

EGFR cDNA was obtained from the Addgene plasmid 11,011 (gift from Matthew Meyerson (Greulich et al. 2005)). The cDNA was amplified by Q5 hot start DNA polymerase (New England Biolabs (NEB) Ipswich, MA, USA) using the following nucleotide primers: TT-CTC-GAG-ATG-CCA-CCC-TCC-GGG (sense primer containing the XhoI restriction site) and AA-CTC-GAG-TGG-TCC-AAT-CCT-CTC-CTC-CTC (antisense primer containing the XhoI restriction site). cDNA was cloned into pcDNA4™myc-HisA (Invitrogen, Waltham, MA, USA). Truncated EGFR cDNA, which encodes a protein that matches the cathepsin cleavage product, was generated by the addition of the EGFR signal peptide to truncated EGFR cDNA, starting with Glu2. 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 (Invitrogen, Waltham, MA, USA) and pcDNA4 EGFR, and pcDNA4 t-EGFR using PolyJet™ in vitro DNA transfection reagent (Sangen Laboratories, Rockville, MD, USA) following the manufacturer’s instructions. The cells were incubated with the transfection reagent for a minimum of 6 h, after which the medium was replaced with fresh complete medium. After an additional 32 h, the transfected cells were washed with Dulbecco’s phosphate-buffered saline (DPBS) (BioWest, Nuaillé, France) and either collected for cell lysis or further treated, depending on the experiment.

5.4 Treatment of cells with recombinant cathepsin L

MDA-MB-231 cells were grown to confluence, after which the cell plates were washed twice with DPBS, and the cells were detached using a Hanks-based enzyme-free dissociation solution (Millipore). 30 million MDA-MB-231 cells per condition were incubated for 1 h at 37 °C 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 (final concentration of 25 μM)-inhibited cathepsin L as a negative control. After incubation, the residual activity of recombinant cathepsin L was blocked by the addition of E64 (final concentration of 20 μM). The cells were centrifuged at 500g for 5 min, and the supernatants were collected and prepared for Western blotting analysis.

5.5 Determination of the cathepsin L cleavage site in the EGFR ectodomain

MDA-MB-231 cells were grown to confluence and detached with a Hanks-based enzyme-free dissociation solution (Millipore). 30 million MDA-MB-231 cells per condition were incubated for 1 h at 37 °C 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 (final concentration of 25 μM)-inhibited cathepsin L as a negative control. After incubation, the residual activity of recombinant cathepsin L was blocked by the addition of E64 (final concentration of 20 μM). The cells were centrifuged at 500g for 5 min, and the supernatants were collected and centrifuged at maximum speed for 5 min. Amicon 0.5 Ultra-centrifugal filters (Millipore) were used to concentrate the supernatants to a final volume of 40 μl. The same amount of supernatant was separated on a 12.5% SDS–PAGE Precast gel (Lonza, Basel, Switzerland).

The separated proteins were fixed and stained using Coomassie Brilliant Blue for 1 h, followed by incubation with a destaining solution (30% ethanol, 10% acetic acid) overnight. The gel was rinsed in water, and sample lanes were cut. Each lane was further cut into a total of six bands, representing six protein samples per condition. The bands were cut into 1 mm gel cubes and further destained by incubation with 150 μl of 50% acetonitrile (ACN) in 25 mM ammonium bicarbonate (ABC) for 30 min at 25 °C and 1200 rpm (2 rounds of incubation). The gel cubes were dehydrated in 100% ACN and vacuum-dried. Free N-termini, which were created by cathepsin L cleavage, were labelled using N-hydroxysuccinimide ester of trideuterio-acetate (AcD3-NHS, made in-house Staes et al. (2011)). AcD3-NHS solution was prepared by dissolving 20 mg of AcD3-NHS in 2 ml of PBS (0.1 M, pH 8.5). The gel cubes were incubated in 150 μl of AcD3-NHS solution for 1 h, 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 10 mM DTT in 25 mM ABC for 1 h at 56 °C, while alkylation of the same residues was performed using 150 μl of 55 mM iodoacetamide (IAM) in 25 mM ABC for 30 min at 30 °C in the dark. The unreacted iodoacetamide was quenched by an additional incubation with 150 μl of 20 mM DTT in 25 mM ABC for 30 min at 37 °C. Then, the gel cubes were 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 approximately 1 μg of sequencing-grade porcine trypsin (Promega, Madison, WI, USA) in 25 mM ABC at 37 °C overnight. The next day, trypsin was inactivated by the addition of formic acid FA (final concentration of 5%), and peptides were extracted from the gel by the addition of extraction solution (150 μl of 50% ACN, 5% FA). The samples were further concentrated to 50 μl by vacuum drying and desalted using a C18 tip. The C18 tip was prepared by loading four Empore/C18 disks (Varian, Palo Alto, CA, USA) into the 200-μl tip and further activated as described by (Vidmar et al. 2017). The samples were applied to the C18 tip and washed using 0.1% FA. In three subsequent steps, peptides were eluted using 300 μl 60% ACN with 0.1% FA. Acetonitrile was removed by vacuum drying the samples to a total volume of 12 μl and sent for LC–MS/MS analysis.
5.6 Analysis by LC–MS/MS

The samples were analysed using an Orbitrap LTQ Velos mass spectrometer (Thermo Fisher Scientific) coupled to an EASY-nanoLC II HPLC unit (Thermo Fisher Scientific). The peptide samples dissolved in solvent A (0.1% FA (v/v) in water) were loaded on a C18 trapping column (Proxeon EASY-Column™, Thermo Fisher Scientific) and separated with a C18 PicoFrit™ AQUASIL analytical column (New Objective, Woburn, MA, USA). The peptides were eluted with a 90 min gradient of 5–50% solvent B (100% CAN in 0.1% FA) at a flow rate of 300 nl/min. Using an Orbitrap mass analyser, the full MS mass spectra were obtained at a resolution of 30,000, with a mass range of 300–2000 m/z. The nine most intense ions from the full MS spectra were fragmented by HCD fragmentation. The MS/MS spectra were recorded at a resolution of 5000. 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 30 s, and exclusion duration of 30 s.

5.7 Data analysis

Tandem mass spectra were analysed using 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 three missed cleavages. The parent and fragment mass error tolerance was 15.0 ppm and 0.02 Da, respectively. Carboxamidomethylation was used as a fixed modification, and for variable modifications, oxidation, acetylation and trideuterio-acetate labelling reagent (heavy, +3 amu) were used. The resulting protein list was filtered by a 1% false discovery rate (FDR).

5.8 Western blotting analysis and antibodies

Cells were washed twice using DPBS and lysed using lysis buffer (20 mM TRIS/HCl, 150 mM NaCl, 1 mM EDTA, 1% (v/v) NP-40, 10% (v/v) glycerol, 1 mM sodium orthovanadate, 10 mM NaF, 10 mM β-glycerophosphate, pH 7.4). The samples were centrifuged at 16,000×g for 15 min at 4°C. The lysates were collected, and their protein concentrations were determined using the Bradford assay. Nuclear and cytoplasmic fractions were generated using the REAP protocol as previously described by (Suzuki et al. 2010). Equal amounts of proteins were boiled at 95°C for 6 min, separated by 10% SDS–PAGE and transferred to nitrocellulose membranes (Cytiva, Amersham Place, UK). The membranes were blocked in a 5% skim milk solution in TBST for 1 h and incubated with primary antibodies overnight at 4°C. The membranes were washed with TBST 3 times and incubated with HRP-conjugated secondary antibodies for 1 h at room temperature. The bands were visualized using enhanced chemiluminescent ECL Western Blotting Detection Reagent (GE Healthcare, Chicago, IL, USA) and the G: BOX Chemi XRQ gel dock system (Syngene). Pixel densities were determined using ImageJ’s extension for the Dot Blot Analyzer. The average signal of paired dots was determined, while the reference 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 by pcDNA4 sample values.

5.9 Human phospho-kinase array

HeLa cells were transfected with the pcDNA4, pcDNA4 EGFR or pcDNA4 t-EGFR plasmid, as previously described. Thirty-two hours after transfection, the complete growth medium was replaced with serum-free medium, and the cells were starved overnight. Transfected HeLa cells were lysed and incubated with phospho-kinase array membranes according to the manufacturer’s instructions (Human Phospho-Kinase Array (R&D System)). Dots were visualized using ECL Western Blotting Detection Reagent (GE Healthcare) and the G: BOX Chemi XRQ gel dock system (Syngene). Pixel densities were determined using ImageJ’s extension for the Dot Blot Analyzer. The average signal of paired dots was determined, while the reference 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 by pcDNA4 sample values.

5.10 Erlotinib and cetuximab treatment

Twenty-four hours after transfection, HeLa cells expressing empty vector, wtEGFR or t-EGFR were starved overnight in serum-free medium. The starved cells were treated with the tyrosine kinase inhibitor (TKI) erlotinib (Selleck Chemicals, Houston, TX, USA) (1 μM) for 1 h. After the treatment, the cells were additionally treated with 100 ng/ml EGF for 8 min. The cells were lysed and prepared for Western blotting analysis. The same experiment was performed using four different concentrations of erlotinib (0, 10, 100 or 1000 nM). For the negative control, cells were treated with serum-free medium containing DMSO. Similarly, after transfection of HeLa cells, the cells were treated with serum-free medium or serum-free medium containing cetuximab (Selleck Chemicals) (50 μg/ml) for 48 h. After the treatment, the cells were additionally treated with 100 ng/ml EGF for 8 min. The cells were lysed and prepared for immunological analysis.

5.11 Cell viability assay

One day after transfection, HeLa cells were seeded at a density of 2 × 10^4 cells per well in a 12-well plate. The cells were grown in complete DMEM or in the same medium containing 10 μM erlotinib for 48 h. Staurosporine-treated samples were treated with 1 μM staurosporine for the last 16 h. Cell suspensions and adherent cells were collected and treated with Annexin V and propidium iodide (PI). Cell viability was investigated by flow cytometry using a previously described protocol (Crowley et al. 2016). The samples were analysed using a FACSCalibur flow cytometer and FlowJo software. Annexin V- and PI-negative cells were considered viable.

5.12 Statistical analysis

GraphPad Prizm (Dotmatics) software was used for visualization and statistical analysis.
Research ethics: Not applicable.

Author contributions: All the authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: The authors state no conflict of interest.

Research funding: The work was supported by grants from the Slovene Research Agency (P1-0140 and J1-1710 to B.T.)

Data availability: The raw data can be obtained on request from the corresponding author.

References


**Supplementary Material:** This article contains supplementary material (https://doi.org/10.1515/hsz-2023-0213).