PhD degree in Molecular Medicine, curriculum in Molecular Oncology,
European School of Molecular Medicine (SEMM),
University of Milan and University of Naples "Federico II"
Faculty of Medicine (BIO/10)

Integration of modeling and experiments to define principles of EGFR activation and ubiquitination

under physiological and pathological conditions

Alexia Conte

IFOM, Milan
Matricola n. R09859

UNIVERSITÀ DEGLI STUDI DI MILANO

Supervisor: Prof. **Pier Paolo Di Fiore**
IFOM, Milan
IEO, Milan
University of Milan

Added supervisor: **Sara Sigismund**
IFOM, Milan

Anno Accademico 2014-2015

This work was supported by:

FONDAZIONE ITALIANA
PER LA RICERCA SUL CANCRO

TABLE of CONTENTS

Figure and tables index

Abbreviations

ADAM	A Disintegrin And Metalloproteinase
AP180	Adaptor protein 180
AP2	Adaptor Protein 2
AREG	Amphiregulin
Arf6	ADP ribosylation factor 6
ATP	Adenosine tri-phosphate
BTC	Betacellulin
c-Cbl	Casitas B-lineage lymphoma
CaMK	Calmodulin-dependent protein kinase
Cas	CRISP associated protein
Cbl-b	Casitas B-lineage lymphoma b
Cbl-c	Casitas B-lineage lymphoma c
CCP	Clathrin-coated pit
CCV	Clathrin-coated vesicle
Cdc42	Cell division control protein 42 homolog
CHC	Clathrin heavy chain
CLC	Clathrin light chain
CLIC	Clathrin-independent carriers
CME	Clathrin mediated endocytosis
CRISPR	Clustered regularly interspaced short palindromic repeats
CtBP1	C-terminal binding protein 1
CTxB	Cholera toxin B
Cys	Cysteine
DAG	Diacylglycerol
DEG	Delayed early genes
DELFIA	Dissociation enhanced lanthanide florescent immunoassay
DSB	Double strand break
DUB	Deubiquitinating enzymes
Dyn	Dynamin
EAM	Early activation model
EE	Early endosome
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EGFR-pY	EGFR phosphorylation
EGFR-Ub	EGFR ubiquitination
ELISA	Enzyme-linked immunosorbent assay
EPGN	Epigen
Eps15	Epidermal growth factor receptor substrate 15
Eps15L1	Epidermal growth factor receptor substrate 15-like 1
Epsin	EPS15-intercating protein
ER	Endoplasmic reticulum
EREG	Epiregulin
ERK	Extracellular signal regulated kinase
ERME	ER-mediated endocytosis

ESCRT	Endosomal sorting complex required for transport
Eu	Europium
EV	Empty vector
FACS	Fluorescence-activated cell sorting
FAK	Focal adhesion kinase
FEME	Fast Endophilin-mediated endocytosis
FGFR	Fibroblast growth factor receptor
GDP	Guanosine 5’ diphosphate
GEEC	GPI-AP enriched early endosomal compartment
GFP	Green fluorescent protein
GPCR	G protein-coupled receptor
GPI	Glycosyl-phosphatidylinositol
GPI-AP	Glycosyl-phosphatidylinositol anchored proteins
Grb2	Growth factor receptor-bound protein 2
gRNA	guide RNA
GTP	Guanosine triphosphate
HB-EGF	Heparin-binding EGF-like growth factor
HDR	Homology-directed repair
HECT	Homologous to the E6-AP Carboxyl Terminus
HER	Human epidermal growth factor receptor
HGF	Hepatocyte growth factor
IB	Immunoblotting
IEG	Immediate early genes
IF	Immunofluorescence
IGF-I	Insulin-like growth factor 1
IGF-IR	Insulin-like growth factor 1 receptor
IgG	Immunoglobulin
IL-2	Interleukin 2
ILV	Intraluminal vesicles
in/del	Insertion and deletions
IP	Immunoprecipitation
IP3	Inositol triphosphate
IR	Insulin receptor
JAK	Janus kinase
JNK	c-Jun NH2-terminal kinase
KCAT	phosphorylation rate constant
KD	Knock-down
kDa	kilo Dalton
KO	Knock-out
K_{PTP}	dephosphorylation constant
LE	Late endosome
lys	Lysine
mAb	Monoclonal antibody
MAPK	Mitogen-activated protein kinase
MEK	Mitogen activated extracellular signal regulated kinase
MET	Mesenchymal epithelial transition factor
MHC	Major histocompatibility complex

MPM	Multisite phosphorylation model
MPM-B	MPM plus binding
mRNA	Messenger RNA
MVB	Multivesicular body
MW	Molecular weight
NA3VO4	Sodium orthovanadate
NCE	Non-clathrin endocytosis
NGF	Nerve growth factor
NHEJ	Non-homologous end joining
NRG	Neuregulin
NSCLC	Non small cell lung cancer
ODE	Ordinary differential equation
PDGF	Platelet-derived growth factor
PDGFR	Platelet-derived growth factor receptor
pEGFR	Phosphorylated EGFR
PI3K	Phosphatidylinositol 3-Kinase
PI3P	Phosphatidylinositol 3-phosphate
PIP2	Phosphatidylinositol 4,5-bisphosphate
PIP3	Phosphatidylinositol 3,4,5-trisphosphate
PKC	Protein kinase C
PLCγ	Phospholipase C γ
PM	Plasma membrane
PTB	Phosphotyrosine binding
PTP	Protein Tyr-specific phosphatases
PTP1B	Protein-Tyr phosphatase 1B
PTRF	Polymerase I and transcript release factor
pY	Phosphotyrosine
qPCR	Quantitative polymerase chain reaction
RING	Really Interesting Gene
RNAi	RNA interference
RSK	Ribosomal S6 kinase
R_T	Total EGFR number per cell
RTK	Receptor tyrosine kinase
RTN3	Reticulon-3
RTPT	Receptor-like PTPs
SD	Standard deviation
SDPR	Serum deprivation response
SDS	Sodium dodecyl sulphate
SH2	Src homology 2 domain
SHC	Src homology 2 domain containing
siRNA	Small interfering RNA
SOS	Son of sevenless
SRG	Secondary response genes
SS	Serum starved
STAT	Signal-transducer and activator of transcription
SV40	Simian virus 40
TfR	Transferrin receptor

TGFα	Transforming growth factor α
TGN	Trans-Golgi network
TKB	Tyrosine kinase binding
TKI	Tyrosine kinase inhibitors
TM	Transmembrane
TRPC1	Transient receptor potential channel 1
Tyr	Tyrosine
Ub	Ubiquitin
UBD	Ubiquitin binding domain
UIM	Ubiquitin interacting motif
WB	Western blot
WT	Wildtype
Y	Tyrosine

Abstract

Epidermal growth factor receptor (EGFR)-dependent signaling is involved in numerous physiological processes, and its deregulation leads to cellular dysfunctions and pathologies, first and forecast, cancer. Endocytosis has a crucial impact on the downstream EGFR signaling response and it is regulated by ligand concentration. Indeed, depending on the EGF dose, the EGFR can be internalized through clathrin-mediated endocytosis (CME) or non-clathrin endocytosis (NCE). The switch between these two internalization mechanisms occurs over a narrow range of EGF concentrations (1-10 ng/ml). Importantly, EGFR ubiquitination shows a threshold response over the same range of EGF doses and is responsible for the commitment of EGFR to NCE, and thus, for EGFR signal extinction through receptor degradation.

In this project, we were interested in elucidating the cellular mechanisms that regulate and coordinate the choice between these two endocytic routes, in addition, we aim to clarify how the integration of the two pathways influences EGFR downstream signaling. In order to deal with the complexity of the system, we adopted an integrated research approach combining mathematical modeling with wet-lab experiments. To this purpose, in collaboration with the Systems Biology group at our Institute, we developed a mathematical model of early EGFR activation that quantitatively accounts for the ubiquitination threshold observed at 2 minutes of EGF stimulation. The 'early model' was able to generate important predictions; in particular, it predicts a weakness in the system that is unveiled in the presence of high EGF concentrations and EGFR overexpression, two conditions frequently observed in cancer.

We tested these predictions using different cell-based model systems subjected to varying perturbations. A challenge in the biological validation of the model, was obtaining quantitative reproducible data. To this aim, we optimized a quantitative ELISA-based assay to

measure EGFR ubiquitination/phosphorylation upon different perturbations. This assay revealed to be powerful and allowed us to validate the predictions generated by the model. Thanks to our integrative approach, we identified Cbl as the limiting and weak element of the system.

We expect that our model of EGFR activation will provide novel insights into the role of EGFR endocytosis, controlling the balance between EGFR signaling and downmodulation, frequently altered in cancer.

1 - Receptor Tyrosine Kinases

Receptor Tyrosine Kinases (RTKs) are a class of proteins important for the control of cell growth and proliferation, differentiation, metabolism and migration. These processes stand at the base of organisms' development and homeostasis, but are also involved in cancer development when deregulated [for complete reviews, see (Lemmon and Schlessinger 2010; Maruyama 2014)].

RTKs are a large family of transmembrane receptors with intrinsic tyrosine kinase activity (Lemmon and Schlessinger 2010). This family comprises a large number of surface receptors, such as the ErbB receptor family, the insulin receptor (IR) family, the platelet-derived growth factor (PDGF) receptor (PDGFR) family, the fibroblast growth factor receptor (FGFR) family and many others (**Fig. 1**; Lemmon and Schlessinger 2010). These receptors are usually activated upon the binding of an extracellular ligand, which induces the auto-phosphorylation of the intracellular part of the receptor, thus allowing for downstream signaling events. RTKs show a high affinity for many specific polypeptides, including growth factors, hormones and cytokines.

Each RTK recognizes its specific ligands via diverse extracellular ligand-binding domains, but all RTKs present a highly conserved intracellular domain, which harbors the kinase activity. Among the vast family of cellular tyrosine kinases, 58 out of the 90 tyrosine kinase genes encode for an RTK. Different subfamilies of RTKs are categorized according to their amino acid sequence identities and extracellular structural similarities; in humans, 20 distinct subfamilies of RTKs have been identified (**Fig. 1**; Lemmon and Schlessinger 2010).

The overall structure of an RTK consists of four domains: an N-terminal ligand-binding domain outside the cell, a single membrane-spanning domain, a tyrosine kinase domain inside the cell and a regulatory C-terminal region (Lemmon and Schlessinger 2010). The protruding

extracellular N-terminal domain, termed ectodomain, is specific for each RTK and confers specificity towards the ligand: it contains a ligand-binding site that can interact with extracellular molecules. The N-terminal region can exhibit a variety of conserved elements that confer ligand specificity. For example, cysteine-rich repeats are typical of the ErbB family of RTKs whose prototype is the epidermal growth factor (EGF) receptor (EGFR). Immunoglobulin-like domains are found in the PDGFR family receptors of class III. Moreover, other domains such as leucine-rich repeats, EGF-like domains and fibronectin type III repeats, are found in different combinations in the extracellular regions of other RTKs.

Each receptor shows a single hydrophobic transmembrane (TM) domain that consists of 25-40 amino acids. The TM domain is then followed by intracellular C-terminal region of the protein that displays the highest level of conservation among RTKs and comprises the catalytic domain responsible for both the kinase activity of the receptor and for the phosphorylation of downstream substrates. The C-terminal portion comprises also a regulatory region containing autophosphorylation residues, which serve as docking sites for

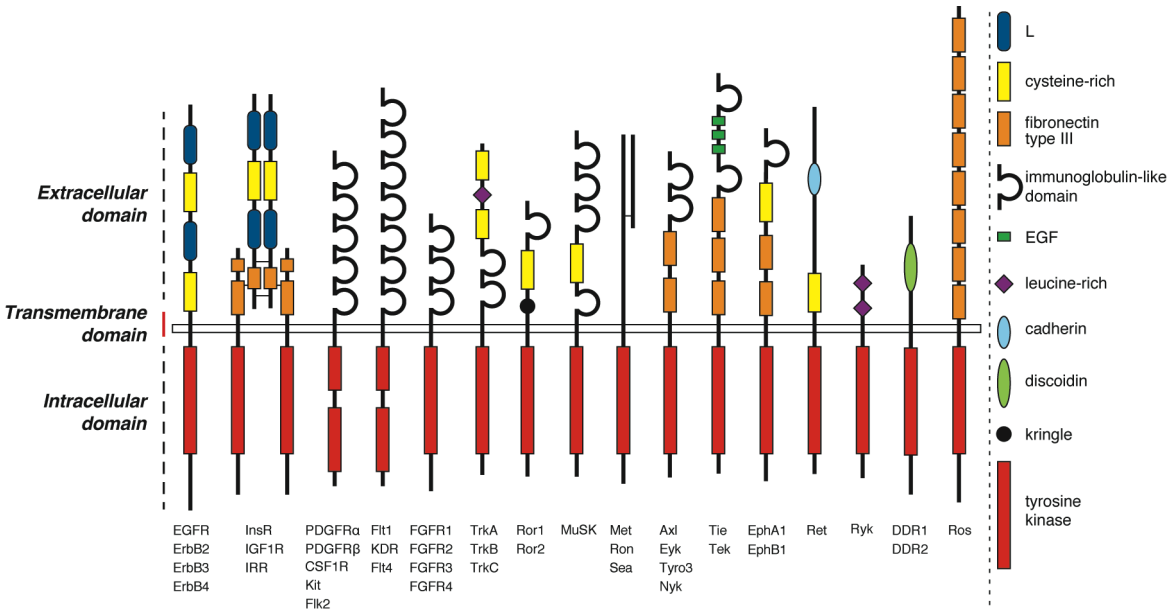


Figure 1: Schematic representation of receptors belonging to the receptor tyrosine kinase family, with their domain organization. All RTKs consist of an extracellular portion, a transmembrane (TM) domain and an intracellular cytoplasmic portion that harbors the tyrosine kinase (TK) domain (red rectangle). The extracellular domain is responsible for ligand specificity and receptor dimerization, and thus determines the cellular response to extracellular signals that are translated inside the cell through the TM domain. The lengths of the receptors as shown are only approximately to scale (adapted from Hubbard and Till 2000).

signaling molecules, endocytic adaptors and the ubiquitination machinery (Lemmon and Schlessinger 2010).

1.1 - RTK activation and signal transduction

RTKs are activated upon binding of growth factors by a mechanism of receptor dimerization (Ullrich and Schlessinger 1990; Lemmon and Schlessinger 2010). Typically, inactive RTKs are found at the PM as monomers and only the binding of the ligand can induce receptor dimerization. More recently, it was also found that a subset of RTKs form oligomers even in the absence of the ligand. This is the case of the IR and the insulin-like growth factor 1 receptor (IGF-1R), which are both expressed on the PM as disulphide-bond dimers; binding of insulin or IGF-1, respectively, triggers structural rearrangements within the dimer that activate the tyrosine kinase domain (Tatulian 2015). Moreover, there is evidence indicating that EGF can bind to and activate pre-existing EGFR dimers (Moriki, Maruyama et al. 2001; Tao and Maruyama 2008) or even tetramers (Clayton, Walker et al. 2005), while the activation of other RTKs, such as Tie2 and Eph receptor, relies on the formation of even larger oligomers (Himanen and Nikolov 2003; Barton, Tzvetkova-Robev et al. 2006). Regardless of whether the inactive state of an RTK is monomeric or oligomeric, binding of the ligand is required to stabilize the dimeric/oligomeric form (Schlessinger 2000). This event is necessary for intracellular structural rearrangements of the kinase domain that activate the kinase activity of the receptor, permitting the trans-phosphorylation of one or more tyrosine residues in the C-terminal region of the neighboring RTK, leading to a pattern of phoshotyrosines (pY) specific for each receptor.

Most RTKs present an activation loop in the cytoplasmic region that is phosphorylated upon ligand binding. This event is necessary for the activation of the tyrosine kinase (TK) domain and the phosphorylation of tyrosine residues, and this is the case - for example - of the IR and the FGFR (Hubbard and Miller 2007). One exception to this mechanism is

represented by the ErbB family of receptors, which includes the EGFR and ErbB2 receptor. The activation of these RTKS does not require the phosphorylation of an activation loop, but rather necessitates asymmetric rearrangements of the intracytoplasmic domain, lending this class of receptor the name "not-so-prototypical" RTK (Lemmon, Schlessinger et al. 2014).

The first response to autophosphorylation of an RTK is the recruitment and subsequent activation of signal-transducing molecules. These molecules contain Src homology 2 (SH2) and phosphotyrosine binding (PTB) domain-containing proteins (Zwick, Bange et al. 2001) that can bind the pYs of the RTK (Schlessinger and Lemmon 2003) and elicit the activation of signaling cascades. Signaling molecules can be directly recruited to the pYs of the activated receptor, or can also be indirectly recruited to the receptor through the action of docking proteins that recognize the RTK-pY, and which are then phosphorylated by the RTK with which they associate with (Schlessinger 2000). With the presence of multiple tyrosine residues in the intracytoplasmic domain of many RTKs, and the involvement of many docking proteins, activated RTKs can clearly recruit and affect a large number of signaling effectors. Thus, stimulated RTKs can be considered as a central node from which complex signaling events originate, and which transmit information from the exterior compartments of the cell to the interior. Initially thought to be linear chains of events (Noselli and Perrimon 2000), RTK-dependent signaling pathways are now recognized as an authentic network of reactions that act through positive and negative feedback mechanisms (Lemmon and Schlessinger 2010; Volinsky and Kholodenko 2013), with the common aim to regulate cell proliferation and survival.

Signaling pathways based on RTKs are multiple and interconnected, and an essential effector cascade required for most RTK function is the mitogen-activated protein kinase (MAPK) cascade that comprises also the extracellular signal-regulated kinase (ERK) kinases. This pathway relies on the activation of the small G protein Ras, and is crucial for RTK-

induced proliferation due to the activation of nuclear transcription factors (McKay and Morrison 2007).

As will be discussed later, docking proteins not only facilitate, or even permit, the interaction of the RTK with the signaling components of the cell, but also link the phosphorylated RTK to other protein modifying enzymes, such as the E3 ligases responsible for RTK ubiquitination. This is the case of the E3 ligase Casitas B-lineage Lymphoma (c-Cbl or Cbl) that is recruited directly to the phosphorylated EGFR via its tyrosine kinase binding (TKB) domain; only upon interaction with the docking protein Growth factor Receptor-Bound protein 2 (Grb2) is Cbl stably recruited to the EGFR and the receptor fully ubiquitinated (see Chapter 3).

2 - The EGFR

The EGFR is a member of the RTK subfamily of the ErbB receptors (Holbro and Hynes 2004), which has critical roles in physiological and pathological processes in epithelial cells (Citri and Yarden 2006). EGFR is the leading member of this subfamily, which includes four closely related proteins: alongside EGFR (also called ErbB1) the family also includes ErbB2 (or HER2/neu), ErbB3 and ErbB4, which are normally expressed in epithelial cells (Citri and Yarden 2006).

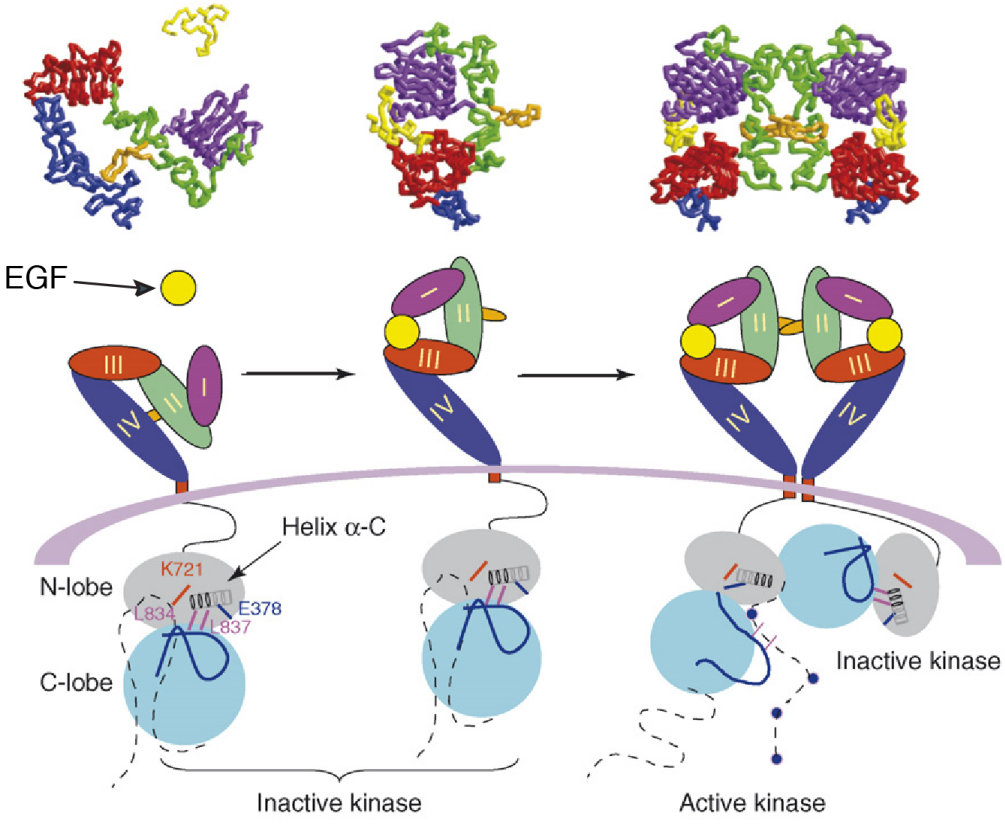
The EGFR can form functional homodimers upon the binding of multiple ligands, as well as three functional heterodimers with the other ErbB family members. Indeed, the EGFR is known to preferentially heterodimerize with ErbB2 (Li, Macdonald-Obermann et al. 2012), and this combination plays a critical role in cancer (Hynes and Lane 2005). Moreover, heterodimerization regulates the activation of specific intracellular signaling pathways and contributes to signaling diversification [for a review on ErbB family members see (Sorkin and von Zastrow 2009)].

The gene that encodes the EGFR is located on the short arm of chromosome 7 in the

Figure 2 (next page): Mechanism of activation of the EGFR upon ligand binding. Schematic representation of the different stages of EGFR activation. The plasma membrane (PM) is represented by the purple arc**.** Above the PM two different representations of the ectodomain are shown. The upper shows a ribbon structure determined by X-ray crystallography of the soluble extracellular domain, the lower shows a schematic representation of those structures. The different subdomains within the ectodomain are indicated with different colors: domain I is purple, domain II green, domain III red, and domain IV dark blue. The ligand (EGF) is yellow. The dimerization arm, a part of domain II, is in orange. The receptor is proposed to exist at the PM as a monomer prior to ligand binding (left). EGF binding to domains I and III in the tethered closed conformation induces a conformational change that exposes the 'dimerization arm' (orange) of domain II, determining the open configuration (center). This dimerization arm interacts with the counterpart on the other receptor moiety to form a dimer (right) in which two intracellular kinase domains assume the asymmetric active structure.
Below the PM diagrams of the intracellular domain are shown: the N-lobe of the kinase domain is colored grey and the C-lobe light-blue. The activation loop is blue, and the C-terminal tail is represented by a dashed line. Left and middle: the inactive conformation of the kinase is depicted. Right: ectodomain dimerization juxtaposes the kinase domains, with the C-lobe of one kinase attached to the N-lobe of the other. The activated kinase phosphorylates tyrosine residues (blue circles) in the C-terminal tail of its dimerization partner. In this scheme, only one receptor is activated; however, in reality, the flexibility of the hinge region allows the two receptors to switch positions and thus, each may activate the other. (Adapted from Bublil and Yarden 2007)

human genome; it is 186 kb long and comprises 28 exons. The mature EGFR is composed of 1186 residues, and it is synthesized starting from a precursor protein of 1210 residues, which, after cleavage of the N-terminal sequence, is inserted into the cell membrane (Ullrich, Coussens et al. 1984).

EGFR can be divided into three main topological regions: an extracellular globular region, a TM domain that spans over the PM and an intracellular portion that harbors the kinetic activity. The EGFR ectodomain is composed of 621 amino acids and contains four distinct subdomains (I-IV) (**Fig. 2**). Domains I and III are leucine-rich regions responsible for ligand binding, with both domains simultaneously contacting the same ligand, whereas domains II and IV are cysteine-rich regions that do not contact the ligand (Burgess, Cho et al. 2003; Lemmon 2009). The 23 amino acid TM domain consists of a patch of hydrophobic residues that extends from the outer part of the PM to the inside, while the C-terminal portion of the protein consists of 542 residues extending into the cytoplasm. This last cytoplasmic portion



contains the catalytic tyrosine kinase domain, as well as the tyrosine residues that can be phosphorylated upon external stimulus.

At the molecular level, EGFR activation and the subsequent signaling cascade can be divided into three essential steps: (1) ligand binding and receptor dimerization, (2) trans-phosphorylation of the intracytoplasmic tyrosine residues and (3) communication with downstream effectors.

2.1 - Ligand-induced EGFR dimerization and activation

The EGFR exists on the cell surface as a monomer, where it waits for activation by a specific ligand that can bind to its ectodomain and trigger the dimerization with another monomer (**Fig. 2**). A series of EGFR crystal structures are available in literature, which have helped to decipher the mechanism of EGFR activation (Garrett, McKern et al. 2002; Ogiso, Ishitani et al. 2002). It is now well accepted that binding of EGF (or other agonists) to EGFR shifts the monomer-dimer equilibrium to favor the dimeric state (Schlessinger 2000; Lemmon and Schlessinger 2010).

Based on crystal structures of the soluble extracellular domain (Ogiso, Ishitani et al. 2002) the four domains can fluctuate between a tethered and *closed conformation*, and an *open conformation* (**Fig. 2** upper part). In the closed conformation, domains II and IV are close together and a tethering arm in domain IV is exposed (Burgess, Cho et al. 2003), while the dimerization arm that resides in domain II is occluded. In the absence of the ligand, the closed conformation is prevalent and is stabilized by disulfide bonds between the dimerization and the tethering arms (**Fig. 2** left), preventing dimerization of the receptor with another monomer. EGF binding to the pocket between domains I and III, at a site remote from the dimer interface, induces a conformational change through a ~130° rotation of domains I and II, until the *open* or *extended conformation* is reached (**Fig. 2** center). This conformational change exposes the dimerization arm in domain II that can subsequently

interact with a counterpart dimerization arm in another ligand-bound and open monomer, thereby driving the dimerization of the receptor (**Fig. 2** right) (Lemmon 2009). This creates an active-state dimer that is able to trans-autophosphorylate the intracytoplasmic region and initiate a signaling events.

There is also evidence indicating that pre-formed EGFR dimers are able to bind EGF. In a study by Chung *et al.* (Chung, Akita et al. 2010) that used quantum dot-based optical tracking of single molecules, it was observed that before ligand addition, EGFRs can spontaneously form finite-lifetime dimers, kinetically stabilized by the dimerization arms. The dimers were primed both for ligand binding and for signaling, such that after EGF addition the dimers displayed a rapid activation. Although the kinetic stability of EGF unloaded dimers was in principle sufficient for EGF-independent activation, ligand binding was still required for signaling (Chung, Akita et al. 2010).

Once the receptor has been activated by the ligand, an asymmetric conformational change takes place in the cytoplasmic domain of the receptor, allowing trans-phosphorylation reactions that result in phosphorylation of specific tyrosine residues. The intracytoplasmic kinase domain can be divided into an N-lobe and a C-lobe, and structural studies have revealed that autophosphorylation takes place in an asymmetric kinase domain dimer (**Fig. 2** lower part), in which the C-lobe of one kinase domain activates the second kinase domain by binding to the N-lobe of the latter (Zhang, Gureasko et al. 2006).

Interestingly, we can highlight two features that differentiate EGFR (and ErbB family receptors) from the rest of the RTK family: (1) the activating ligand does not interact with the dimerization surface, differently to what has been observed for the other RTKs; (2) activation of the tyrosine kinase domain relies on an asymmetric conformational change in the TK domain, rather than on phosphorylation of an activation-loop. These characteristics make EGFR, and the other ErbB family members, non-prototypical RTKs.

2.2 - The EGFR ligands and their role in physiology

The EGFR can bind to 7 ligands that exert different effects on the receptor and that are available under different circumstances (Harris, Chung et al. 2003). EGF is the best-characterized EGFR ligand, but alongside it, the transforming growth factor-α (TGFα), amphiregulin (AREG), epiregulin (EREG), betacellulin (BTC), heparin-binding EGF-like growth factor (HB-EGF), and epigen (EPI) can all act as EGFR ligands (**Fig. 3**).

EGFR ligands are synthesized as membrane-anchored precursors that are later processed by metalloproteases to generate soluble ligands. Metalloproteases of the ADAM (a disintegrin and metalloproteinase) family are thought to be responsible for the shedding of certain EGFR ligands (Sahin, Weskamp et al. 2004). Under certain circumstances, the membrane-anchored isoforms, as well as the soluble growth factors, may also act as biologically active ligands. It is for this reason that these ligands have been observed to act in a justacrine, autocrine, paracrine and/or endocrine fashion (Singh and Harris 2005).

Justacrine signaling is a type of cell/cell signaling that is based upon the interaction of a non-cleaved ligand precursor and the EGFR. This kind of cell/cell communication produces spatially non-uniform patterns in the expression of genes that guide the development of tissues and organs. The first ligand to be discovered to activate EGFR in a justacrine fashion was the TGFα (Anklesaria, Teixido et al. 1990), followed by HB-EGF (Raab, Kover et al. 1996) and AREG (Inui, Higashiyama et al. 1997).

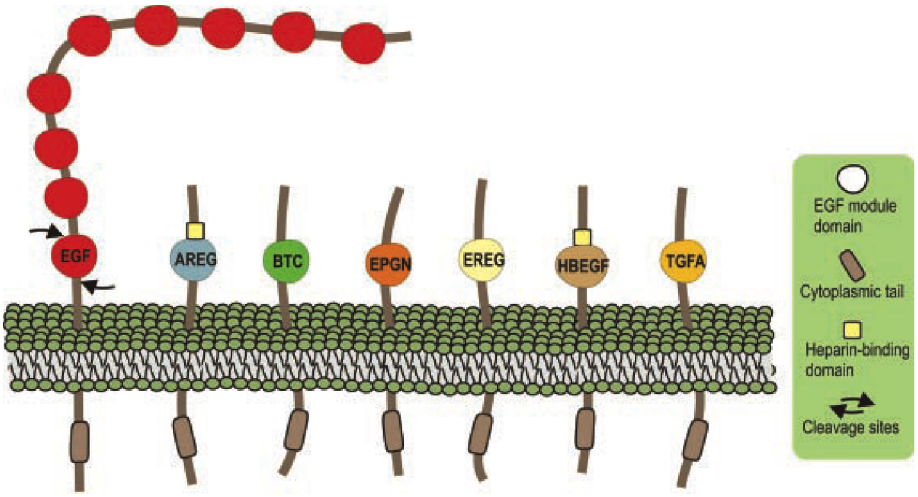


Figure 3: The EGFR ligands. Structural features and schematic representation of the membrane-anchored precursor forms of the seven EGFR ligands. All the EGFR ligands are synthesized as precursor transmembrane proteins that comprise an N-terminal extension, the specific ligand module, a hydrophobic transmembrane domain and a carboxy-terminal fragment, also known as the cytoplasmic tail. Through a proteolytic cleavage, a ‘‘soluble’’ growth factor is released into the extracellular milieu. Overall protein sequence identity between EGFR ligands is low (about 25%) and these proteins also differ concerning the distribution of glycosylation sites, the presence of a heparin-binding domain, and other biochemical characteristics (Harris et al., 2003). EGF is the only ligand harboring a total of nine EGF motifs, although only the one adjacent to the cell membrane is functional as an EGFR-binding domain after two proteolytic cleavages. Abbreviations: Epidermal Growth Factor (EGF), Transforming Growth Factor α (TGFα), Heparin-Binding EGF-like Growth Factor (HBEGF), AmphiREGulin (AREG), betacellulin (BTC), EpiREGulin (EREG) and EPpiGeN (EPGN). (Adapted from Schneider and Wolf 2009)

The other modes of receptor activation (autocrine, paracrine, endocrine) are also important in development and are necessary for different cellular functions. These modes of activation all involve a ligand that has been released from the cell that generated it, and which acts in the fluid phase. Similarly to justacrine signaling, autocrine signaling also exhibits space-restricted dynamics, with the ligand activating receptors on the cell that produced it. In contrast, longer distances can be covered by paracrine signaling, where the ligand is released into the tissue and can interact with the extracellular matrix and cell surface receptors as it spreads through the tissue. Paracrine stimulation can exert effects on the same cell type or on a different cell type. Finally, when the ligand is released systemically, it is known as endocrine signaling. This type of signaling is a frequent mechanism in hormone-dependent stimulations (Singh and Harris 2005). An example of such a hormone-like mechanism is given by the role of EGF

in newborn mammals. In the developing animal, pro-EGF mRNA, EGF mature ligand, as well as TGFα and EGFR, are present in many tissues. EGF is also produced and secreted by the maternal mammary gland, and mammary-derived EGF appears to be important in gut development in the neonatal rodent (Brown, Lam et al. 1990; Mroczkowski and Reich 1993).

EGF and TGFα are the best-studied EGFR ligands that specifically bind to and activate the EGFR. A parallel field of study has also highlighted an important role of HB-EGF in EGFR activation. It was observed that EGFR transactivation, mediated by G-protein coupled receptors, requires metalloprotease-dependent cleavage of pro-HB-EGF, suggesting an autocrine or paracrine mode of EGFR activation by this ligand (Prenzel, Zwick et al. 2000). However, HB-EGF can also activate ErbB4, another member of the ErbB family, making HB-EGF a non-specific EGFR ligand.

Earlier, we mentioned the importance of EGF in the development of newborn mammals, but EGF, together with TGFα and HB-EGF, have an important role in the development of the nervous system (Xian and Zhou 1999). An interesting question arises from the fact that although EGF has been detected in the majority of body fluids of several mammalians, neither EGF antibody administration to newborn animals, nor passive immunization of pregnant rodents against EGF, caused major detrimental effects (Brown, Lam et al. 1990). Moreover, to date, no pathological EGF deficiency disorder has been characterized (Singh and Harris 2005). These observations can be explained by the possible redundancy between the EGFR ligands in a defined body context. Further evidence suggests that TGFα and HB-EGF could be alternative functional growth factors in fetal development.

2.2.1 - Different ligands correlate with distinct biological outcomes

Various studies have observed how EGFR activation by different receptor ligands is translated into distinct biological activities (as schematized later, in **Fig. 4**) (Yarden and Sliwkowski 2001; Harris, Chung et al. 2003). However, the mechanism that underlies these

diverse modes of action is unknown. In biochemical terms, the ligand dynamics involved in activating the receptor appears to be the same for all ligands, encompassing ligand binding, EGFR dimerization and intracytoplasmic trans-phosphorylation. Thus, how can these ligands, which interact in similar way with the receptor, induce distinct biological effects? As mentioned above, EGFR is known to form homo- and heterodimers upon ligand binding, and this has been postulated to affect the binding affinity of the ligand to the EGFR. Therefore, differential dimerization of receptor moieties could be a possible system to induce ligand-dependent signaling specificity, and thus different outputs. Another possible mechanism for discriminating between signaling induced by the different EGFR ligands could involve the strength of the ligand–receptor interaction, which could determine differential trafficking of the EGFR after ligand-dependent internalization. Indeed, as will be discussed in Section 5, following ligand binding, EGFR undergoes internalization through different endocytic pathways, in which the EGFR is sorted into early, recycling and/or late endosomes. For all EGFR ligands, it has been reported that following activation, the receptor is internalized and trafficked to early endosomes (Roepstorff, Grandal et al. 2009). However, the fate of receptors after reaching the early endosomes appears to vary depending on the stimulating ligand. It has been shown that the different ligands vary in their potential to stimulate EGFR degradation or recycling. For example, EPI and TGFα were observed to stimulate endocytosis of EGFR followed by complete recycling to the cell surface, whereas EGF, BTC and HB-EGF stimulate EGFR degradation (Roepstorff, Grandal et al. 2009).

Consequently, some studies have proposed that specific EGFR signaling pathways are triggered within different types of endosomes. In contrast to the requirement of kinase activity for the recruitment into internalization structures (Dikic 2003), regulated endosomal sorting appears to be independent of receptor kinase activity and the strength of the binding between EGFR and its ligands has been proposed to be the mechanism responsible for signaling diversification (Singh and Harris 2005).

In the present work, we followed EGFR activation, endocytosis and signaling in the context of EGF-dependent stimulation. In the next paragraph, the EGF ligand will be described in more detail, and in the following sections, the sequential molecular steps that lead to EGF-dependent EGFR activation and endocytosis will be detailed.

2.2.2 - EGF concentration as a variable element in the organism

EGF is a 6 kDa protein and in humans it is composed of 53 amino acid residues. EGF is a growth factor whose concentration is regulated locally and not systemically like hormones. Indeed, under physiological conditions, EGFR-expressing cells seem to be exposed to a wide range of EGF (and EGF-like ligands) concentrations. There has been a huge debate in the field regarding the establishment of accepted physiological concentrations of EGF, since historically the erroneous perception has been that only low doses of EGF (below 1 ng/ml) are physiological.

The notion that EGF might act locally, rather than systemically as an endocrine factor, is substantiated by findings that various organs seem to independently regulate their levels of EGF. This notion is further supported by the vast differences in EGF concentrations in various bodily fluids (**Table I**): from low concentrations (1-5 ng/ml) in plasma, serum, and saliva (with the exception of mice that have high EGF in saliva), to medium concentrations (5–50 ng/ml) in tears, follicular fluid, sperm, and seminal plasma, to high concentrations (50–500 ng/ml) in bile, urine, milk, and prostate fluid. Bodily EGF might further be derived from exocrine (salivary gland, especially in mice), local/endocrine, and alimentary (milk, where EGF concentrations are very high) sources (reviewed in Carpenter and Cohen 1990). In addition, EGF is produced as a transmembrane precursor, which is not obligatorily processed, but might still act (especially in the kidney) as a juxtacrine stimulator (also reviewed in Carpenter and Cohen 1990).

	Tissue/bodily fluids	Concentration
LOW	Plasma	~1 ng/ml
	Serum	~5 ng/ml
	Saliva	1-3 ng/ml
MEDIUM	Tears	10-30 ng/ml
	Follicular fluid	3-30 ng/ml
	Sperm	20-40 ng/ml
	Seminal plasma	~50 ng/ml
HIGH	Bile	~150 ng/ml
	Urine	~100 ng/ml
	Milk	~400 ng/ml
	Prostate fluid	150 ng/ml

Table I: Concentration of EGF in human tissues and bodily fluids. The concentration of EGF is locally regulated and varies in different tissues and bodily fluids.

In physiological conditions, epithelia prevent these EGF-containing fluids from reaching EGFRs that are expressed basolaterally, but when the tight junctions of the epithelia become leaky, e.g. as a result of a lesion or in premalignant neoplasia, elevated doses of EGF will reach and activate EGFR present on the cell membranes (Mullin 2004).

Furthermore, elevated levels of EGFR ligands have been found in different cancer types (Thogersen, Sorensen et al. 2001; Revillion, Lhotellier et al. 2008; reviewed in Normanno, Bianco et al. 2005). Although an exact ligand concentration has not been measured in tumors, due to the autocrine and paracrine nature of the system, local EGFR ligand concentrations in the tumor microenvironment may reach very high levels, such as in breast and bone marrow cancers (Normanno, Campiglio et al. 2008; Revillion, Lhotellier et al. 2008; Roepstorff, Grandal et al. 2009).

2.3 - Other ErbB family members

Class I RTKs are frequently co-expressed in various combinations and, depending on the activating ligand, they can form various homodimers or heterodimers, generating a complex signal transduction network, as schematized in **Fig. 4** (Alroy and Yarden 1997; Riese and Stern 1998). ErbB receptors are activated by a large group of EGF-related polypeptides, which all contain a conserved EGF-like motif. ErbB receptor ligands include the 7 EGFR ligands that we have described previously (see Introduction Section 2.4), plus 4 members of the neuregulin family (NRG1-4), for a total of 11 ErbB ligands (Garrett, McKern et al. 2002).

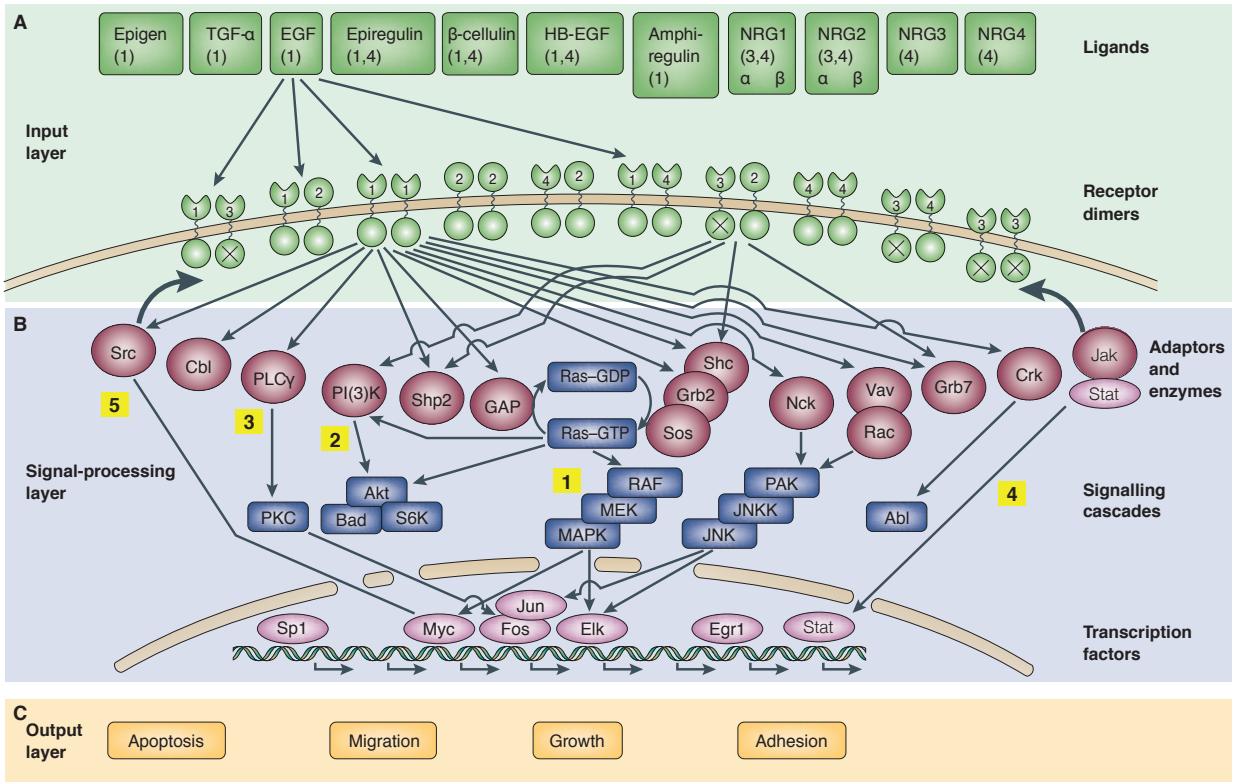


Figure 4: The ErbB signaling network. A - ErbB receptor ligands and the ten dimeric receptor combinations constitute the input layer. Numbers under each ligand indicate the respective high-affinity ErbB receptors. Receptors are indicated as follows: 1=EGFR, 2=ErbB2, 3=ErbB3, 4=ErbB4. For simplicity, specificities of receptor binding are shown only for the epidermal growth factor (EGF). ErbB2 binds no ligand with high affinity, and ErbB3 homodimers are catalytically inactive (crossed kinase domains). **B -** Signaling to the adaptor/enzyme layer is shown only for two receptor dimers: the weakly mitogenic ERGF homodimer, and the relatively potent ErbB2–ErbB3 heterodimer. Only some of the pathways and transcription factors are represented in this layer. Numbers 1 to 5 refer to the main text. C How the signal from the above layers are translated to specific types of output; many signaling cascades are involved in these complex signal transduction pathways. (Adapted from Yarden 2001)

These growth factors bind with different specificities and affinities to EGFR, ErbB3 and ErbB4, while no ligand for ErbB2 has been identified yet (**Fig. 4**). Of note, EGF, TGFα, EPI and AREG are specific for the EGFR (Hynes and Lane 2005; Schneider and Yarden 2014), while BTC, HB-EGF and EREG have been shown to have a dual specificity towards EGFR and ErbB4. Finally, the neuregulins can bind to ErbB3 and/or ErbB4 according to their own specificity.

Despite the fact that ErbB2 is an orphan receptor (Lonardo, Di Marco et al. 1990), it is well known that it acts as a co-receptor for the other EGFR family members. In particular, the EGFR ligands and the neuregulins that bind to ErbB3 or ErbB4 are able to induce heterodimer formation with ErbB2. These heteromolecular interactions are of pathophysiological relevance, because such receptor combinations show strong mitogenic signaling and tumorigenicity. Indeed, signaling pathways involving Ras, Src, phosphatidylinositol 3-kinase (PI3K), MAPK/ERK and c-Jun NH2-terminal kinase (JNK) have been shown to be activated by ErbB2-expressing cell lines (Yamamoto, Saito et al. 2011).

2.4 - EGF-dependent signal transduction

Trans-autophosphorylation of intracytoplasmic tyrosine residues determines the presence of pY signals that are interpreted by the cell through molecules harboring SH2 or PTB domains (see Introduction Section 1.1; Schlessinger 2000). There are 10 tyrosine residues in the EGFR that can be phosphorylated upon ligand binding and which could serve as docking sites for SH2 and PTB domain-containing proteins (Shoelson 1997). Only seven of these tyrosine residues have been demonstrated to be linked to signaling effector activity. A signal-specific array of molecules is recruited to the PM, where the activated receptor still resides, and can initiate distinct signaling pathways. Ongoing research in the field is still attempting to link specific phosphosites to the different effectors. It has been reported that many pYs can

interact with multiple effectors, and conversely, some effectors bind to multiple pYs, as in the case of Grb2 and Shc (Schulze, Deng et al. 2005).

The major EGF/EGFR signaling pathways include the MAPK/ERK, the PI3K/AKT, the phospholipase C (PLC)/protein kinase C (PKC), the Janus kinase (JAK) and the signal transducer and activator of transcription (STAT) proteins (Kloth, Catling et al. 2002; Andl, Mizushima et al. 2004), the JNK and p38 MAPKs (Mialon, Sankinen et al. 2005), and the Ca^{2+}/calmodulin-dependent protein kinase (CaMK; Sengupta, Bosis et al. 2009).

The transcriptional response to EGF can be divided into three temporal phases. The initial wave, up to 45 minutes from stimulation, consists of a limited set of genes, called *immediate early genes* (IEGs), which include transcription factors and cell cycle regulators, such as FOS and JUN (Kujubu, Norman et al. 1991). The second wave is activated 45-120 minutes after stimulation, and drives the activation of *delayed early genes* (DEGs) and includes many newly induced phosphatases, as well as DNA and RNA-binding proteins. This second wave regulates components that positively or negatively influence the EGFR signal (Amit, Wides et al. 2007). DEGs act by strongly shutting down IEGs either by inhibiting upstream signal transduction pathways or by promoting transcriptional attenuation. The late, *secondary response genes* (SRGs) are activated after 120 minutes from stimulation and confer stable phenotypes, which are crucial for fate determination.

2.4.1 - The MAPK/ERK pathway

The MAPK/ERK cascade (**Fig. 4**, yellow label 1) is a critically important signaling pathway that transmits the signal of extracellular stimuli to the nucleus, to regulate cellular processes such as proliferation, differentiation and motility. The signaling pathway that acts through the ERK cascade is mediated by sequential phosphorylation and activation of protein kinases. The central phosphorylation stream of the cascade includes Raf kinases, MEK1/2, ERK1/2 and ribosomal S6 kinases (RSKs): once the EGFR has been activated, the complex formed by

the adaptor proteins Grb2 and son-of-sevenless (SOS) can bind the EGFR through the pY signals (Lowenstein, Daly et al. 1992; Batzer, Rotin et al. 1994). This interaction leads to a conformational modification of Sos, which is then able to activate Ras. In turn, Ras activates Raf-1 that, through intermediate steps, phosphorylates the MAPKs, ERK1 and ERK2 (Hallberg, Rayter et al. 1994; Liebmann 2001). As a result of this phosphorylation cascade, ERK kinases act on a dual level. First, they can phosphorylate cytoplasmic proteins, such as protein kinases and phosphatases, cytoskeletal elements, regulators of apoptosis, and a variety of other signaling-related molecules. Second, they migrate into the nucleus to phosphorylate target transcription factors, thus promoting transcription of mitogenic genes (Hill and Treisman 1995; Gaestel 2006).

2.4.2 - The PI3K/AKT pathway

The PI3K/AKT pathway (**Fig. 4**, yellow label 2) is involved in cell growth, apoptosis resistance and invasion. PI3K is a dimeric enzyme composed of a regulatory p85 subunit and a catalytic p110 subunit. p85 is responsible for the anchorage to ErbB receptor-specific docking sites, whereas p110 generates the second messenger, the phosphatidylinositol trisphosphate (PIP_3). This messenger molecule in turn promotes phosphorylation and activation of the serine/threonine kinase AKT (**Fig. 4**) (Vivanco and Sawyers 2002). Although docking sites for p85 are absent on EGFR, they are instead abundant on ErbB3 (Carpenter, Auger et al. 1993; Yarden and Sliwkowski 2001). Indeed, the dimerization of the EGFR with ErbB3 represents the principal mechanism that drives EGFR-dependent PI3K pathway activation. Alternatively, the p85 subunit can interact with EGFR through the docking protein Grb2-associated-binding protein 1 (Gab-1; Mattoon, Lamothe et al. 2004).

2.4.3 - The PLCγ pathway

The EGFR also activates PLCγ (**Fig. 4**, yellow label 3), an enzyme that drives the hydrolysis of phosphatidylinositol bisphosphate (PIP_2) to inositol triphosphate (IP_3). This hydrolysis event is critical for the intracellular release of Ca^{2+} from the endoplasmic reticulum (ER) and subsequent activation of calcium-regulated pathways, and the release of diacylglycerol (DAG), a cofactor in PKC activation (Chattopadhyay, Vecchi et al. 1999; Patterson, van Rossum et al. 2005). PLCγ interacts directly with the phosphorylated EGFR via its SH2 domain and its activation results in MAPK and JNK activation (Schonwasser, Marais et al. 1998; McClellan, Kievit et al. 1999).

The EGF-induced increase in the cellular concentration of Ca^{2+} exhibits two components: a release of Ca^{2+} from intracellular stores, such as the ER, and a net influx of Ca^{2+} from the extracellular milieu through the channel transient receptor potential channel 1 (TRPC1). Ca^{2+} entry through TRPC1 conversely activates EGFR, suggesting that TRPC1 is a component of a Ca^{2+}-dependent amplification of EGF-dependent cell proliferation (Tajeddine and Gailly 2012).

2.4.4 – The JAK/STAT pathway

Another signaling pathway relies on the activity of the JAK/STAT pathway (**Fig.4**, yellow label 4). Upon ligand-induced receptor phosphorylation, JAK forms homodimers that are capable of trans-autophosphorylation. Phosphorylated JAK activates the cytosolic transcription factor STAT, which translocates to the nucleus and initiates transcription of specific target genes (Haura, Zheng et al. 2005). Importantly, constitutive activation of STAT proteins and especially STAT3 has been observed in numerous primary tumor cells and tumor-derived cell lines. Indeed, increased activity of membrane-associated tyrosine kinases, such as the EGFR, promotes persistent STAT3 activation, which contributes to oncogenesis and tumor progression (Bromberg 2002).

2.4.5 – The Src kinase pathway

The proto-oncogene, non-receptor tyrosine kinase Src (**Fig. 4**, yellow label 5) plays a critical role in the regulation of cell proliferation, migration, adhesion, angiogenesis, and immune function. Src, which is located in the cytosol, activates a series of substrates, including focal adhesion kinase (FAK), PI3K, and STAT proteins (Summy and Gallick 2006;Yeatman 2004). Although Src can function independently, it also cooperates with other tyrosine kinases, such as RTKs. Src interaction with the EGFR is complex: on the one hand, Src serves as a signal transducer and enhancer of EGFR activation (Jorissen, Walker et al. 2003; Leu and Maa 2003), and on the other, Src appears to be involved in resistance to anti-EGFR therapies through its EGFR-independent activation, either by Src-dependent EGFR transactivation through distinct pathways involving integrins and G-protein coupled receptors (GPCRs) (Kopetz 2007).

2.5 - EGFR dephosphorylation

For what regards EGFR phosphorylation kinetics, many data are available in literature (Guo, Kozlosky et al. 2003; Boeri Erba, Matthiesen et al. 2007). In contrast, the kinetics of EGFR dephosphorylation are little understood, although much is known about the specific phosphatases that can regulate EGFR (Tonks 2006). The first protein tyrosine phosphatase (PTP) to be purified and characterized was PTP1B in 1988 (Tonks, Diltz et al. 1988; Tonks, Diltz et al. 1988). Since then, over 100 PTPs have been identified in the human genome. PTP1B was shown to localize to the ER and to dephosphorylate several RTKs, including EGFR (Frangioni, Beahm et al. 1992; Lammers, Bossenmaier et al. 1993). This finding led to the initial hypothesis that EGFR can only be regulated by PTPs after internalization. This hypothesis was corroborated by a successive study that demonstrated that EGFR could only interact with PTP1B after EGFR endocytosis (Yudushkin, Schleifenbaum et al. 2007). These

observations have contributed to the widespread assumption that EGFR dephosphorylation occurs only in the cell interior, after EGFR internalization. However, the identification of multiple cell surface-localized PTPs that are able to dephosphorylate EGFR, suggested that EGFR regulation by PTPs might not be restricted to the cell interior (Monast, Furcht et al. 2012). This notion is supported by a study in MCF7 and COS7 cells, in which EGFR phosphorylation was reduced to basal levels after approximately 2 min of treatment with an EGFR tyrosine kinase inhibitor (TKI), regardless of the EGFR cellular localization (Offterdinger, Georget et al. 2004). The same study demonstrated that EGFR could become phosphorylated at the cell surface in the absence of stimulatory ligand if cells were treated with pervanadate, a potent and irreversible inhibitor of PTPs (Huyer, Liu et al. 1997).

These data suggest that EGFR pYs are under the control of PTPs, regardless of EGFR cellular localization, in contrast to the classic view of EGFR dephosphorylation. These data also support the possibility that PTPs are important regulators of EGFR-mediated signaling and internalization.

2.6 - From input to output

One big task in the study of EGFR signaling pathways, and more broadly RTK signaling pathways, is to decode the exact mechanisms, and their timings, by which an extracellular signal is translated to a specific cellular outcome. This goal is particularly challenging when dealing with receptor signaling pathways, because they are highly complex and interconnected (Citri and Yarden 2006). For instance, a decade ago, a review on the ErbB signaling network assigned 122 proteins to the network with 211 interactions among these proteins (Oda, Matsuoka et al. 2005). It is plausible that this network has been enriched with new players and interactions since the publication of the review.

In addition to the intricate network of signaling proteins associated with the EGFR signaling pathway, another level of regulation exists that influences EGFR signaling, i.e.,

endocytosis of the receptor (see Chapter 5). Following activation, the EGFR undergoes cycles of endocytosis and recycling, while still phosphorylated and active (Sigismund, Confalonieri et al. 2012), thus complicating the regulation of the signaling pathway.

Although biochemistry and molecular biology offer tools to study local and direct interactions within networks, and cell biology allows the interpretation of spatial input–output connections, this experimental approach is unable to draw connections that link long-range or systemic elements, as well as to infer the dynamic aspects of complex networks (Enuka, Feldman et al. 2015). A possible solution aimed at overcoming the limitations of experimental approaches, is the application of computational modeling to biological data to reveal the hidden functional and topological properties of the protein under analysis, as will be discussed in Chapter 7.

3 - EGFR Ubiquitination

The post-translational modification of a wide array of proteins consisting of the covalent attachment of one or more ubiquitin (Ub) moieties has emerged as one of the major mechanisms involved in protein regulation. Signaling receptors are tightly regulated by ubiquitination, and Ub is responsible for receptor trafficking, sorting and downregulation (reviewed in Haglund and Dikic 2012). In the next pages, the structure of Ub and the signaling network generated by this small protein, which is responsible for complex and various cellular responses, will be briefly discussed.

Ub was first identified in 1975 as a protein with an unknown function, being expressed ubiquitously in all eukaryotic cells (Schlesinger, Goldstein et al. 1975). In the early 1980s, the basic functions of Ub and some components of the ubiquitination pathway were elucidated by Aaron Ciechanover and colleagues (Ciehanover, Hod et al. 1978). The importance of this research was formally recognized when Ciechanover was awarded the Nobel Prize in Chemistry in 2004.

Ub is a small regulatory protein of only 8.5 kDa that is composed of 76 amino acid residues and is encoded in mammals by four different genes, reflecting the importance of its role in cellular dynamics. It is highly conserved among eukaryotic organisms, sharing an extremely high sequence identity: human and yeast share a 96% identity in their Ub amino acid sequences, differing only in 3 residues out of 76. No ubiquitination machinery is known to exist in prokaryotes, with the exception of the Prokaryotic Ub-like Protein found in *Mycobacterium tuberculosis* (Pearce, Mintseris et al. 2008).

3.1 - The ubiquitination machinery

Ubiquitination is a post-translational modification in which the carboxylic acid of the di-glycine motif in the activated Ub moiety is covalently attached to the ε-amino group of a lysine residue in the target protein.

The sequential activity of three types of proteins is necessary for substrate ubiquitination (**Fig. 5**; Woelk, Sigismund et al. 2007):

- **Ub-activating enzyme - E1**: this enzyme catalyzes the ATP-dependent activation of the carboxyl terminus of Ub and its conjugation to the active site cysteine residue in the E1.
- **Ub-conjugating enzyme – E2**: this enzyme accepts the Ub molecule from E1 in a trans-thioesterification reaction with an active cysteine in E2.
- **Ub-protein ligase – E3**: this class of protein is involved in substrate recognition and determines the specificity of the reaction. Indeed, there are few E1 enzymes, despite the fact that there are dozens of E2s and hundreds of E3s, creating a network that easily modulates the response after receiving a specific stimulus. The E3 ligase class of proteins includes hundreds of members, and

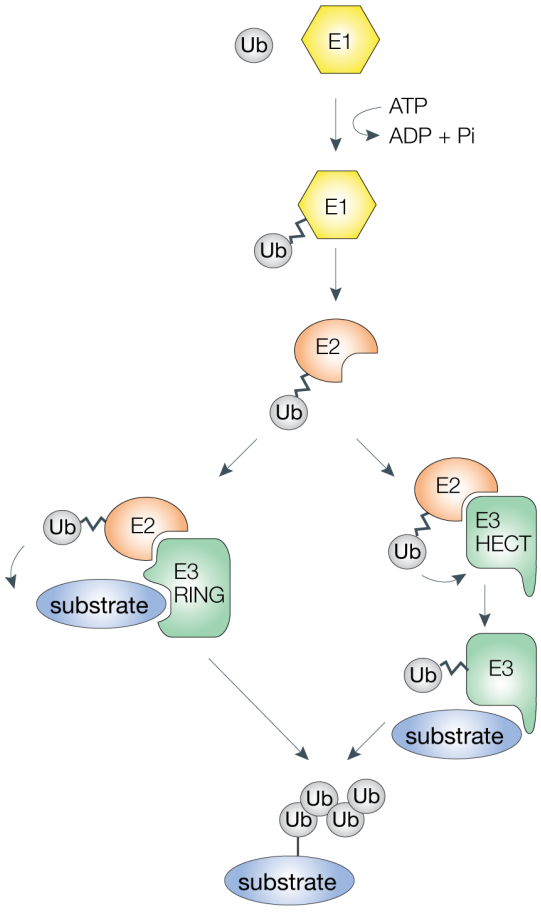


Figure 5: Schematic representation of the ubiquitination process. Hierarchical set of three types of proteins is required for substrate ubiquitination: the ubiquitin-activating enzyme (E1), the ubiquitin-conjugating enzyme (E2) and ubiquitin-ligase (E3). The two major classes of E3 ligases are shown: the RING-type E3-ligases function as adaptors, bringing the ubiquitin (Ub) moiety close to the substrate, whereas the HECT-type E3-ligases act as enzymes catalyzing directly substrate ubiquitination. (Adapted from Woelk, Sigismund et al. 2007).

can be subdivided in two main branches: the HECT ligases and the RING domain-containing ligases. E3 ligases that possess the RING domain work as adaptors, bringing the E2 close to the substrate and allowing the direct transfer of Ub from the E2 enzyme to the substrate, whereas those possessing the HECT domain act as real enzymes creating an E3-Ub intermediate state before transferring the Ub molecule to the target protein. In both cases, an isopeptide bond between a lysine residue in the target protein and the C-terminal glycine of Ub is formed.

3.1.1 - Different types of Ub signal

The amino acid sequence of Ub comprises seven lysine residues (**Fig. 6A**) that can be potentially used as acceptors for the attachment of other Ub moieties, allowing the formation of different types of Ub chains. However, it is currently not known if all of the linkages have a specific function (Woelk, Sigismund et al. 2007).

Ubiquitination has long been studied for its role in proteasomal degradation. The **polyubiquitination** (i.e., attachment of multiple Ub moieties or of a Ub chain - see next page) of target proteins to be destroyed by the proteasome is a well-known pathway, necessary for a number of physiological aspects of the cell, such as homeostasis and cell cycle progression, as well as cellular differentiation and apoptosis (**Fig. 6B**). More recently, new roles for Ub were highlighted, not involving proteasomal degradation. For example, an important process in which Ub plays a central role is the endocytosis and trafficking of PM receptors. In this case, ubiquitination has been shown to be involved in receptor internalization mechanisms or in targeting receptors for lysosomal degradation.

Substrate proteins can also be modified by the attachment of a single moiety of Ub, in a process denoted as **monoubiquitination**. This reversible modification is not a proteolytic signal, but is involved in various cellular mechanisms, such as endocytosis, endosomal sorting, histone regulation, DNA repair, virus budding and nuclear export (**Fig. 6B**). Moreover,

substrate proteins can also be ubiquitinated on several lysine residues, giving rise to **multiple monoubiquitination**, which is involved mainly in receptor internalization and endocytosis (**Fig. 6B**). In addition, the seven lysine residues of Ub themselves can be potentially used as acceptors of other Ub molecules, determining the formation of different types of Ub chains. As above mentioned, the process by which a protein is modified by the addition of one or more Ub chains is known as polyubiquitination, and the potential number of modifications that can be achieved is impressive. It has been shown that all the seven lysine residues in Ub can be used for chain formation *in vitro*, but the physiological relevance of the different types of Ub chains is far from being fully understood.

There is evidence indicating that structurally different polyUb chains contribute to the generation of functionally distinct signals. Indeed, the importance of the specific lysine-dependent linkage was first highlighted in yeast, where a mutation that abolished the Lys63 linkage (Ub-Lys63Arg) resulted in a DNA repair-defective phenotype (Spence, Sadis et al. 1995). Lys63 linkage has since been shown to be a non-proteolytic signal involved in DNA repair, transcriptional regulation, endocytosis and activation of protein kinases (Bach and

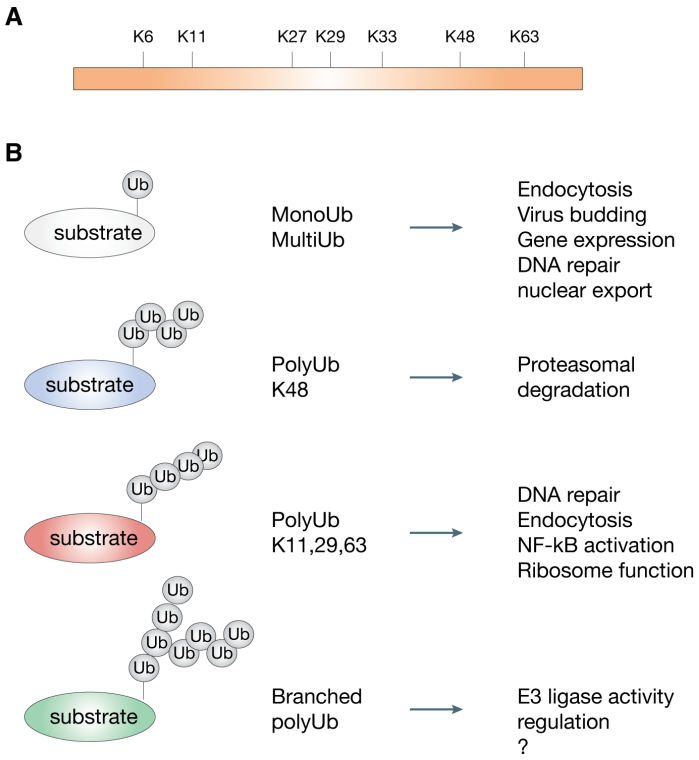


Figure 6: Schematic representation of the different ubiquitin modifications.
A - Schematic representation of Lys residues (K) present in the ubiquitin amino acid sequence **B** - The different types of ubiquitin (Ub) modification with associated biological functions are shown. Target proteins can be modified by a single or by multiple monoubiquitins (MonoUb and MultiUb, respectively), or by different types of polubiquitin (PolyUb) chains, both linear and branched (K48-linked, linear K11,29,63-linked, and branched polyUb chains). The question mark indicates that the functions of branched PolyUb chains are mostly unknown (adapted from Woelk, Sigismund et al. 2007).

Ostendorff 2003; Huang and D'Andrea 2006) see also Introduction Section 3.2 for details on its role in endocytosis). Instead, Ub-mediated degradation via the 26S proteasome, the most well-known Ub function, requires the presence of four or more Lys48-linked Ub molecules. Structural studies showed that there is a conformational difference between Lys63- and Lys48-linked chains, highlighting the possibility that chain recognition could be based on the topology adopted by different chain types. This hypothesis could explain how specific cell components interpret the diversity of Ub signals, and translate them into specific biological responses. This concept also raises the idea that ubiquitination is a code, exploited by cells to diversify cellular responses starting from a narrow range of different molecules.

Indeed, the Ub signal is emerging to be as versatile as the phosphorylation-based signal in transducing intracellular outputs. Similarly to phosphorylation (see Introduction Section 2.1), a reversible post-translational modification that can be recognized by specific protein domains (e.g. the SH2 domain), ubiquitination is an inducible signal (e.g. in response to growth factors or DNA damage) that can be recognized by proteins harboring Ub-binding domains (UBDs). These Ub-UBD interactions generate a network of Ub-based connections.

Ubiquitination is also reversible due to the presence of deubiquitinating enzymes (DUBs), which modulate the cellular response to the Ub signal. Compared to phosphorylation, the ubiquitination signal is even more complex because the 76-residue molecule appended to the target proteins is obviously bigger than a phosphate group. In addition, as we discussed above, Ub can form chains that increase the number of, and differentiate between, the possible readouts.

3.2 - Ligand-induced ubiquitination of the EGFR

One of the most studied examples of a receptor undergoing Ub-mediated downregulation following ligand stimulation is the EGFR. EGFR is a well-characterized model for receptor internalization and for ligand-induced trafficking (for a recent review see Tomas, Futter et al.

2014). Once the EGFR is activated by its cognate ligand EGF, it undergoes structural and biochemical modifications including intracytoplasmic tail phosphorylation (see Introduction Section 2.1). At this stage, EGFR can subsequently be ubiquitinated by a specific machinery that is recruited to the receptor tail by signals transmitted from the activated receptor itself. EGFR ubiquitination starts at the plasma membrane (PM), however, it can also continue all along the endocytic route (Stang, Blystad et al. 2004; Umebayashi, Stenmark et al. 2008).

EGFR ubiquitination is exerted by the E3 RING ligase Cbl (Levkowitz, Waterman et al. 1998; Waterman, Levkowitz et al. 1999). The Cbl family is composed of three genes: c-Cbl, Cbl-b and Cbl-c (Schmidt and Dikic 2005; Lipkowitz and Weissman 2011). c-Cbl (henceforth Cbl) is best characterized for its role in EGFR ubiquitination and is responsible also for the ligand-induced ubiquitination of other RTKs, for example the mesenchymal epithelial transition factor (MET) receptor (Peschard, Fournier et al. 2001; Mettlen, Pucadyil et al. 2009). For these receptors, it has been shown that Cbl recognizes and binds directly to a phosphotyrosine (Emde, Kostler et al. 2012) in the activated receptor, through its N-terminal TKB domain and drives the attachment of Ub moieties to the intracytoplasmic tail of the same receptor. In addition, the adaptor protein Grb2, which is recruited to other pY residues on the cytoplasmic tail of the RTK, was found to be an additional, indirect docking site for Cbl to the receptor, both in the case of the EGFR and MET. Cooperation of the direct and indirect binding sites has been proposed (Capuani, Conte et al. 2015), in the case of the MET receptor, to be necessary for the correct positioning of Cbl on the receptor or for full receptor ubiquitination (Peschard, Fournier et al. 2001).

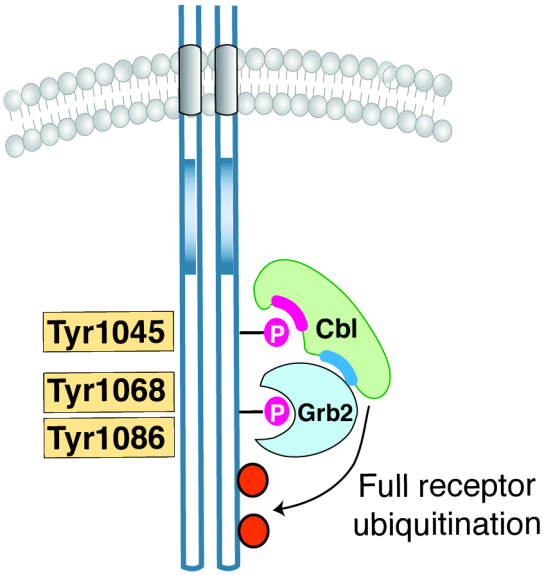


Figure 7: EGFR ubiquitination. Schematic representation of cooperative binding between the Cbl/Grb2 complex and the EGFR. Cbl can bind directly to the receptor through pY1045, and indirectly through the interaction with Grb2, which binds pY1068/86 on the receptor. Only the binding of both Cbl and Grb2 to the receptor allows for full ubiquitination of the EGFR.

The interaction between Cbl and the EGFR has also been hypothesized to be a bi-dented contact. Cbl is recruited to the activated EGFR by two distinct mechanisms (**Fig. 7**): it can interact directly with a specific site in the receptor tail - the pY1045 (Waterman, Levkowitz et al. 1999), or it can indirectly interact with the receptor through the Grb2 adaptor that binds to pY1068 or pY1086 (Waterman, Katz et al. 2002; Jiang, Huang et al. 2003). We have shown that a two-pronged cooperative interaction between Cbl and the EGFR, involving both pY sites, is needed for optimal recruitment of Cbl to the receptor and for an efficient ubiquitination (Sigismund, Algisi et al. 2013).

We have previously described how EGF concentration varies among different body tissues and fluids, being present from 1 ng/ml up to hundreds of ng/ml even in physiological or pathological conditions (see Introduction Section 2.2.2). Interestingly, it was observed by our lab that in HeLa cells, treated with increasing doses of EGF, the ubiquitination of the EGFR increases sharply over a narrow range of EGF concentrations: EGFR-Ub is minimal at 1 ng/ml and nearly maximal at 10 ng/ml (**Fig. 8A**). Conversely, the EGFR-pY content, used as a surrogate for receptor activation, follows a typical hyperbolic dose-response curve, which can be translated into a linear behavior when a log scale is used for EGF concentration (**Fig. 8B**). This behavior was referred to as the threshold effect for EGFR ubiquitination due to the

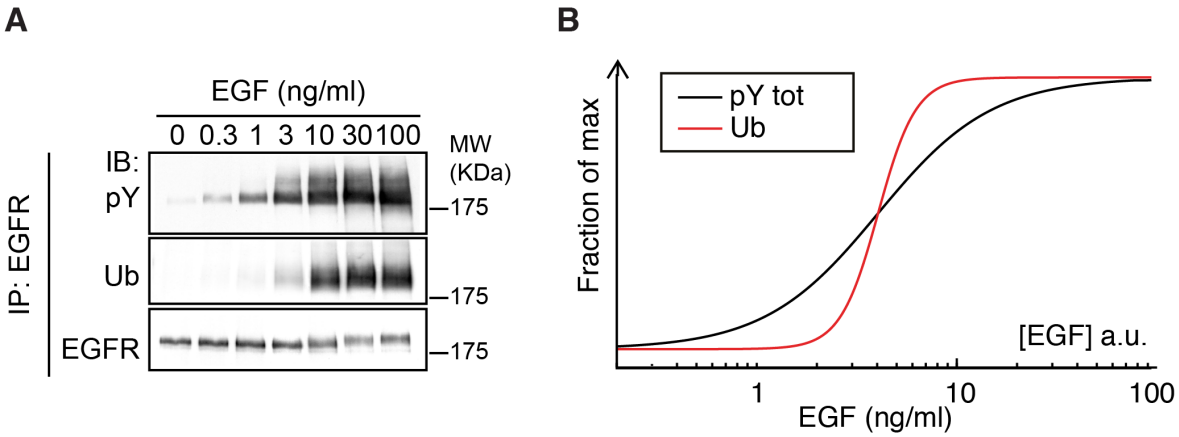


Figure 8: EGFR and ubiquitination. A - HeLa cells were stimulated with increasing concentrations of EGF for 2 min, as indicated. Total cell lysates were subjected to immunoprecipitation (IP) using an anti-EGFR antibody and immunoprecipitates were then analyzed by immunoblotting (IB) using the indicated antibodies: anti-ubiquitin (Ub), anti-total phosphotyrosine (pY), anti-EGFR intracellular domain (EGFR). **B -** Quantitation of the pY and Ub immunoblots shown in A by densitometry. A threshold effect for EGFR ubiquitination is evident, whereas EGFR phosphorylation presents a gradual increment over EGF concentrations. (From Sigismund, Algisi et al. 2013; Capuani, Conte et al. 2015).

fact that ubiquitination is minimal at low doses of EGF, maximal at high doses of EGF, but presenting a rapid increase in a very narrow interval of EGF concentrations.

The mechanism of threshold generation depends in part on the cooperativity of Cbl recruitment to the two pY sites, but the exact machinery that determines the ubiquitination threshold as a function of EGF concentration remains to be elucidated and represents the subject of the present thesis work.

Initial studies on EGFR ubiquitination showed that when stimulated with high doses of EGF, the receptor presented a multi-monoubiquitination pattern (Haglund, Sigismund et al. 2003). Moreover, a single Ub moiety was sufficient to drive internalization of the receptor, although multiple Ub signals were more efficient at driving the process. This is possibly due to a higher affinity of the ubiquitinated proteins for ubiquitin receptors, that contain a Ub-interacting motif (UIM) domain, able to specifically recognizes and bind to ubiquitin. Instead, thanks to mass spectrometry analysis approaches, it is now widely accepted that EGFR can be both mono- and poly-ubiquitinated through Lys63 chains (Huang, Kirkpatrick et al. 2006), but not through Lys48, amplifying the diversity of EGFR-Ub signals. It has been speculated

that Lys63 linked chains can increase the avidity of binding of the receptor to UIM-containing proteins or act as a completely distinct signal with respect to mono-Ub.

Ubiquitination of the EGFR was initially shown not to be essential for its internalization. Indeed, different studies in which ubiquitination of the EGFR was abolished, either by mutagenizing the Cbl binding site (Huang, Kirkpatrick et al. 2006) or the acceptor lysine residues in the receptor (Haglund, Sigismund et al. 2003), demonstrated that internalization was impaired. However, this simplistic view of EGFR internalization has now been replaced by a much more complicated scenario in which the receptor can be internalized by multiple and redundant pathways with different Ub requirements (Goh, Huang et al. 2010).

Although not essential for EGFR internalization, EGFR ubiquitination plays a fundamental role both in receptor endocytosis and intracellular fate. Indeed, we have previously shown that when EGFR is stimulated with low doses of EGF, it almost exclusively undergoes internalization through a clathrin-dependent pathway and, importantly, it is not ubiquitinated. Conversely, as EGF doses reach higher concentrations, a substantial fraction of the receptor becomes ubiquitinated and is internalized through a clathrin-independent pathway (Sigismund, Woelk et al. 2005). Later, it was discovered that the two internalization pathways are associated with distinct receptor fates: the clathrin-dependent pathway is mainly associated with EGFR recycling to the PM and sustained signaling (Sigismund, Argenzio et al. 2008), whereas clathrin-independent endocytosis targets the receptor to lysosomal degradation resulting in signal attenuation (Sigismund, Argenzio et al. 2008; Sigismund, Algisi et al. 2013). Thus, the balance between these two pathways controls the fate (recycling *vs.* degradation) of the EGFR, and consequently the level of signaling and the final biological outcome. Together, these findings highlight the importance of the cell's response to increasing levels of ligand and how this response is converted into a biological outcome, as discussed in more detail in the following sections.

4 - Endocytosis: a brief overview

The PM is a dynamic structure that separates the intracellular environment from the extracellular space, regulating the transport of small and big molecules. Whereas small molecules, such as amino acids, sugars and ions, can cross the membrane, reaching the intracellular compartment through integral membrane proteins like pumps or channels, macromolecules are internalized by vesicles derived from invagination of the PM itself, in a process known as endocytosis.

EGFR trafficking is a well-characterized series of events determined by EGF stimulation: following ligand-induced activation, the EGFR is internalized by the cell through different pathways. Plenty of evidence supports the notion that endocytosis is the major mechanism by which the cell attenuates RTK signals, by removing active receptors from the PM. However, recent data has demonstrated that endocytosis is also involved in sustaining signaling, thus dramatically changing our perspective on this process (Sorkin and von Zastrow 2009; Scita and Di Fiore 2010). In the next paragraph (4.1), a general overview on endocytosis will be presented, while in the following section (5), the mechanisms underlying the different EGFR internalization pathways and their downstream effects, will be described.

4.1 - The endocytic pathways: different entry portals and sorting routes

Endocytosis is tightly linked to almost all aspects of cellular signaling, leading to the notion that endocytosis is actually the master organizer of cellular signaling, providing the cell with understandable messages, resolved in space and time. In essence, endocytosis provides the communications and supply routes (the logistics) of the cell (Le Roy and Wrana 2005; Platta and Stenmark 2011; Sigismund, Confalonieri et al. 2012).

Based on the present knowledge, it is thought that endocytosis has evolved for the purpose of bringing nutrients inside the cell. Indeed, early prokaryotic life forms used relatively simple transport mechanisms, such as pumps or channels, to transport essential molecules through the PM. As prokaryotes evolved to eukaryotes, more complex "entry portals" began to appear, perhaps due to the selective pressure provided by the transition from a situation in which nutrients were present in a concentrated form, to a new and more diluted environment. This transition might have selected for those life forms capable of actively searching for and concentrating nutrients (Sigismund, Confalonieri et al. 2012).

Regardless of how endocytosis evolved, in mammalian cells it can occur by distinct mechanisms, whose complexity already starts at the PM, where multiple entry portals have been described (**Fig. 9** and **Table II**). A traditional way to differentiate between the different endocytic pathways is on the basis of cargo size. Particles larger than 500 nm are taken up by phagocytosis, a specific form of endocytosis occurring in specialized cells, such as immune system cells, which can actively remove solid particles, e.g. pathogens (bacteria) and cell debris (apoptotic cells). Phagocytosis is dependent on the receptor-ligand interaction, a process through which the specialized cell recognizes a ligand on the particle, which initiates the endocytosis of the whole particle. The internalized particle is then broken down by enzymes and absorbed into the cell. The uptake of fluids and solutes generally occurs by macropinocytosis in which smaller particles are brought into the cell within vesicles up to 500 nm in diameter, that subsequently fuse with lysosomes to hydrolyze the cargo particles. Both phagocytosis and macropinocytosis involve large rearrangements of the PM guided by extensive actin cytoskeleton remodeling, and coordinated by the stepwise involvement of the RHO-GTPase.

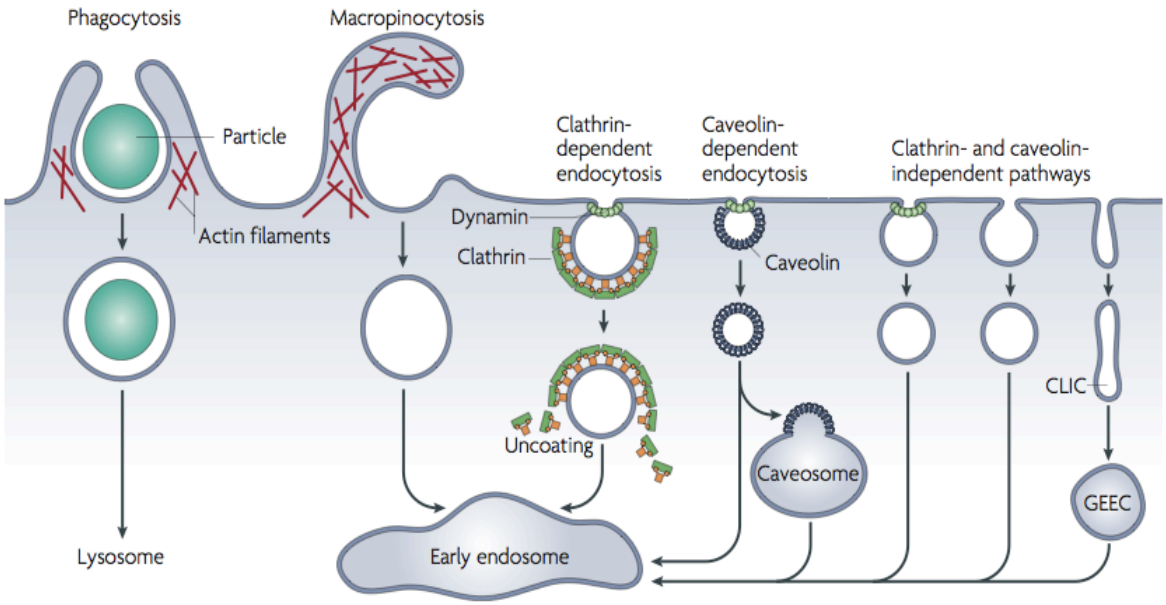


Figure 9: Pathways of entry into the cell. The first distinction between the different entry routes is based upon the size of the entering cargo and thus the type of membrane rearrangements. Large particles can be taken up by phagocytosis, whereas fluid uptake can occur by macropinocytosis. These two processes are dependent on extensive actin-mediated remodeling of the PM. Compared with the other endocytic pathways, phagocytosis and macropinocytosis lead to bigger endocytic vesicles. Molecules and receptors residing on the PM can be internalized through mechanisms that are dependent on coat proteins (e.g. clathrin- and caveolin-dependent pathways) or independent of coat proteins (e.g. clathrin- and caveolin-dependent pathways, and the CLIC-GEEK pathway). Most internalized cargoes are delivered to the early endosome via vesicular or tubular intermediates that are derived from the plasma membrane. (Adapted from Mayor and Pagano 2007).

Smaller invaginations (< 200 nm), are instead characteristic of micropinocytosis, which includes clathrin-mediated endocytosis (CME) and non-clathrin endocytosis (NCE). Both CME and NCE are required for the endocytosis of transmembrane receptors, a wide category of proteins that carry out many different roles and which have important implications for cell physiology. In the late 70s, Brown and Goldstein established the concept of receptor-mediated endocytosis in the optimization of the uptake of nutrients. They discovered a clever strategy used by cells for increasing the efficiency of the internalization of macromolecules, by concentrating them in focused regions of the PM, and thus reducing the energy consumption (Anderson, Brown et al. 1977).

Receptor-mediated endocytosis also harbors other interesting implications connected with the signaling network downstream of the internalized receptor. As illustrated in the following sections, RTKs are internalized by ligand-mediated endocytosis, a process that is predicted to

remove active signaling receptors from the PM, and to target them for lysosomal degradation. However, this process was also found to recycle the receptor back to the membrane, determining an extraordinary plasticity of the PM in response to various external stimuli, and allowing the cell to polarize the signal (Tomas, Futter et al. 2014). Moreover, it has been found that endocytosis of active RTKs could sustain signaling, exposing the receptors to substrates that were inaccessible at the PM (Sigismund, Confalonieri et al. 2012).

The vast impact of endocytosis on cellular homeostasis highlights the importance of understanding this process and its many complex pathways. In the following sections, CME and NCE, along with their impact on RTK signaling will be discussed in more detail.

4.2 - Clathrin-mediated endocytosis

Among the different cell entry routes, the best-characterized endocytic pathway is CME (McMahon and Boucrot 2011), whereby PM-resident cargoes are removed from the cell surface via clathrin-coated pits, a multi-component endocytic machinery (Schmid and McMahon 2007; Kirchhausen, Owen et al. 2014). The CME pathway involves the action of adaptor molecules (e.g., adaptor protein 2; AP2) that bridge the internalizing cargo to clathrin. Polymerization of the latter drives the progressive invagination of the pit, which is later released into the cytoplasm as an endocytic vesicle through the action of the GTPase dynamin (Marks, Stowell et al. 2001; Mettlen, Pucadyil et al. 2009). Indeed, the formation of invaginated pits and the subsequent budding of the vesicles is an energetically demanding process and the contribution in energy, above the structural contribution, is provided by the cooperation of clathrin with many different proteins (McMahon and Boucrot 2011). Although the most abundant proteins found in coated pits are clathrin and AP2, many other proteins, such as EPS15 (which derives its name from ‘EGFR pathway substrate clone number 15’), EPS15-interacting proteins (epsins) and adaptor protein 180 (AP180), associate with clathrin-coated pits and assist in coat formation and vesicle release. Some of these proteins, are found

in coated pits, but are not enriched in clathrin-coated vesicles (CCV), suggesting that they play assisting roles during clathrin-coat assembly (Mousavi, Malerod et al. 2004). The wide variety of proteins involved in CME has raised the possibility that they might be required for the formation of distinct types of clathrin pits, specialized in cargo-selection and in specific intracellular fate (Sigismund, Confalonieri et al. 2012).

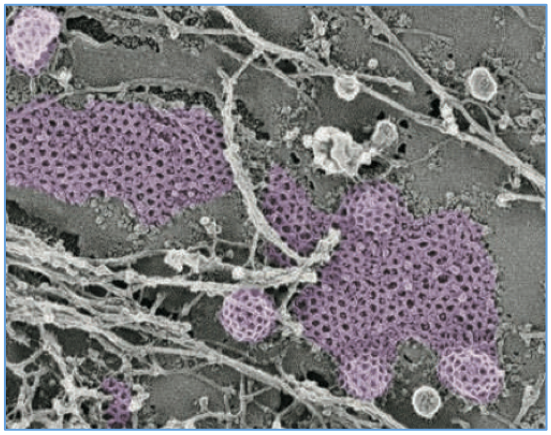


Figure 10: The clathrin lattice. Rapid freeze-etch micrograph views of clathrin lattices on the inner surface of a HeLa cell. It shows large, flat hexagonal sheets and the progression to a deeply invaginated state before membrane scission. (Picture from Traub 2009).

Clathrin-coated pits and CCV were first observed by electron microscopy in the early 1960s (Roth and Porter 1964). They are characterized by a 'bristle-like' appearance of the coat in cross sections of forming vesicles. This is due to the presence of a lattice of hexagons and pentagons on the surface of the PM, making them easily discernible (**Fig. 10**). This distinctive pattern is given by the structure of the assembly unit of clathrin, a three-legged structure called the *triskelion* (**Fig. 11A**) that consists of three clathrin heavy chains (CHC) and three clathrin light chains (CLC) (Kirchhausen, Owen et al. 2014). When triskelia interact they form a polyhedral lattice that surrounds the vesicle: the three CHCs provide the structural backbone of the clathrin lattice, and the three CLCs are thought to regulate the formation of and

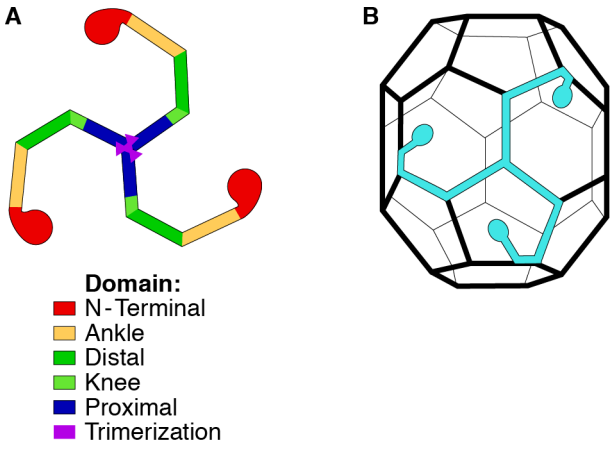


Figure 11: The architecture of clathrin. A - A schematic representation of a clathrin triskelion, which highlights the various domains of the clathrin heavy chain (CHC) with different colors (as indicated in the color legend). **B** - A clathrin barrel with a single triskelion highlighted in blue. (Adapted from Edeling, Smith et al. 2006).

disassembly of the clathrin lattice. When triskelia assemble, they interact with enough flexibility to form 6-sided rings - or hexagons - that yield a flatter lattice, or 5-sided rings - or pentagons - that are necessary for curved lattice formation. When many triskelions connect, they can form a basket-like structure that resembles the panels on a soccer ball (**Fig. 11B**); by constructing different combinations of 5-sided and 6-sided rings, vesicles of different sizes can assemble. These clathrin structures generate the force necessary to curve the membrane and pull it into a bud (Mousavi, Malerod et al. 2004).

As anticipated, clathrin alone is not sufficient to drive the invagination and budding of endocytic pits; AP2 is the principal protein associated with traffic at the PM in CME. AP2 is a heterotetrameric protein that binds clathrin and can interact with membrane proteins and lipids. It is a large protein complex that belongs to the Adaptor Protein family involved in vesicle formation in different subcellular compartments; however, it is the only member of the family involved in the formation of clathrin-coated pits at the PM. AP2 is composed of four subunits tightly bound to each other: α, β2, μ2 and σ2 (Edeling, Smith et al. 2006). The flexibility of the α2 and β2 subunits permits the simultaneous interaction of the complex with phosphoinositides in the PM and with clathrin, whereas μ2 can bind the cargo, thus orchestrating clathrin-coated pit assembly.

The assembly of the clathrin pits originates on the cytoplasmic side of the PM, where

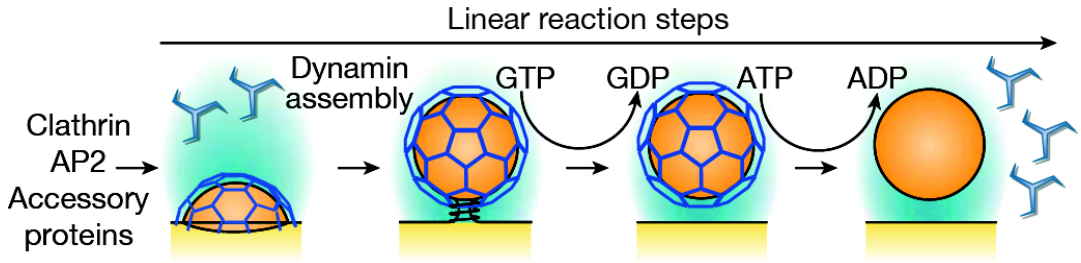


Figure 12: Clathrin dependent endocytosis. Clathrin-dependent endocytosis starts when clathrin and adaptor (AP2) complexes associate with the cargo, thus initiating the formation of a coated pit. As the pit matures, additional adaptor and scaffold proteins join the pit, providing a structural platform that regulates interactions between the adaptors and the other endocytic proteins. The pinching-off of the pit is an event driven by the GTPase dynamin, which releases the clathrin-coated vesicle (CCV) into the cytoplasm. (Adapted from Schmid and McMahon 2007)

triskelia start to assemble. The assembly is a progressive and dynamic event that starts from a small clathrin-coated bud, which grows into a pit that imposes on the membrane a deep curvature (**Fig. 12**). After the formation of deeply invaginated pits, the formation of endocytic vesicles requires fission of the budded membrane from the parent membrane. The large GTPase dynamin is required for this process. Dynamin has been most studied in the context of CCV budding from the PM, but, in addition to acting as a part of the scission machinery for the clathrin-dependent pathway, dynamin function is also implicated in vesicle formation in some clathrin-independent pathways (McMahon and Boucrot 2011).

4.3 - Non-Clathrin Endocytosis

One central obstacle in defining clathrin-independent endocytic pathways has been the lack of specific endocytic markers and pathway-specific cargoes. Analysis of the CME pathway significantly benefited from signature cargoes, such as the transferrin receptor (TfR) (McMahon and Boucrot 2011), which are enriched in the forming endocytic vesicle. Moreover, the transient assembly of clathrin and adaptor proteins marks the forming vesicle specifically in space and time (Kirchhausen, Owen et al. 2014). In contrast, clathrin-independent endocytosis has largely been defined by morphological criteria and by the persistent uptake of cargoes that can utilize multiple pathways following perturbation of the clathrin machinery. Thus, compared to CME, the current picture of NCE is at a much lower level of resolution. Additionally, NCE pathways have been studied in the context of CME perturbation. Abrogation of CME should affect the endocytosis of cargoes that specifically enter through that pathway, leaving relatively unaffected the endocytosis of many other types of cargo. Of note, exploiting methods that inhibit CME as a tool to discover new constitutive uptake pathways is far from a perfect approach, because such treatment might functionally up-regulate a pathway whose role is usually minor or even non-existent under physiological conditions.

The term NCE is used to refer to a heterogeneous group of pathways that share the common property of being insensitive to clathrin depletion, but that frequently depend on cholesterol-rich PM microdomains called rafts. The dependency of these pathways on rafts renders these pathways sensitive to pharmacological depletion of cholesterol (Cheng, Singh et al. 2006). Indeed, there is strong evidence that lipid rafts play a fundamental role in NCE pathways. Lipid rafts are membrane microdomains, particularly rich in cholesterol, glycosphingolipids, glycosylphosphatidylinositol-anchored proteins (GPI-APs), and signaling molecules (Pike 2009). They consist of a dynamic assembly of these structural elements that form liquid-ordered domains that float in the less-ordered surrounding membrane. Their size, lifetime, biogenesis and lipid/protein composition have yet to be defined. Nevertheless, it is thought that lipid rafts probably cluster into large platforms that can segregate membrane components, and they have been implicated in the regulation of various physiological processes, including protein trafficking and signal transduction.

The current classification for NCE pathways is based on three major criteria (**Table II**):

1. dependency on dynamin for vesicle release: clathrin-independent pathways are divided according to whether they use a dynamin-mediated scission mechanism (dynamin-dependent) or not (dynamin-independent) (Mayor and Pagano 2007);
2. presence of coat-like proteins, such as caveolins or flotillins in the case of caveolae-mediated or flotillin-mediated internalization, respectively;
3. dependency on small GTPases involved in the endocytic process, which control the entry of specific cargoes. In fact, the terminology Cdc42-, RhoA- or Arf6-regulated endocytic pathway indicates that modifying the function of these GTPases affects the internalization or trafficking of one set of NCE markers (Ellis and Mellor 2000; D'Souza-Schorey and Chavrier 2006).

Pathway	Morphology and size	Coat	Dynamin dependence	Small GTPase involved	Internalized cargoes	Associated/ regulatory proteins
Phagocytosis	Cargo shaped > 500 nm	None	No	RAC1/RHOA/ CDC42 (depending on type)	Pathogens, apoptotic cells, FcRs	Actin, ARP2/3; Formins; PI3K; WASP; WAVE2; amphiphysin; coronin; others
Macro-pinocytosis	Ruffled 0.2-10 μm	None	In some cases	RAC1, CDCD42, ARF6, RAB5	RTKs; fluids, some bacteria	Actin, ARP2/3, cortactin; PI3K; SRC; PAK1; Ras, CTBP1/BARS; others
Clathrin-mediated	Vesicular 150-200 nm	Clathrin	Yes	RAB5	RTKs; GPCR; TFR; some toxins; bacteria; viruses	AP-2, EPSINS, EPS15, intersectin, amphiphysin (plus many others, >50
Caveolae-mediated	Flask-shaped 50-120 nm	Caveolin 1 and 2	Yes	Not clear	GPI-linked proteins; CTxB; SV40	PTRF/cavin; SRC, SDPR
Flotillin-dependent	Vesicular	Flotillin 1 and 2	No	None	CTxB, CD59, proteoglycans (Glebov, Bright et al. 2006; Frick, Bright et al. 2007; Payne, Jones et al. 2007)	None as yet
CLIC/GEEC	Tubular	None	No	CDC42, ARF1	Fluids, bulk membrane, GPI-linked proteins	Actin; GRAF1; ARHGAP10
IL-2Rb	Vesicular 50-100 nm	None	Yes	RHO-A, RAC1	IL-2Rb	PAK1 and 2, cortactin, N-WASP
Arf6-dependent	Tubular	None	No (so far)	ARF6	MHC I; MHC II; CD59; CD55, GLUT1	None as yet

Table II: Pathways of internalization. Endocytic pathways are mainly classified based on the type of PM rearrangements that take place in the internalization event, which depend on the size of the cargo to be internalized. Large cargoes (> 500 nm) are internalized through phagocytosis, whereas fluids and solutes are generally internalized through macropinocytosis, which involves extensive actin rearrangements. Micropinocytic events are responsible for the internalization of smaller cargoes, such as PM resident proteins. These pathways are differentiated based on the presence of coat proteins, the requirement of dynamin or other small GTPases. For details on clathrin-mediated, caveolae-mediated, flotillin-dependent, and CLIC-GEEK pathways refer to the main text. The IL-2Rb pathway is specific for the uptake of IL-2Rb receptor and is dependent on the small GTPase RHO-A, whereas no coat proteins have been identified. The GTPase ARF6 seems to be involved in the clathrin- and dynamin-independent internalization of several proteins, such as the major histocompatibility complex I (MHC I), β1-integrin, E-cadherin and GPI-anchored proteins, although a direct role of ARF6 in a non-clathrin endocytosis (NCE) pathway has not yet been established. The similitude in cargo specificity with other internalization pathways suggests that IL-2Rb and ARF6-dependent pathways could represent variations of the same process.

By far, the best-characterized NCE pathway is the **caveolae-mediated endocytosis** pathway (Rothberg, Heuser et al. 1992; Pelkmans and Zerial 2005), which is dynamin-dependent and sensitive to cholesterol depletion (**Table II**). Caveolae are 50-80 nm flask-shaped PM invaginations that are marked by the presence of a coating element that belongs to the caveolin protein family (**Fig. 9**).

The presence of a coat protein is also a peculiar characteristic of the **flotillin-dependent pathway**, an NCE internalization mechanism in mammalian cells distinct from caveolae-mediated endocytosis (but still morphologically similar to it), and not inhibited by overexpression of GTPase-null dynamin mutants. It was also shown that flotillin1 is necessary for cholera toxin (CTxB) uptake (Glebov, Bright et al. 2006). Both caveolin and flotillin are inserted into the inner leaflet of the PM and it has been proposed that they induce membrane curvature.

The **CLIC/GEEC pathway** (CLathrin-Independent Carrier/GPI-anchored protein-Enriched Early endosomal Compartment) was discovered while studying the internalization of CTxB (Torgersen, Skretting et al. 2001) and GPI-anchored proteins (Kirkham, Fujita et al. 2005) (for a review see Lundmark, Doherty et al. 2008). This pathway is characterized by long and wide tubular invagination, in contrast to the spherical structures that are typical for the clathrin and caveolar pathways (**Fig. 9**). In the CLIC/GEEC pathway, a large volume of fluid-phase is co-internalized in a single endocytic event. Of note, the type of cargo internalized by this pathway is the same as that depicted in the case of caveolar internalization, underlining the redundancy between internalization pathways for some molecules. In addition, disruption of the CLIC/GEEC pathway appears to induce compensatory CME of GPI-linked proteins (Sabharanjak, Sharma et al. 2002), but whether this allows internalized proteins to be trafficked to their appropriate destinations remains unclear. The intracellular destination for the CLIC/GEEC pathway internalization route appears to differ between cell types, and

includes the lysosomal and the pericentriolar recycling compartments (Fivaz, Vilbois et al. 2002; Sabharanjak, Sharma et al. 2002).

One emerging NCE pathway relies on the activity of the coat-like protein endophilin, initially ascribed as a component of the CME. Recent evidence indicates that endophilin marks and controls a fast-acting tubulovesicular endocytic pathway, referred to as fast endophilin-mediated endocytosis (**FEME**; Boucrot, Ferreira et al. 2015). This pathway is independent of clathrin and AP2, and is activated upon ligand binding to cargo receptors, resulting in rapid (within seconds) ligand-induced internalization (Boucrot, Ferreira et al. 2015). The FEME pathway was shown to be dynamin-dependent and the main cargoes internalized by this pathway, so far, include some G-protein-coupled receptors, some RTKs (including the EGFR), the Interleukin-2 (IL-2) receptor, CTxB and Shiga toxins (Renard, Simunovic et al. 2015). The FEME pathway is prominent at the leading edge of the cell and, importantly, is also dependent on actin polymerization (Boucrot, Ferreira et al. 2015; Renard, Simunovic et al. 2015).

This link between an endocytic route and a cytoskeleton component highlights how endocytosis and actin dynamics are intimately connected. On the one hand, endocytosis and recycling can exert spatial and temporal control over a number of critical regulators of actin dynamics (Sigismund, Confalonieri et al. 2012), thus influencing the biological outcome of a variety of actin-based processes. On the other hand, the actin cytoskeleton is a strong requirement along the various steps of the endocytic process, both to provide structural support for the intermediates of membrane trafficking, but also to generate forces aimed to deform membranes into invaginations or permit the scission of vesicles and their motility.

4.4 - The endosomal sorting station

The budding of the vesicle into the cytoplasm is the first step in a series of events that are defined as endosomal sorting (**Fig. 13**). The endocytic compartments of mammalian cells are mainly endosomes and lysosomes and, for the purpose of endocytosis, they travel from the periphery to the center of the cell. In contrast, endosomes can also be directed backward to the external limits, in the case of recycling or when they are carrying exocytic cargoes.

Endosomes are compartments containing cargoes internalized at the PM, and which represent a station where cargoes can be sorted to the lysosome or recycled back to the cell surface. Endosomes can be divided into three categories - early, late and recycling - characterized by the time it takes the internalized cargo to reach them or by their function.

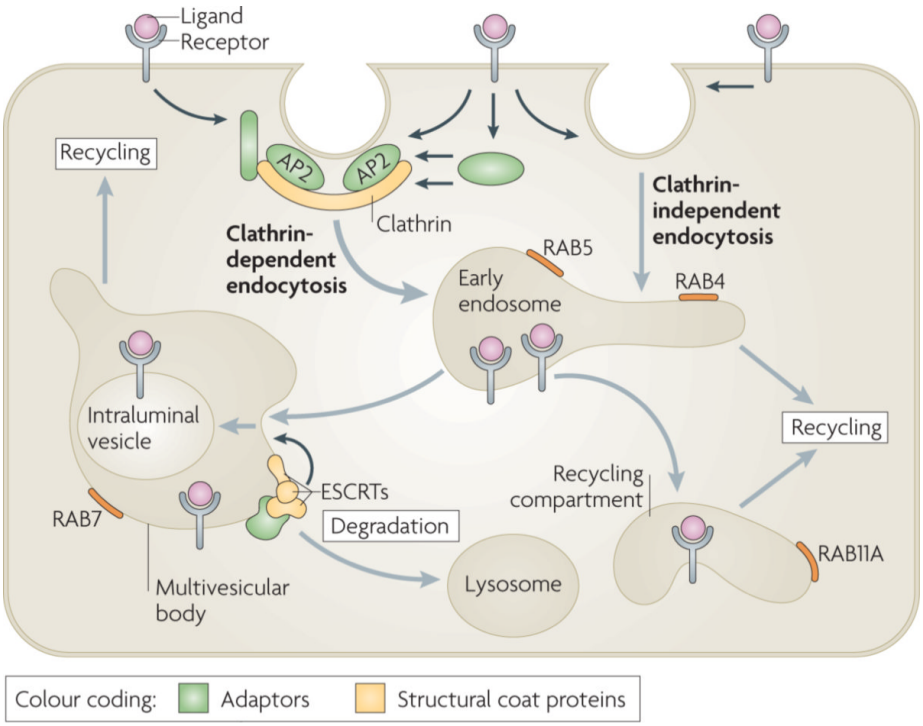


Figure 13: The endosomal sorting station. Endocytic vesicles derived from both clathrin-dependent and clathrin-independent internalization fuse at the level of early endosomes. Endosomal trafficking is controlled by several Rab proteins (small GTP- binding proteins of the Ras superfamily), which specifically reside in a particular type of endosome and which function by recruiting specific effector proteins. Following their internalization into early endosomes, receptors can recycle back to the plasma membrane or proceed to the late endosomes (LEs), also known as multivesicular bodies (MVBs). In the MVBs, cargo destined for degradation is incorporated into intraluminal vesicles (ILVs). The following fusion of this compartment with the lysosomes - which host proteolytic enzymes - results in cargo degradation. (Adapted from Sorkin and von Zastrow 2009)

Early endosomes (EEs) are temporally and spatially the first station on the endocytic pathway, and they are often located at the periphery of the cell, receiving most types of vesicles from the surface. The EEs represent the major sorting station responsible for:

- ensuring that receptors that need to be recycled to the PM, reach the recycling endosomes;
- sorting material towards the trans-Golgi network (TGN);
- shuttling receptors and internalized materials, which require downregulation/degradation, to the late endosomes (LEs).

The EE environment is mildly acidic causing the dissociation of certain ligands from their receptor (not in the case of the EGFR). Moreover, the structure of EEs directly relates to their function. EEs present a tubulovesicular morphology with vesicles up to 1 μm in diameter. Cargoes destined for the recycling pathway or the TGN are localized in the tubule portion of the EE, through which they can be recycled back to the PM. In contrast, cargoes destined for degradation are localized in the vesicular portion of the EE, through which they proceed to the LE compartment (**Fig. 13**).

The LE compartment receives material from EEs, from the TGN on the biosynthetic pathway, and from phagosomes on the phagocytic pathway, and are characterized by low levels of endocytosed recycling proteins and high levels of endocytosed proteins destined for degradation. LEs are mainly spherical, although not entirely lacking tubules. Inside the late endosomes, molecules are also sorted into smaller vesicles that bud from the perimeter membrane into the endosomal lumen, forming intraluminal vesicles, where cargoes destined for degradation are entrapped. This leads to the multivesicular appearance of LEs, and so they are also known as multivesicular bodies (MVBs). Mature LE/MVBs fuse with lysosomes, pouring vesicles present in the endosomal lumen directly into the lysosomal lumen.

Lysosomes are the final compartment of the endocytic pathway, and they are considered the main hydrolytic and degradative compartment of the cell. Their principal role is to break down discarded cellular products and to return new building material to the cytoplasm.

Lysosomes have a high content of lysosomal hydrolases provided from the Golgi apparatus via LEs, and their internal environment is acidic - pH is around 4.8. By electron microscopy, lysosomes are characterized as an electron-dense material and they appear as large vacuoles (1-2 μm in diameter), where proteins are efficiently degraded.

The decision to target cargoes to recycling or degradation is critical for cell homeostasis and viability, and thus determinants for cargo fate needed to be identified. Ubiquitination was found to be the key signal directing cargoes to the degradative pathway (Raiborg and Stenmark 2009; Wollert, Yang et al. 2009), via the action of several proteins harboring UBDs that recognize ubiquitinated cargoes and lead them along the degradative route. Ubiquitinated cargoes first bind to the UIM-containing protein HRS and then to the endosomal sorting complexes required for transport (ESCRTs) in the endosomal station, which sort them to the luminal LE vesicles that are then delivered to lysosomes. The ESCRT multiprotein complexes play a key role in MVB sorting of ubiquitinated receptors, and they have two important features: first they recognize ubiquitinated cargoes, preventing their recycling and retrograde trafficking; then they drive the sorting of the cargo in endosomal membrane invaginations, forming the intraluminal vesicles that will later fuse with lysosome.

4.5 - Role of Ub in endocytosis

Considering the plasticity of Ub signals, it is not surprising that in recent years ubiquitination has emerged as a fundamental player in the endocytosis and downmodulation of receptors (Dikic 2003; Haglund, Sigismund et al. 2003), thus, having also implications on signaling. Many signaling receptors were found to be modified by the attachment of one or more Ub moieties, including RTKs, GPCRs and others (Hicke and Riezman 1996).

Ub-mediated endocytosis was first identified in studies using the yeast *Saccharomyces cerevisiae* that demonstrated that internalization of several PM cargoes required Ub (Kolling and Hollenberg 1994; Hicke and Riezman 1996; Galan and Haguenauer-Tsapis 1997). Indeed,

monoubiquitination of several yeast membrane receptors is necessary and sufficient for endocytosis, causing their internalization and degradation in the vacuole (Hicke and Riezman 1996).

In mammalian cells, the role of Ub in endocytosis is somewhat more complicated. In the case of ion channels, ubiquitination is essential for their internalization (Sorkin and von Zastrow 2009). Instead, for what concern many endocytic cargoes - including RTKs and GPCRs - ubiquitination appears to be sufficient for endocytic uptake (Haglund, Sigismund et al. 2003; Sigismund, Woelk et al. 2005). Importantly, although many of these cargoes exhibit ligand-dependent Ub modification, they can also undergo Ub-independent endocytosis. Thus, in mammalian cells, ubiquitination is often sufficient - but not required - for internalization, and the involvement of Ub depends on the type of receptor and the internalization route it takes. However, as observed in yeast, ubiquitination is required for receptor degradation in a ligand-induced Ub-mediated sorting mechanism, where Ub-receptors specifically recognize the ubiquitinated cargo through their UBDs. This inexorably destines receptors to lysosomal degradation and drives them away from a recycling pathway.

5 - Endocytosis of EGFR

When no external stimuli are supplied, the EGFR at steady-state is mainly localized at the PM. Binding of EGF to an inactive EGFR results in the dimerization of two EGFR monomers and subsequent receptor activation (see Section 2.1), leading to the phosphorylation of Tyr residues that can function as docking sites for signaling molecules, as well as for proteins involved in internalization pathways. Depending on the EGF dose, the EGFR can be internalized through the CME pathway, or via the NCE pathway, with distinct signaling outcomes: CME is crucial to sustain EGFR signaling, while NCE is essential for EGFR degradation in the presence of high EGF doses, safeguarding cells against excessive signaling (Sigismund, Algisi et al. 2013). The existence of two distinct EGFR internalization routes raises the question: *how does the cell sense the signal and target the EGFR to the correct internalization pathway?*

Of note, the switch between CME and NCE occurs over a narrow range of EGF concentrations (1-10 ng/ml), and it has been proposed that this fine modulation is transmitted to the cell in a pY-based cooperativity mechanism (Sigismund, Algisi et al. 2013). The molecular determinants for distinguishing between the CME and NCE internalization pathways have been elucidated by our group in a study that demonstrated that the EGFR ubiquitination pattern is responsible for committing the receptor to the different internalization routes (Sigismund, Woelk et al. 2005). At low doses of EGF, the receptor is not ubiquitinated and is internalized through CME, whereas at higher ligand concentrations the receptor becomes ubiquitinated and, concomitantly, the NCE internalization pathway is activated (**Fig. 14**). At this point, the receptor is internalized both through CME and NCE, with a ratio that is about 60 and 40% respectively (Sigismund, Argenzio et al. 2008). Both CME and NCE direct the EGFR to EEs, where its fate is decided: either it is recycled back to

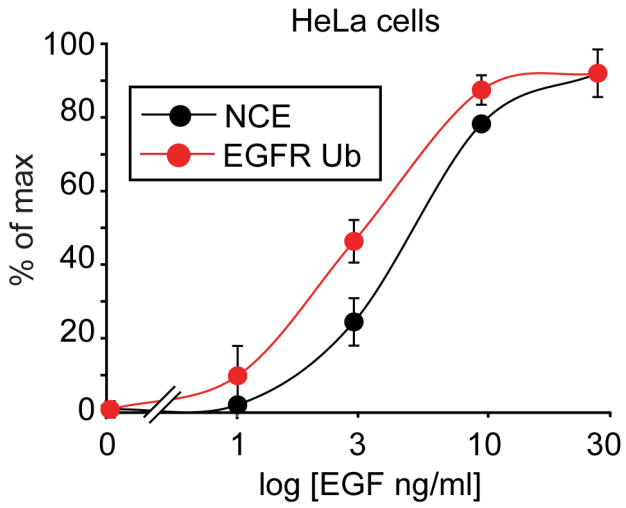


Figure 14: **EGFR ubiquitination and EGFR-NCE activation occur in the same EGF concentration interval.** Dose-response curves of EGFR-NCE (black) and EGFR ubiquitination (red) in HeLa cells. Cells were stimulated with increasing doses of EGF and ubiquitination was measured. NCE activation was measured with an iodinated internalization assay at the indicated EGF concentrations. (From Sigismund, Algisi et al. 2013)

the PM or it is destined to degradation via the LEs and lysosomes (**Fig. 15**). Thus, distinguishing between the two endocytic pathways is essential for understanding EGFR biology, since both pathways have profound and distinct effects on receptor fate, and hence on downstream signaling.

5.1 - CME sustains EGFR recycling to the cell surface and prolonged EGFR signaling

The first EGFR internalization route to be characterized was the CME, through which the receptor is rapidly removed from the cell surface and internalized via clathrin-coated pits. The phosphorylated receptor is recognized by CME adaptors that recruit the clathrin apparatus and drive pit formation (**Fig. 15**, see also Introduction Section 4.2). In particular, a direct interaction between the phosphorylated EGFR and the μ2 subunit of the adaptor protein AP2 has been observed (Sorkin, Mazzotti et al. 1996). Apart from AP2, other adaptors play a role in EGFR endocytosis, namely EPS15/EPS15L1 and epsins (Carbone, Fre et al. 1997; Sigismund, Woelk et al. 2005), which, however, also have a role in EGFR-NCE. It has also been found that Grb2 and Cbl are necessary for CME of the EGFR (Jiang, Huang et al. 2003; Stang, Blystad et al. 2004). Since Cbl-mediated ubiquitination of the EGFR was not necessary

for its internalization via CME (Huang, Goh et al. 2007), Cbl and Grb2 may be involved in CME via a distinct mechanism independent of EGFR ubiquitination.

In the presence of low EGF concentrations, EGFR is internalized almost exclusively through the CME route, and preferentially destined to recycling to the PM (around 70%), although there is a minor component that is directed to degradation (around 30%) (Sigismund 2008). CME-dependent recycling could represent an important aspect of EGFR signaling, conferring to the receptor spatial and temporal specificity; CME represents a way to redirect the receptor to precise membrane locations where a specific signal is required. Continuous

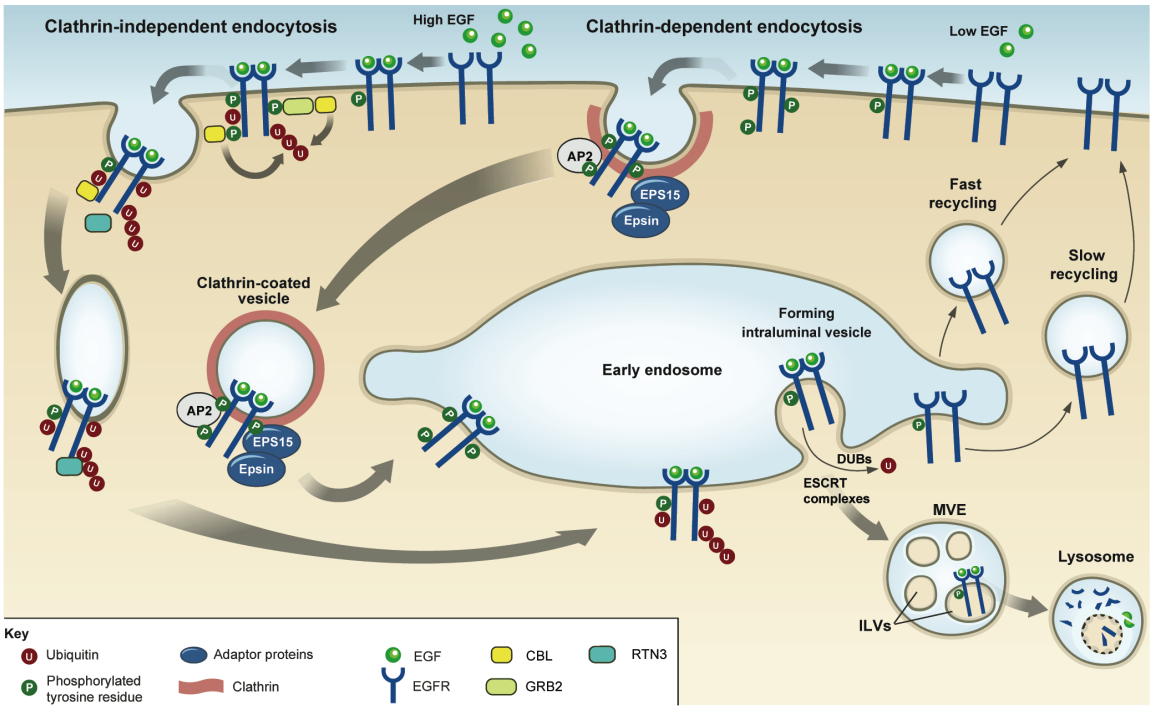


Figure 15: EGFR endocytosis and endosomal sorting. Following ligand binding, EGFR can undergo clathrin-dependent endocytosis (CME, right) or clathrin-independent endocytosis (NCE, left). In both cases, endocytic vesicles derived from both clathrin-dependent and -independent pathways fuse with early endosomes (EEs). From here, ubiquitinated receptors are sorted into late endosomes (LEs) and intraluminal vesicles (ILVs) of multivesicular endosomes (MVEs) and subsequently targeted for lysosomal degradation. In contrast, non-ubiquitinated receptors are recycled to the plasma membrane. Ubiquitin is an essential signal for endosomal sorting of EGFRs into the ILVs of MVEs. Components of the endocytic machinery, including Eps15, epsins, and the ESCRT components, contain ubiquitin-binding domains (UBDs) that have been implicated in recognizing and sorting ubiquitinated receptors either at the plasma membrane or at endosomes. The ESCRT-I and -II components (not shown) assemble in super-complexes and have been proposed to organize buds at the endosomal membrane. The ESCRT-III (not shown) complex associates with ESCRT-II and forms polymers that drive membrane scission and ILV biogenesis. Deubiquitinating enzymes (DUBs) catalyze the removal of ubiquitin from receptors before their translocation into the ILV, without allowing cargo to escape. (Adapted from Haglund and Dikic 2012)

recycling would also allow for multiple cycles of signaling and could also protect the receptor from degradation in conditions of limited ligand availability. Indeed, recycling of the receptor to the PM is a way to achieve sustained signaling.

CME was also found to be necessary for downstream EGFR signaling: in fact, some downstream signaling pathways, such as AKT and ERK, require intact CME machinery (Sigismund, Argenzio et al. 2008; McMahon and Boucrot 2011). In addition, the execution of complex biological functions, such as EGFR-activated DNA synthesis, depends on functional CME (Sigismund, Argenzio et al. 2008). An interpretation of these findings could be that the initial phase of signaling, or the peak-phase, is endocytosis-independent and probably occurs at the PM, whereas the later decay-phase requires CME.

5.2 - NCE is preferentially coupled to EGFR degradation

Functional ablation of clathrin in HeLa cells has been shown to have little effect on EGFR degradation, implying that CME is not the major pathway responsible for receptor downmodulation and degradation (at least in the model system under scrutiny; Sigismund, Algisi et al. 2013). Indeed, an alternative cholesterol-dependent NCE internalization pathway for the EGFR has been detected in the same cells following stimulation with high EGF dose under CME-deficient conditions (i.e., lacking clathrin or AP-2; Sigismund, Argenzio et al. 2008). Inhibition of this pathway using cholesterol-interfering drugs such as filipin, showed that NCE directs around 90% of the EGFRs that are internalized through this pathway to degradation, thereby resulting in downmodulation of EGFR signaling.

Despite recent efforts, the EGFR-NCE pathway has not yet been fully elucidated and its existence is still questioned by some researchers in the field. In our laboratory, we have directed much effort towards the characterization of the EGFR-NCE pathway and the molecular mechanisms that govern it. EGFR-NCE that gets activated at high EGF doses and responsible for receptor degradation, is apparently distinct from other forms of clathrin

independent endocytosis, as evidenced by the peculiar dependence on dynamin 1/2 and Cdc42, but not other known genetic NCE determinants (Sigismund, Argenzio et al. 2008; Sigismund, Algisi et al. 2013; Caldieri, Barbieri et al. 2015). However, the lack of knowledge of the precise molecular determinants involved in EGFR-NCE has prevented in the past years, the use of molecular genetics tools to selectively inhibit this pathway, and thus hampered the study of the role of EGFR-NCE in cellular homeostasis.

To address this knowledge gap, our group has attempted to characterize the NCE molecular players using an integrated approach based on immunopurification of EGFR-NCE vesicles, SILAC-based mass spectrometry, and RNA interference (RNAi)-based validation screening (Caldieri, Barbieri et al. 2015). In this work, novel regulators of EGFR-NCE have been identified and validated. Among the regulators of the pathway, the ER-resident protein Reticulon-3 (RTN3) was identified, suggesting a role of the ER in EGFR internalization. This finding led to the discovery that EGFR-NCE represents a completely novel form of endocytosis that requires the formation of specific contacts between the PM and the cortical ER, mediated by RTN3 (**Fig. 16**). This novel NCE pathway has been termed ER-mediated endocytosis (**ERME**; Caldieri, Barbieri et al. 2015). By knocking down RTN3 expression in

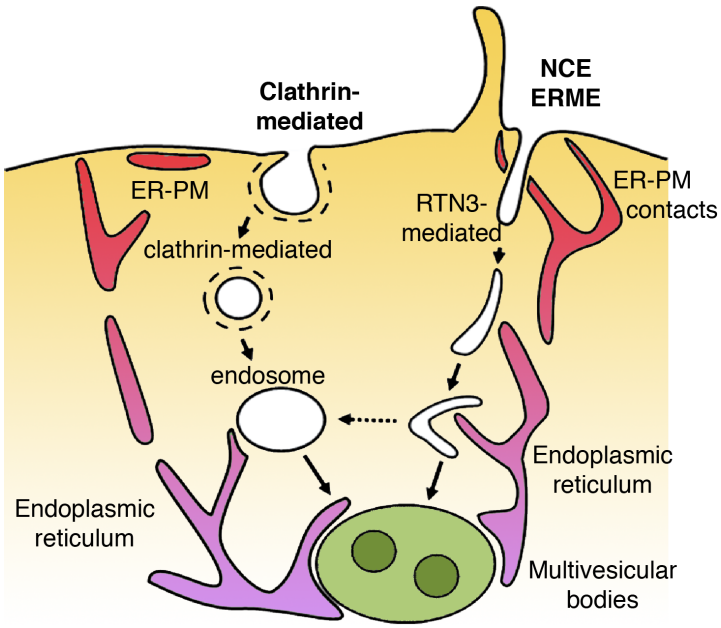


Figure 16: A novel model for EGFR endocytosis and the involvement of the ER. At high doses of EGF, EGFR is internalized both through CME and NCE. The proposed NCE pathway ERME is mediated by tubular structures that, differently from clathrin-coated pits, require contact sites between the plasma membrane and the endoplasmic reticulum (ER-PM) in order to progress. The establishment of these contact sites is Reticulon-3 (RTN3)-dependent. By inhibiting EGFR entry via NCE, RTN3 KD affects subsequent EGFR targeting to the lysosomal compartment and delays EGFR degradation. (Adapted from Caldieri, Barbieri et al. 2015)

HeLa cells, we confirmed that this specific NCE pathway is the major endocytic pathway responsible for EGFR degradation and long-term attenuation of EGFR signaling. Thus, this pathway might serve as a crucial negative regulator of EGFR signaling in response to excessive stimulus. It's worthy to note that hereafter we refer to NCE as the specific EGFR clathrin independent, RTN3 dependent pathway described in this section.

5.2.1 - EGFR ubiquitination and EGFR-NCE are mechanistically linked

As stated above, the switch between CME and NCE occurs over a narrow range on EGF concentrations (1-10 ng/mL). Notably, EGFR ubiquitination also occurs in this EGF concentration range (Sigismund, Algisi et al. 2013). Indeed, the EGFR-Ub and the EGFR-NCE curves are almost superimposable (**Fig. 14**), indicating that EGFR ubiquitination correlates with NCE activation. This notion is supported by the finding that only the EGFR-NCE pathway displays a threshold-controlled dose-response curve over the EGF concentration range, while both EGFR-CME and EGFR-total internalization display completely different behaviors, not matching that of EGFR ubiquitination (Sigismund, Algisi et al. 2013). In addition, EGFR mutants that affect EGFR ubiquitination also affect the capacity of the mutants to be internalized through NCE (Sigismund, Algisi et al. 2013). The fact that EGFR ubiquitination and NCE activation only occur at high EGF concentrations suggests that this pathway serves to prevent overstimulation of the cell in the presence of excess ligand by directing receptor to degradation.

6 - EGFR and other RTKs in cancer

In general, RTKs have been shown not only to be key regulators of normal cellular processes, but also to have a critical role in the development and progression of many types of cancer (Zwick, Bange et al. 2001). The first connection between RTKs and human cancers was made in 1984 (Ullrich, Coussens et al. 1984), when the predicted viral-*erbB (v-erbB)* mRNA oncogene product, which was identified in chickens infected with avian erythroblastosis virus, was found to be homologous to the amplified EGFR gene in human A431 epidermal carcinoma cell line. This discovery connected, for the first time, an animal oncogene with a human gene that encodes a cell-growth-controlling membrane protein.

In the same study, the *EGFR* gene was found to be amplified and rearranged in the cell line under study, resulting in the expression of a receptor with a truncated extracellular binding domain (Ullrich, Coussens et al. 1984). Since this discovery, the link between RTKs and cancer became increasingly evident and now it is well known that aberrant RTK signaling is critically involved in human cancer and other hyperproliferative diseases (Gschwind, Fischer et al. 2004, see also Section 6.1.1).

As with other types of oncogenes, oncogenic activation of RTKs can occur through several mechanisms, including gene amplification, protein overexpression, amino acid deletions and mutations. These alterations result either in the expression of a constitutively active form of the receptor or of a non-degradable receptor. Deletions and mutations can occur within the extracellular region, as well as in the catalytic domain or in the juxtamembrane region, resulting in a constitutive active RTK (for a recent review of EGFR mutations in lung cancer see Siegelin and Borczuk 2014). Additionally, mutations in the TM domain of the EGFR can also result in ligand-independent kinase activation, as reported for the other ErbB family member, ErbB2 (HER2; Bargmann, Hung et al. 1986). Somatic and germline mutations that

are associated with distinct inherited and spontaneous human cancer syndromes, have been observed in at least ten different RTK families (Robertson, Tynan et al. 2000).

In many cases, alterations affect the ability of the RTK to be properly ubiquitinated and downregulated, thereby causing sustained signaling, as in the case of the EGFR and the MET receptor (Kon, Kobayashi et al. 2014). The most frequent genetic alterations, in these cases, consist of mutations that affect the region encoding the intracellular domain of RTKs, usually encompassing the binding region for Cbl (as in the case of MET receptor), the major E3 ligase involved in RTK ubiquitination (Sigismund, Confalonieri et al. 2012). An alternative mechanism for preventing receptor downmodulation is proposed to be the occlusion of relevant binding sites on the EGFR, for instance by heterodimerization with other ErbB family members that renders the EGFR inaccessible to effectors, such as Cbl (Shtiegman, Kochupurakkal et al. 2007, see also page 60).

Another way by which RTK functionality can be deregulated is to establish an autocrine growth factor loop, as frequently described for the EGFR and the IGF-IR (Smith, Derynck et al. 1987; Kaleko, Rutter et al. 1990; Saeki, Salomon et al. 1995). This mechanism of oncogenic transformation occurs when an RTK is aberrantly expressed or overexpressed in the presence of its cognate ligand, or when overexpression of the ligand occurs in the presence of its associated receptor. In many solid tumors it has been shown that elevated levels of both growth factor receptor and its ligand are expressed concomitantly (Salomon, Normanno et al. 1995; Scala, Saeki et al. 1995).

6.1 - ErbB family receptors and cancer

Over the past two decades, it has become evident that the ErbB family members play an important part in the initiation and maintenance of several solid tumors. This has led the scientific community to invest much effort into the development and expansion of specific ErbB inhibitors for cancer therapies. In the following paragraphs, the role of ErbB receptors

in human cancer will be briefly discussed, whereas in Section 6.2 an overview of the possible intervention strategies and therapies targeting ErbB receptors will be given.

6.1.1 - EGFR in human tumors

Since the milestone discovery in 1984, when the sequence of EGFR was recognized to be closely related to the sequence of the *v-erbB* oncogene, many studies have been dedicated to characterizing EGFR deregulation in cancer. EGFR is now known to be deregulated in a number of tumors, including a variety of lung tumors, breast, prostate and ovarian carcinomas, gliomas and others (Arteaga and Engelman 2014). There are different ways through which EGFR can be involved in neoplastic transformation or in cancer progression, as outlined above for RTKs in general. These mechanisms include *EGFR* gene amplification or mutation, genetic rearrangements, and production of different spliced variants. The subsequent changes in EGFR expression are often correlated with poor prognosis in patients and usually determine increased activity of the EGFR, which is able to be activated both in a ligand-dependent and -independent manner.

Overexpression of EGFR is frequently found in non-small cell lung cancer (NSCLC), head and neck cancer, glioma, esophageal cancer, and colorectal cancers (Yarden and Pines 2012), as well as in bladder, cervical, ovarian, kidney, and pancreatic cancer (Zwick, Bange et al. 2001). One important mechanism leading to EGFR overexpression is *EGFR* gene amplification, with more than 15 copies per cell reported in certain tumors (Velu 1990). Moreover, mutations in regulatory regions of the gene (e.g. the promoter) can be responsible for EGFR increased expression levels (Hudson, Thompson et al. 1990; Hou, Johnson et al. 1994). In general, elevated levels of EGFR expression are associated with the late stages of disease progression and often correlate with high metastatic rate and increased tumor proliferation (Nicholson, Gee et al. 2001; Selvaggi, Novello et al. 2004; Zimmermann, Zouhair et al. 2006).

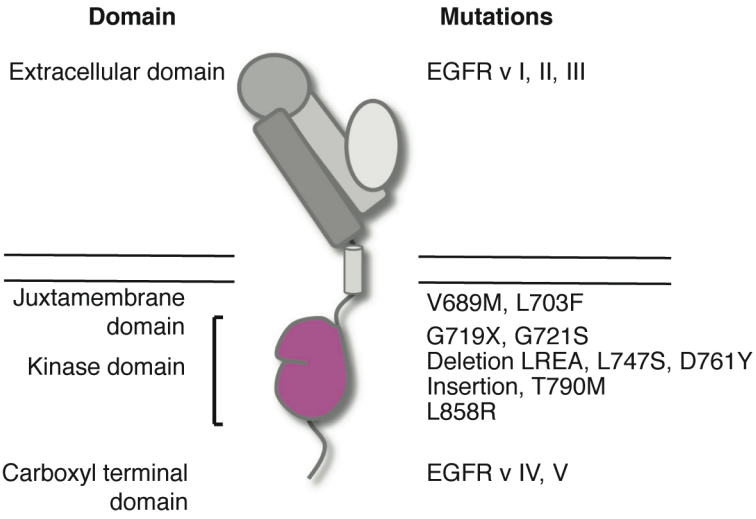


Figure 17: Oncogenic EGFR variants. Cartoon shows the positions of key EGFR mutations/variants in the corresponding domains. Large rearrangements of the extracellular domain of the EGFR, as well as single base mutations in the juxtamembrane or intracellular domains, can cause deregulated EGFR-dependent signaling activity.EGFR vI, EGFR vII, and EGFR vIII represent the three major types of deletion mutations in the extracellular region. Of these mutant forms, EGFR vIII is the most common mutation in gliomas. V689M and L703F mutations are localized in the EGFR juxtamembrane (Huang, Nijman et al. 2006) region and were identified as constitutively active in NSCLC patients. JM seems to play an activating role, enhancing the formation of the asymmetric dimer and promoting allosteric activation of the acceptor kinase.
The most prevalent EGFR mutations occur in the kinase domain of the receptor: the small in-frame deletion around the conserved LREA motif of exon 19 and the L858R point mutation in exon 21, account for more than 90% of all EGFR kinase mutations. Point mutations G719 account for approximately another 5% of EGFR mutations. These EGFR mutations activate the EGFR signaling pathway and promote EGFR-mediated pro-survival and anti-apoptotic signals through down-stream targets. In contrast to the activating mutations, a point mutation T790M was identified as a secondary mutation responsible for acquired resistance to EGFR inhibitors, accounting for more than 50% of primarily EGFR TKI- sensitive lung tumors that become resistant to EGFR inhibitors. Other resistance mutations, such as D761Y and L747S, have been reported but seem to be rare. Finally, two kinds of deletion mutants in the C-terminal domain have been reported in gliomas: EGFR vIV and vV. These mutants seem to be constitutively active. (Adapted from Zhang, Stiegler et al. 2010)

Mutations affect different mechanisms of EGFR regulation (**Fig. 17** and **Table III**), leading to enhanced EGFR activity. These mutations can be single base mutations or larger genetic rearrangements. In gliomas, for example, EGFR amplification is often accompanied by structural rearrangements that cause in-frame deletions in the extracellular domain of the receptor. This is the case for an oncogenic mutation that deletes exons 2–7 in the receptor ectodomain, denoted EGFR variant III (EGFRvIII), that is found in up to 60% of glioblastomas (Nishikawa, Ji et al. 1994) and in other types of cancer, including lung, breast, ovarian and prostate cancer (**Fig. 17**) (Wikstrand, Hale et al. 1995; Gan, Cvrljevic et al. 2013). EGFRvIII exhibits constitutive dimerization, impaired downregulation, and aberrant tyrosine kinase activity, all resulting in enhanced tumorigenicity (Nishikawa, Ji et al. 1994). EGFRvIII

is not the only variant resulting from a deletion mutation in the extracellular region; two other major variants exist (EGFRvI, EGFRvII, originally discovered in gliomas). However, EGFRvIII remains the most common mutation in gliomas (Zhang, Stiegler et al. 2010).

Somatic activating mutations in EGFR have been discovered in the last 10 years in a subset of NSCLCs (**Fig. 17**) (Lynch, Bell et al. 2004; Paez, Janne et al. 2004; Pao, Miller et al. 2004). Like the EGFRvIII mutation, these mutations enhance EGFR activity, causing ligand-independent firing and/or boosting receptor-dependent signaling up to 50-fold over the basal unliganded receptor activity (Yun, Boggon et al. 2007). One kind of somatic mutation that boosts EGFR-dependent signaling acts through impairment of Cbl binding. This mutation is represented by the EGFR L834R (or L858R, accordingly to whether the numbering system includes the signal peptide) mutant (**Fig. 17**) (Shtiegman, Kochupurakkal et al. 2007) and was identified in NSCLC patients. Although the L834R mutation does not directly disrupt the Cbl binding site on the receptor (pY1045), which shows a normal - or even increased - phosphorylation, it causes reduced Cbl recruitment to the EGFR and defective EGFR ubiquitination (Shtiegman, Kochupurakkal et al. 2007). The authors proposed that this reduction in Cbl recruitment could be due to occlusion of its binding site through the formation of a heterodimer between L834R-EGFR and ErbB2, which appears to form even in the absence of ligand stimulation (for L834R mutant see also Discussion Section 7, page 163). Impairment of Cbl-mediated ubiquitination of the EGFR causes persistent activation of downstream signaling molecules, including Ras, MEK and ERK signaling effectors (Kon, Kobayashi et al. 2014).

Another very potent mechanism for constitutive EGFR activation in a variety of human cancers is the autocrine/paracrine stimulation via growth factor loops. The most prominent ligand, which is involved in autocrine growth receptor activation is TGFα, whose coexpression with the EGFR is frequently observed in glioblastomas and squamous cell

carcinomas of the head and neck, as well as in lung, pancreas, ovarian and colon cancers (Grandis, Chakraborty et al. 1998; Yarden 2001).

The abovementioned mechanisms are all frequently associated with different types of cancer, and the identification and characterization of these non-physiological alterations is fundamental to the development of an effective anti-cancer therapy and to discover new druggable targets. Even though this thesis is focused on the study of the EGFR, it is interesting to highlight how the ability of the EGFR to heterodimerize with the other ErbB receptors, in particular ErbB2, plays an important role in cancer.

6.1.2 – ErbB2/HER2

The involvement of ErbB2 in cancer was first inferred from a study on its rat orthologous Neu (Schechter, Stern et al. 1984). ErbB2 is typically amplified in human cancers, such as breast,

ErbB receptor	Alteration	Cancer type	References
EGFR	Mutation (L834R, etc.)	Lung NSCLS	(Lynch, Bell et al. 2004; Paez, Janne et al. 2004; Pao, Miller et al. 2004)
EGFR	EGFR-vIII	Glioma	(Sugawa, Ekstrand et al. 1990)
EGFR	Amplification	Lung NSCLC, head and neck, glioma, esophageal, colorectal	(Yarden and Pines 2012)
ErbB2 (HER2)	Amplification	Breast, gastric, esophageal	Cancer genome atlas network 2012
ErbB2 (HER2)	Mutation	Breast, lung, gastric, bladder, endometrial	Cancer genome atlas network 2012
ErbB3 (HER3)	Mutation	Breast, gastric	(Jaiswal, Kljavin et al. 2013)
ErbB4 (HER4)	Mutation	Melanoma, lung NSCLC, medulloblastoma	(Prickett, Agrawal et al. 2009; Gilbertson, Hernan et al. 2001)
TGFα	Overexpression	Prostate, lung, pancreas, ovary, colon, head and neck	(Yarden and Sliwkowski 2001)
Neuregulin1	Overexpression	colorectal, head and neck	(Wilson, Lee et al. 2011; Yonesaka, Zejnullahu et al. 2011)

Table III: Alterations of ErbB Receptors and Ligands in Human Cancer. (Adapted from Arteaga and Engelman 2014)

ovarian, gastric and esophageal tumors (**Table III**; Arteaga and Engelman 2014). In approximately, 30% of breast cancers ErbB2 was found to be amplified 2- to 20-fold and was associated with poor prognosis (Slamon, Clark et al. 1987). This was the first time that an oncogenic alteration had been associated with poor prognosis in cancer patients, suggesting a causal relationship to cancer. Additional evidence linking ErbB2 with cancer progression comes from the improvement in survival of patients with ErbB2-amplified early-stage breast cancer treated with the ErbB2 antibody Trastuzumab (Gajria and Chandarlapaty 2011).

As previously discussed, ErbB2 is an orphan receptor, meaning no ligands have been identified that bind to and activate the receptor. The mechanism of activation of ErbB2 is indeed peculiar: it acts as a co-receptor for the other EGFR family members (**Fig. 4**, Introduction Section 2.4). The mechanism underlying the oncogenic potential of overexpressed ErbB2 may relate to the increased potential for heterodimer formation. In particular, the specific EGFR ligands or the neuregulins (that bind ErbB3 or ErbB4) are able to induce heterodimer formation between a high-affinity ligand binding receptor and ErbB2. ErbB2 possesses the strongest tyrosine kinase activity among the ErbB receptors (Moasser 2007), and the heterodimers that is forms are pathophysiologically relevant since a plethora of signaling pathways, including MAPK, PI3K, Ras and Src, are activated by such heterodimers and appear to be important for tumor growth (Prenzel, Zwick et al. 2000; Arteaga and Engelman 2014).

ErbB2-containing heterodimers are also characterized by a higher affinity for ligands with respect to the other heterodimeric receptor complexes, owing to slow rates of growth factor dissociation (Aria-Romero 2010). Moreover, ErbB2-containing heterodimers undergo slow endocytosis, and are more frequently recycled back to the cell surface (Sorkin, Di Fiore et al. 1993). These features contribute to the generation of potent mitogenic signals from ErbB2-containing heterodimers, through the simultaneous and prolonged activation of multiple signaling pathways.

6.2 - Possible drugs targets and therapy

In the last decades, the increasing knowledge of RTK structure and the activation mechanisms, as well as of the signaling pathways they control, allowed the development of target-specific drugs and new anti-cancer therapies. Approaches towards the prevention or interference of deregulated RTK signaling include the development of selective compounds that target either the extracellular ligand-binding domain or the intracellular tyrosine kinase domain or substrate-binding region (**Fig. 18**).

6.2.1 - Monoclonal antibodies

One of the most successful classes of targeted therapies designed to selectively kill tumor cells are monoclonal antibodies (mAbs) directed against the extracellular domain of RTKs. At the end of last century, recombinant antibody technology made immense progress in the design, selection and production of newly engineered antibodies (Farah, Clinchy et al. 1998; Hudson 1999). This research field has since been accelerated by the generation of humanized antibodies for targeted cancer therapies, which are now widely used in the clinic (Nelson,

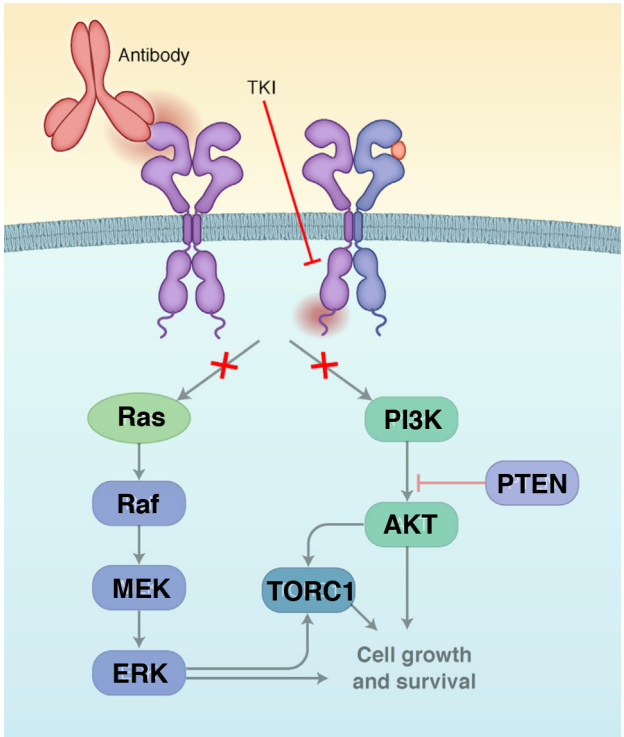


Figure 18: Schematic representation of EGFR and ErbB2 inhibitors. In the cartoon, EGFR or ErbB2-addicted cancer cells are treated with an ErbB small-molecule inhibitor (tyrosine kinase inhibitor; TKI) or monoclonal antibody, resulting in suppression of downstream signaling. EGFR homodimers (left) and EGFR-ErbB2 heterodimers (right) are shown. Adapted from Arteaga and Engelman 2014)

Dhimolea et al. 2010).

Anti-RTK mAbs work by blocking the ligand:receptor interaction, and therefore inhibiting ligand-induced RTK downstream signaling and/or increasing receptor internalization and downregulation. Alteration of the intracellular signaling pattern inside the targeted tumor cell may lead to growth inhibition and/or apoptosis (Zwick, Bange et al. 2001). The most widely used EGFR-neutralizing mAbs are Cetuximab and Panitumumab (**Table IV**). These antibodies specifically bind to the extracellular domain of the EGFR, thereby preventing ligand binding (Bou-Assaly and Mukherji 2010; (Dubois and Cohen 2009), and are most effective in cancers that express ligand-activated, wild-type EGFR. They have also been proposed to enhance EGFR endocytosis and downregulation (Doody, Wang et al. 2007); however, the exact mechanism of their action is still under investigation.

Cetuximab, commercialized as Erbitux, is an approved treatment for head and neck squamous cell carcinomas and colorectal cancer as indicated by the information provided by the company that produced it (MerkSerono, www.merckserono.com). It has also been

Drug	Type of molecule	Mechanism of action	Notes
Cetuximab	human-murine chimeric IgG_2, binds ligand-binding domain	inhibits ligand-dependent activation of EGFR	
Panitumumab	human IgG_1, binds ligand-binding domain	inhibits ligand-dependent activation of EGFR	
Lapatinib	small molecule	reversible, ATP-competitive TKI	
Erlotinib	small molecule	reversible, ATP-competitive TKI	
Afatinib	small molecule	irreversible, ATP-competitive TKI	
Neratinib	small molecule	irreversible, ATP-competitive TKI	Trials in patients with tumors expressing ErbB2 mutants are ongoing.

Table IV: ErbB receptor inhibitors. Mechanisms of action and key clinical trials for the principal clinical-grade ErbB targeted therapies. (Adapted from Arteaga and Engelman 2014)

evaluated as a first-line treatment for advanced NSCLC in clinical trials, with the result that combination of chemotherapy plus Cetuximab improves overall survival in advanced NSCLC patients with respect to chemotherapy alone (Yang, Liu et al. 2014). Panitumumab is commercialized as Vectibix and to date is approved for the treatment of patients with metastatic colorectal cancer (**Table IV**; Arteaga and Engelman 2014).

In the case of ErbB2, the best-characterized and most widely used blocking antibody is Trastuzumab, approved for the use in breast cancers since early 2000s (Slamon and Pegram 2001) and more recently used also for HER2-positive advanced gastric or gastro-esophageal junction cancer (Bang, Van Cutsem et al. 2010).

6.2.2 - Tyrosine kinase inhibitors (TKIs)

Another effective approach to inhibiting aberrant RTK signaling is to use small molecule inhibitors that selectively interfere with the intrinsic tyrosine kinase activity of the EGFR. Kinase inhibition is achieved by two alternative mechanisms: by competing with ATP for binding to the catalytic site or by blocking the catalytic activity of the receptor. These mechanisms block receptor autophosphorylation and thus activation of downstream signal transducers. Two EGFR-targeting reversible small molecule inhibitors, Gefitinib (commercialized as Iressa) and Erlotinib (commercialized as Tarceva) received fast-track approval from the US Food and Drug Administration (FDA) as treatment for patients with advanced NSCLC who had failed to respond to conventional chemotherapy (**Table IV**; Sharma, Bell et al. 2007). These molecules were designed to act as competitive inhibitors of ATP binding to the active site of the EGFR (Barker, Gibson et al. 2001; Wakeling, Guy et al. 2002), so that receptors are inhibited at lower concentrations of the drug resulting in a favorable therapeutic index. Early clinical data showed that 10% of patients with NSCLC responded to Gefitinib or Erlotinib, followed by the observation that sensitivity to Gefitinib and Erlotinib strongly correlated with the presence of a new class of somatic activating

mutations in the EGFR kinase domain (Lynch, Bell et al. 2004; Paez, Janne et al. 2004; Pao, Miller et al. 2004), including the L834R substitution (see page 60). This finding can be explained by the fact that EGFR kinase domain mutations result in hyper-activation of the kinase, so conferring a dependence of the tumor cells on the mutated kinase. The treatment of TKI-sensitive cells with these targeted therapies appears to cause a form of 'oncogenic shock', which is proposed to result from the gradual degeneration of downstream mitogenic signals, leading to a temporary predominance of apoptotic signals (Sharma, Bell et al. 2007).

EGFR lung cancers carrying the primary activating mutations often develop a secondary mutation T790M that confers resistance to the TKI. Following the identification of this secondary mutation (Pao, Miller et al. 2005) there was a huge effort to develop an inhibitor of the T790M EGFR mutant and thus overcome resistance. This effort led to the development of second-generation EGFR inhibitors, such as Afatinib (**Table IV**). Although these second-generation drugs are able to inhibit the T790M EGFR mutant, they do so at concentrations that also inhibit wild-type EGFR. Thus, treatment with these inhibitors induces severe side-effects, such as acneiform-like rash and diarrhea (Hirsh 2011).

Another approach to inhibition of ErbB kinases has been to develop irreversible inhibitors (reviewed in Sharma, Bell et al. 2007). These inhibitors are able to form a covalent bond with a crucial Cys residue in the active site of the enzyme: Cys797 in EGFR, and Cys805 in ErbB2 (Singh, Dobrusin et al. 1997; Fry, Bridges et al. 1998). Since only EGFR and ErbB2, among the ErbB receptors, have cysteine residues at these corresponding positions, these irreversible inhibitors show very high specificity towards these two enzymes and are thus called dual inhibitors for their dual specificity. Interestingly, resistance to irreversible dual inhibitors does not develop as quickly as it does with reversible inhibitors such as Gefitinib and Erlotinib. Indeed, the irreversible EGFR inhibitors CL-387785 (Discafani, Carroll et al. 1999) and Canertinib (Smaill, Rewcastle et al. 2000) can overcome resistance to L834R-mutated EGFR harboring the T790M mutation, whereas the reversible EGFR and ErbB2 inhibitor Lapatinib

was ineffective in this regard (Carter, Wodicka et al. 2005; Kobayashi, Ji et al. 2005). Similarly, a small subset of NSCLCs that express the EGFRvIII (see Introduction Section 6.1.1) are also insensitive to Gefitinib and Erlotinib, but show sensitivity to the irreversible TKI Neratinib (Ji, Zhao et al. 2006) that is currently in phase I/II clinical trials for NSCLC patients (**Table IV**). Therefore, several independent lines of evidence underscore the use of irreversible ErbB inhibitors, especially for situations in which reversible inhibitors of EGFR lose efficacy (Sharma, Bell et al. 2007).

During the development of this project, we took advantage of Getifinib, the competitive reversible inhibitor of ATP-binding, to study the role of phosphatases in the early steps of EGFR activation upon kinase inhibition (see Results Section 4).

7 - Modeling biological networks

As mentioned earlier, one big challenge in studying complex signaling pathways is the step of deciphering the exact mechanisms that underlie the translation of an extracellular signal to a specific cellular outcome. In particular, when dealing with RTKs, the situation becomes more complex, since pathways that rely on the activation of RTKs are highly complex and interconnected. Cell biology and biochemistry provide tools to study local and short-range interactions within biological networks, but are subjected to technical limitations and not suitable to study long-range connections. Computational modeling aims to bridge this gap by revealing hidden functional aspects and topological properties (Enuka, Feldman et al. 2015).

Mass action modeling, also known as mechanistic computational modeling, is the main approach used to model layers of the ErbB network. Mass action modeling is based on ordinary differential equations (ODEs) through which it is possible to express the overall rate of change in the concentrations of cellular components. The rate equations that form the model are connected; this means that a change in the concentration of one component causes a change in another component.

The dynamics of the signaling cascade is modeled by an ordered set of rate equations. Rate equations generally have the form:

$$\frac{dC}{dt} = generation - consumption$$

which states that the change in concentration of a certain component over time is dependent upon the rate of generation minus the rate of consumption of the component. The generation and consumption terms can be either a constant or a function of the concentration of cellular components, such as mRNAs, proteins, and small molecules. Usually, the set of ODEs is too much complicated for an analytical solution. Therefore, mechanistic models are solved using

numerical methods that are based upon two parameters that must be either measured or estimated for every *species* in the model, namely, initial concentrations and rate constants. When experimental data for these parameters are missing, it is necessary to estimate their values. When parameters must be estimated, simulation techniques are used to ensure that the final output is robust against variations in these parameters.

The modeling process is a recursive procedure where a first model is built based on current knowledge. From this first version of the model, new hypotheses are derived. Later, these hypotheses are examined using available tools (e.g. wet lab experiments, analysis of data available in the literature, databases, etc.), and finally the model is validated or else rejected. In the case where the model is validated, the regulatory mechanisms that were forecast by the model are confirmed. In contrast, if the model is rejected, the model must undergo a refinement process.

7.1 - ErbB receptors: a multi-layered signaling network

Although the ErbB network possesses a relatively high level of complexity in terms of its players and interactions, it is however suited to the development of computational models. The ErbB network topology consists in three layers: (1) the input layer, (2) the signal processing layer and (3) the output layer.

The input layer is the part of the system where the ligand activates the receptor and, as previously discussed, it consists of several extracellular ligands, as well as 4 receptors, of which 3 can bind a ligand, whereas only one is an orphan receptor, able to heterodimerize with the other receptors (see Introduction Section 1). The different receptor-ligand complexes that are formed upon receiving a stimulus, drive distinct biological outcomes (Yarden and Sliwkowski 2001). The four receptors can assemble into several kinds of homo- and hetero-dimers and the situation becomes more complicated due to the fact that each receptor has at least nine tyrosine residues that can be phosphorylated. Moreover, each receptor displays

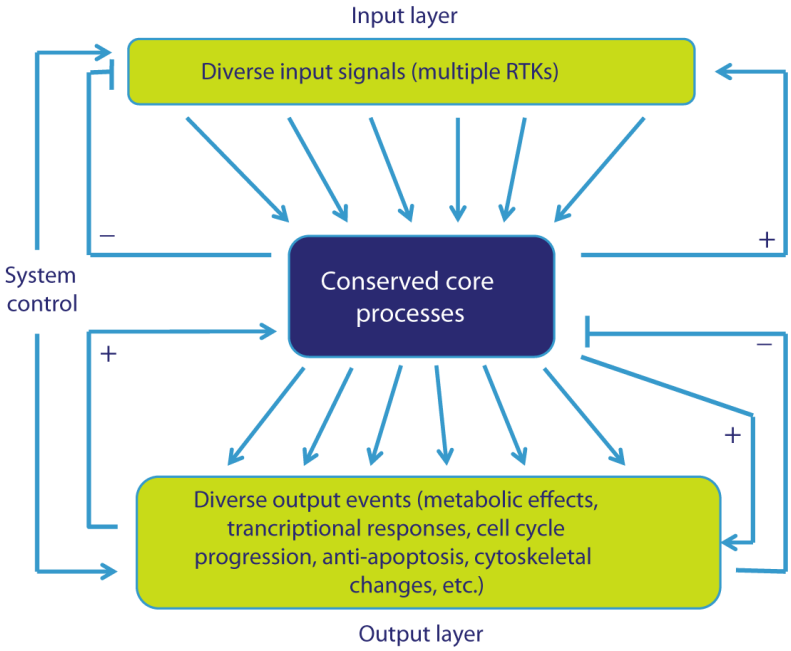


Figure 19: a "bow tie" or "hourglass" network. A useful concept that has emerged is that of a "bow tie" architecture of robust networks, linking the input layer of ligands and receptors to the output layer of biological outcomes, in which modularity and redundancy are important components. (Adapted from Meyts 2015)

distinct functional properties which depend on ligand affinity, increasing the complexity of this network.

The ErbB signal processing unit is composed of several kinase cascades. The first kinase in each cascade is able to phosphorylate, and thereby activate, the successive kinase in the cascade. Interestingly, although the processing unit is composed of components that are shared among several signaling pathways, the system is still able to signal through the correct pathway to produce the correct output. An example of this is the case of the EGF and the nerve growth factor (NGF). While the former activates ErbB receptors, the latter activates the TrkA receptor, but both signals are funneled to the MAPK module. However, while EGF transiently activates ERK, which is the last step in the enzymatic cascade of the MAPK module, NGF causes sustained activation of ERK, leading to profoundly different outputs. In particular, transient activation on ERK by EGF leads to proliferation, while sustained activation by NGF induces differentiation (Marshall 1995). A possible explanation for such a different output in a system that utilized the same players was given by the hypothesis of the

presence of different feedback loops governing this processing layer. Indeed, if cells were stimulated with EGF, the MAPK network exhibits a negative feedback loop between ERK and RAF (two players of the same network module), while with NGF stimulation, ERK positively signals backwards to RAF through PKC, thereby re-feeding the enzymatic cascade, resulting in a switch-like response of ERK activation (Birtwistle, Hatakeyama et al. 2007).

Villasenor *et al.* recently gave another possible mechanism by which cells can differentiate between EGF and NGF stimuli (Villasenor, Nonaka et al. 2015). Their study is based on the evidence that phosphorylated receptors are removed from the PM and packaged in endosomes. With the use of a mathematical model, they were able to determine that different endosomes contain the same mean amount of phosphorylated EGFR (pEGFR). By increasing EGF dose, cells respond to higher EGF concentrations by increasing the number of endosomes, keeping the average number of pEGFR almost constant in each endosome. This mechanism is referred to as an 'analogue-to-digital conversion' that ensures a robust signal and can regulate the signaling of the activated receptors in both space and time.

Moreover, the authors also found that changing how the pEGFR is distributed between endosomes modifies how the cells decipher the signal. Indeed, different growth factors either increase or decrease the number and size of endosomes in various cell types: this is the case of NGF or HGF in PC12 or primary mouse hepatoblasts, respectively (Villasenor, Nonaka et al. 2015). In NGF or HGF treated cells, a prolonged Erk signaling was observed and the authors proposed that this sustainment of signaling is due to the presence of endosomes that are smaller in size but more numerous. In contrast, in the same cell lines they observed that EGF determines a transient ERK activation, likely due to the presence of larger endosomes. The picture that they outlined is that the pEGFRs contained in smaller endosomes are protected from phosphatases and surrounded by a favorable signaling environment, whereas pEGFRs that reside in bigger endosomes are shorter-lived and dephosphorylated faster.

In the ErbB network, the kinetics of the processing unit downstream of the receptor is a major determinant of cell fate because the different network modules that can potentially exist (like positive or negative feedback loops) in the processing layer, determine a number of different signals, later transmitted to the output layer. The output layer receives information from the processing unit and directs specific transcriptional responses that determine cell fate. While much progress has been made in our understanding of the ErbB network, specifically at the input and signal processing unit layers, the output layer still lacks quantitative mechanistic models. Modeling of the output layer would require more detailed data than is currently available in terms of metabolic, biochemical, and transcriptional processes (Enuka, Feldman et al. 2015).

7.1.1 - Positive and negative feedback circuits

Autocrine and paracrine loops are important positive feedback circuits that act at the input layer, enhancing the ErbB network from the very first stimulus, i.e., ligand binding. In fact, activation of this network causes transcription of multiple ErbB ligands, such as TGFα and HB-EGF (Schulze, Lehmann et al. 2001; Brankatschk, Wichert et al. 2012), which then re-feed into the network. Another important feedback loop, in this case a negative circuit, is the endocytosis of activated receptors. Ligand-receptor complexes are indeed internalized through distinct mechanisms that overall result in either recycling or degradation in lysosomes.

At the border between the input layer and the signaling processing unit, phosphatases importantly impinge on EGFR activity and on that of the other ErbB receptors. Indeed, phosphoreceptors can undergo dephosphorylation both at the PM and at later intracellular compartments (Ostman and Bohmer 2001; Roda-Navarro and Bastiaens 2014), and are thus negatively regulated by this class of enzymes. In a work by Kleiman *et al.* it was highlighted how phosphorylated ErbB receptors are short-lived with half-lives of a few seconds (Kleiman, Maiwald et al. 2011). By exploiting mathematical modeling, combined with the use of

phosphatase inhibitor drugs, the authors observed that a rapid phosphoturnover occurred both on PM-resident receptors, as well as on multiple adaptor proteins and signaling kinases, such as AKT and MAPK. Thus, the complexes formed on the cytoplasmic tail of activated receptors and the downstream signaling cascades that they regulate, are highly dynamic and antagonized by potent phosphatases.

For what concerns the signaling processing unit, a major attenuator of signals within this layer is dephosphorylation by specific phosphatases. NRG-stimulated signaling offers an example of the complex role of phosphatases along this signaling pathway. NRG binds to ErbB4, stimulating signals that involve activation of the MAPK and AKT phosphorylation cascades. These two pathways are interconnected since activation of the AKT pathway decreases the level of ERK activation. The protein phosphatase 2A (PP2A) has contradicting effects on ERK signaling: it can negatively regulate the MAPK cascade (by dephosphorylation of MEK) and can simultaneously enhance the MAPK cascade by dephosphorylating AKT (Enuka, Feldman et al. 2015).

7.2 - Existing EGFR models

Computational models of the EGFR have been useful in the past to understand complex biological behaviors of receptor systems. At the end of last century, two complementary approaches were used in different attempts to model EGFR activity. Several groups focused their efforts on the ligand-induced endocytosis of EGFR (Sorkin, Waters et al. 1991; Wiley, Herbst et al. 1991; Reddy, Wells et al. 1998). These models mainly concentrated on receptor trafficking, without including mechanistic details of downstream signaling. This kind of approach did not consider qualitative differences in receptor signaling from the different cellular compartments, either the PM or endosomes.

In contrast, other models focused on the acute response that is characteristic of ligand-activated EGFR, jumping to the signaling processing layer without considering internalization

events. An example of such a modelling approach is the development of an EGFR network model to describe signal transduction events down to the Ras activation (Kholodenko, Demin et al. 1999). This model included a set of experimentally verified rate constants of the biochemical reactions that participate in signal transduction. Another example is given by a model by Kholodenko that presented a detailed outline of ERK activation cascades after EGFR activation (Kholodenko, Kiyatkin et al. 2002), but always disregarding the endocytosis component and the compartmentalization of the signaling receptors.

The need to implement both the input layers and the signaling processing unit into a single model, came at the beginning of this century. Understanding the lack of information that was produced by modeling only different layers, different groups proposed a series of models that were trying to contemplate more aspects of EGFR activation and signaling. In 2003, Wiley *et al.* proposed a model that attempted to give a comprehensive understanding of the EGFR signaling network, including both trafficking and signaling events (Resat, Ewald et al. 2003). By using a multicompartment model, the authors investigated the distribution of the receptors among cellular compartments, as well as their potential signal transduction characteristics (as schematized in **Fig. 20**).

The authors ended up with a model consisting of hundreds of distinct endocytic compartments and about 13,000 reactions/events that occur over a broad spatio-temporal range. Interestingly, in this model physiochemical aspects of ligand-receptor interactions, such as pH-dependent binding in different endosomal compartments were incorporated. To determine the utility of the model for biologically relevant predictions, they simulated and confirmed experimentally the differential activation of the EGFR by either EGF or TGFα ligands. It turned out that TGFα-stimulated receptors preferentially signal from the PM, whereas EGF-stimulated receptors produce a signal that is biased towards the endosomal compartment. The model was also able to predict the kinetics and magnitude of

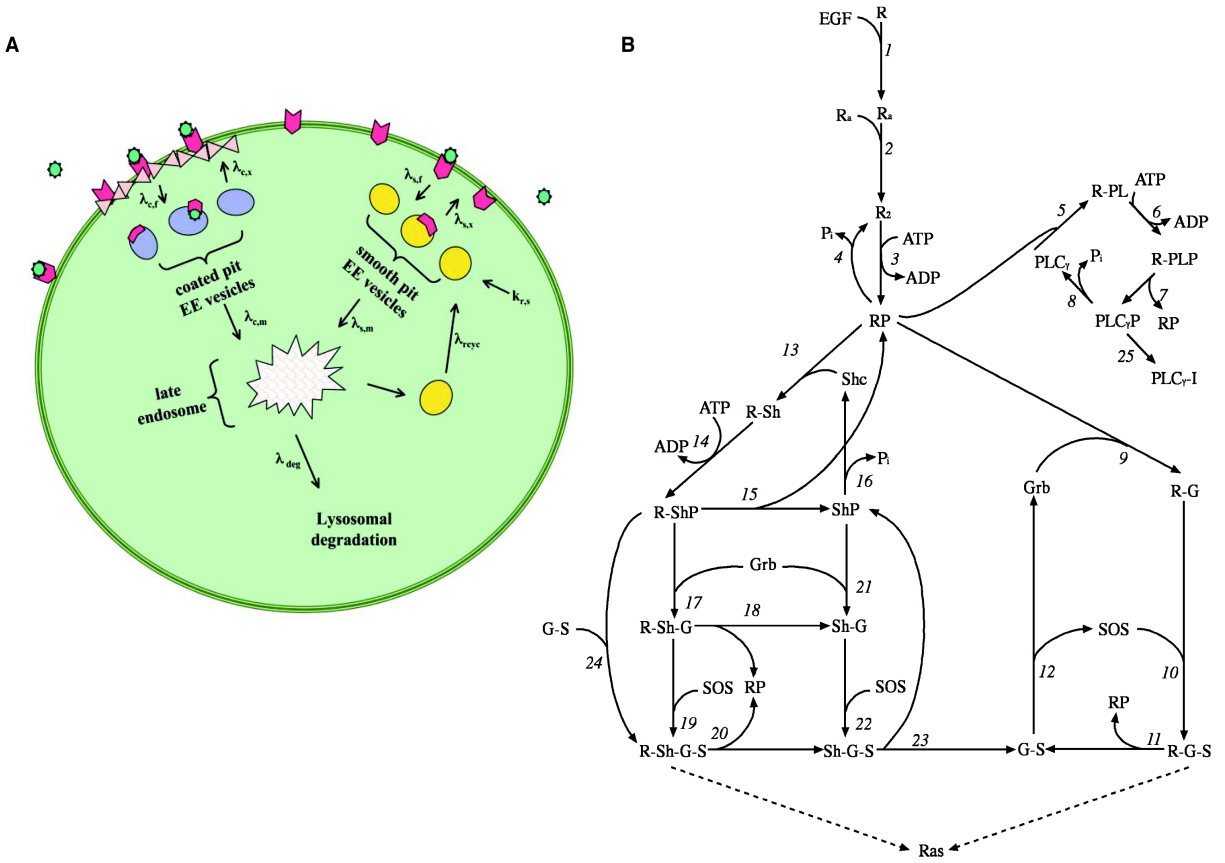


Figure 20: An EGFR model that integrates both internalization and signaling. A - The diagram shows the compartments involved in receptor trafficking. EGFR can be internalized through a ligand induced pathway, where clathrin-coated pits direct the receptor to early endosomes (EEs), or through a constitutive pathway where random non-coated (smooth) vesicles are formed at the PM. Vesicles can recycle back to the PM with different rates. EEs converge into late endosomes, whereby the receptor is mainly degraded and only a few non-stimulated receptors can escape lysosomal degradation and be recycled to the PM. **B -** The signal transduction model that was integrated in this work, was adapted from Kholodenko *et al.* (Kholodenko 1999). From the monomeric form of the EGFR (R) a subsequent series of reactions lead to the formation of the phosphorylated receptor (RP) and to the activation of signaling cascades that converge at the level of Ras activation. (Adapted from Resat, Ewald et al. 2003)

downregulation of EGF- or TGFα-stimulated receptors. Overall, this work suggested that EGFR trafficking controls the modulation and diversification of the signaling cascade that has been shown to be dependent upon the ligand that activates the receptor.

In 2009, the non-linear dependence of downstream signaling on receptor activity was highlighted by a computational model from Lauffenburger and Sorger (Chen, Schoeberl et al. 2009). This model included all the four receptors of the ErbB family, the activity of two ligands (EGF and HRG), and first- or second-order reactions to include protein endocytosis and degradation. Moreover, the model aimed to understand the readouts of the AKT and ERK pathways given different concentrations of the two ligands as input. The authors found that

their model, comprising all the four ErbB receptors, but designed to encompass the internalization events in a simplified approach, was able to recapitulate the activities of the AKT and MAPK modules. In particular, the model predicted, and experiments confirmed, that the system shows an ultrasensitive behavior: due to signal amplification, as the signal descends down the MAPK cascade, the stimulus/response curve becomes steeper and the relationship between EGFR activation and ERK phosphorylation is nonlinear.

Another interesting model that integrates both EGFR activation and internalization is the work by Resat group (Shankaran, Zhang et al. 2012), where authors focused their attention on the integration of EGFR compartmentalization and the activity of phosphatases. The major features of their model were the following: I) less than 40% of EGFR are phosphorylated at any time-point, even after stimulating cells with saturating ligand doses; II) the rate of receptor dephosphorylation at the PM is comparable to that in the EEs; III) dephosphorylation of the receptor in the LEs, which determines signal extinction, is an important component of the cellular response and takes place before lysosomal degradation.

Thus, the EGFR modeling field is constantly updating and integrating novel discoveries and experimental evidence into pre-existing models with the aim of recapitulating *in-silico* a complete scenario of EGFR-dependent signaling. Such an approach is important as it could help shed light on EGFR-related dysfunctions and pathologies.

Rationale of the project

EGFR-dependent signal transduction is a key molecular chain of events important in determining cell fate and growth. EGFR biology is a very complex system, the complexity of which starts already at the PM, where diverse ligands can interact with and activate the receptor. Moreover, heterodimerization of the EGFR with other ErbB receptors increases the number of possible states of the EGFR at the PM and contributes to diversification of the downstream biological responses.

Endocytosis has also emerged as a crucial regulator of the EGFR signaling response, since it is widely accepted that internalized receptors continue to signal from intracellular compartments. In addition, the internalization route is critical in the regulation of receptor fate. Indeed, EGFRs can be internalized through CME, which is fundamental for sustaining signaling, or via NCE, which is essential for receptor degradation. This latter endocytic pathway is activated only when high doses of EGF are applied to cells, thus, safeguarding them against excessive signaling. In particular, NCE activation occurs over a very narrow interval of EGF concentrations and is mechanistically linked to EGFR ubiquitination, which follows a threshold response over the same range of EGF doses. The link between an analogical variable (i.e., the linear input of EGF concentrations) and the digital output (i.e., ubiquitination threshold and NCE activation) needs to be clarified.

A number of modeling approaches aimed at understanding ErbB receptor signaling, particularly that of the EGFR, are present in literature. Until now, more effort has been placed on implementing aspects of the output layer and the signaling processing unit, covering many aspects of EGFR activation and signal transduction. However, EGFR ubiquitination has received scarce attention. In the present thesis, we have integrated this novel aspect with previously published models to obtain a detailed model for EGFR activation at the PM.

In order to deal with the system complexity, we have adopted an integrated research approach, combining wet-lab experiments and mathematical modeling. We designed an Early Activation Model (*EAM*) that integrates all the EGFR dynamics occurring at 2 minutes after EGF stimulation. The novelty that we introduced in our model was to add the EGFR ubiquitination component to the system, as a function of EGFR phosphorylation. Although the aim of this thesis project was not to build the mathematical model itself, but rather to combine experimental evidence and modeling issues in a parallel progression towards model development, throughout the next sections, I have tried to give an in-depth overview of the processes leading to the construction of the EAM. Although I did not tackle the mathematical side of the modeling, I have discussed the preceding more-simplified models (MPM and MPM-B), which were developed before reaching the final model, the EAM. The final model provides an extensive description of the mechanisms that contribute to the creation of the threshold given increasing EGF concentrations, from the dissection of the phosphorylation events to the mechanism of EGFR ubiquitination.

We proceeded in modeling EGFR activation by taking advantage of the many existing models of EGFR activation and dependent signaling (see Introduction Section 7), while also taking into account the ubiquitination component of the system that has so far received scarce attention. This integrated approach aimed to decipher the biological rules underlying the generation of the ubiquitination threshold, but also to generate novel and unexpected predictions of the behavior of the EGFR system that could contribute to aberrant EGFR signaling, as observed in cancer.

Materials and methods

Solutions

Phosphate-buffered saline

NaCl	137 mM
KCl	2.7 mM
Na_2HPO_4	10 mM
KH_2PO_4	2 mM

8 g of NaCl, 0.2 g of KCl, 1.44 g of Na_2HPO_4, and 0.24 g of KH_2PO_4 were dissolved in 800 ml of distilled water. The pH was adjusted to 7.4 with HCl and the volume was brought to 1 litre with distilled H_2O.

Tris-HCl (1 M)

121.1 g of Tris base were dissolved in 800 ml distilled H2O. The pH was adjusted to 7.4, 7.6 or 8.0 with HCl, and distilled H2O was added to bring the volume up to 1 litre.

Tris-buffered saline (TBS)

NaCl	137 mM
KCl	2.7 mM
Tris HCl, pH 7.4	25 mM

8 g of NaCl, 0.2 g of KCl and 3 g of Tris base were dissolved in 800 ml of distilled H_2O. The pH was adjusted to 7.4 with HCl and distilled H_2O was added to bring the volume up to 1 litre.

10X SDS-PAGE running buffer

Glycine	192 mM
Tris HCl, pH 8.3	250 mM
SDS	1%

10X Western transfer buffer

Glycine	192 mM
Tris HCl, pH 8.3	250 mM

For 1X western transfer buffer, the 10X stock was diluted 1:10 with ddH_2O and 20% v/v methanol or ethanol.

50X TAE (Tris-Acetate-EDTA)

Tris base	2 M
Acetic acid	1 M
EDTA, pH 8	10 mM

The pH was adjusted to 8.5 with HCl and distilled H_2O was added to bring the volume up to 1 liter.

Protein buffers

1X JS buffer

HEPES, pH 7.4	50 mM
NaCl	150 mM
Glycerol	10%
Triton X-100	1%
$MgCl_2$	1.5 mM
EGTA	5 mM

1X RIPA buffer

Tris HCl, pH 7.6	50 mM
NaCl	150 mM
NP-40	1%
SDS	0.1% for standard RIPA 1% for RIPA 1% SDS 0% for RIPA w/o SDS
Deoxyxholic acid	0.5%
EGTA	5 mM

500X Protease inhibitor cocktail from Calbiochem, sodium pyrophosphate pH 7.5 20mM, sodium fluoride 250 mM, PMSF 2 mM, and sodium orthovanadate 10 mM were added to the buffer just before use.

1X Laemmli buffer

SDS	2%
Tris HCl, pH6.8	62.5 mM
Glycerol	10%
Bromophenol blue	0.1%
β.Mercaptoethanol	5% (v/v)

SDS-PAGE sample buffer was prepared as a 5X stock solution and stored at -20°C, protected from light.

Reagents and antibodies

Human recombinant EGF was from INALCO. ^{125}I-EGF was from Perkin Elmer. Doxycycline hydrocloride was from SIGMA.

Antibodies were:

- an in-house polyclonal anti-EGFR against aa 1172-1186 of human EGFR;
- a monoclonal anti-EGFR (m108 hybridoma, directed against the extracellular domain of human EGFR, ATCC);
- anti-pY (4G10, Millipore); anti-Ub (P4D1, Santa Cruz Biotechnology, used in all anti-Ub IBs);
- anti-Ub FK2 (BIOMOL, used in all ELISA assays);
- anti-EGFR phospho-specific antibodies (pY1045, pY1068, Cell Signaling);
- anti-tubulin and anti-Grb2 (Santa Cruz Biotechnology);
- anti-Cbl polyclonal (Santa Cruz Biotechnology, used in IB) and monoclonal (BD, used in IF) antibodies.

RNAi oligos

Oligos used in experiments reported in the figures in the Results Section 7.1

- EGFR human (Invitrogen): CCGCAGCAUGUCAAGAUCACAGAUU
- Negative control siRNA: the negative control used in our assays was All Stars from Qiagen.

Cloning techniques

Agarose gel electrophoresis

DNA samples were loaded on 0.8%-1% agarose gels along with DNA markers (1 kb DNA Ladder, NEB). Gels were made in TAE buffer containing Gel Red (Biotium), according to manufacturer's instructions, and run at 100 V until desired separation was achieved. DNA bands were visualized under a UV lamp.

Minipreps

Individual colonies were used to inoculate 3 ml LB (containing the appropriate antibiotic) and grown overnight at 37°C. Bacteria were transferred to Eppendorf tubes and centrifuged for 5 minutes at 16,000*xg* using a 5415 R centrifuge. Minipreps were performed with the Wizard Plus SV Minipreps Kit (Promega) following manufacturer's instructions. The plasmids were eluted in 30 µl nuclease free H_2O.

Diagnostic DNA restriction

Between 0.5 and 5 µg DNA were digested for 2 hours at 37°C with 10-20 units of restriction enzyme (New England Biolabs). For digestion, the volume was made up to 20-50 µl with the appropriate buffer and distilled H_2O.

Large scale plasmid preparation

Cells containing transfected DNA were expanded into 250 ml cultures overnight. Plasmid DNA was isolated from these cells using the Qiagen Maxi-prep kit according to manufacturer's instructions.

Transformation of competent cells

An aliquot of competent cells TOP10 (Invitrogen) were thawed on ice for approximately 10 minutes prior to the addition of plasmid DNA. Cells were incubated with DNA on ice for 30

minutes and then subjected to a heat shock for 45 seconds at 42°C. Cells were returned to ice for an additional 5 minutes. Then, 900 µl of LB medium was added and the cells were left at 37°C for further 60 minutes before plating them onto agar plates with the appropriate antibiotic. Plates were incubated overnight at 37°C.

Constructs and plasmids

N-terminal HA-tagged human c-Cbl WT [kindly provided by Y. Yarden - Weizmann Institute, Israel (Levkowitz, Waterman et al. 1998)] were subcloned into pBABE-puro, pGEX-6P (GE Healthcare) or pSLIK (Invitrogen) vectors, through restriction enzyme digestion (New England Biolabs) and ligation (New England Biolabs), starting from original pcDNA. All clones were sequence verified.

EGFR mutants were generated by site-directed mutagenesis of the human EGFR cDNA and sub-cloned into the pBABE-puro retroviral vector. All clones were sequence verified.

Cell culture

Cell culture media

HeLa cells were grown in GlutaMAX™-Minimum Essential Medium (MEM, Gibco Invitrogen), supplemented with 10% Fetal Bovine Serum South American (FBS SA, Invitrogen), sodium pyruvate 1 mM (Euroclone), non-essential aminoacids (Euroclone), and 2 mM glutamine.

Murine fibroblastic NR6 cells, Phoenix helper cell were grown in Dulbecco's Modified Eagle's Medium (DMEM, Lonza), supplemented with 10% Fetal Bovine Serum South American (FBS SA, Invitrogen) and 2 mM glutamine.

NIH 3T3 cells were grown in Dulbecco's Modified Eagle's Medium (DMEM, Lonza), supplemented with 10% Calf Serum South American (Invitrogen) and 2 mM glutamine. NIN cells stably expressing human EGFR were cultured with supplemental 2.5 ug/ml puromycin; NIN cells stably expressing also human c-cbl were cultured with supplemental 200 ug/ml neomycin.

WI38 were grown in Dulbecco's Modified Eagle's Medium (DMEM, Lonza), supplemented with 10% Fetal Bovine Serum North American (FBS NA, Invitrogen) and 2 mM glutamine.

CASKI cells were grown in RPMI (Lonza) medium supplemented with 10% Fetal Bovine Serum North American (FBS NA, Invitrogen) and 2 mM glutamine.

BT20 cells were grown in GlutaMAX™-Minimum Essential Medium (MEM, Gibco Invitrogen), supplemented with 10% Fetal Bovine Serum North American (FBS NA, Invitrogen) sodium pyruvate 1 mM (Euroclone), non-essential aminoacids (Euroclone), and 2 mM glutamine.

CHO cells were grown in Ham's F12 supplemented with 10% Fetal Bovine Serum South American (FBS SA, Invitrogen).

Transfections

RNAi transfections were performed using LipofectAMINE RNAi MAX reagent from Invitrogen, according to manufacturer's instructions. Oligos for c-Cbl and Cbl-b mRNA silencing were designed with BLOCK-iT™ RNAi Designer from Invitrogen. Cells were subjected or double (in both suspension and adhesion) transfection, treated with 10 nM RNAi oligo and analyzed 5 days after transfection.

Retroviral and lentiviral infection

Human EGFR WT was stably expressed in NIH 3T3 cells by retroviral infection. Briefly, vuses were produced by transfecting the Phoenix helper cell line with 5-10 μg of DNA. 48 hours after transfection, supernatant was collected and passed through a 0.45 μm filter. After the addition of 8 μg/ml polybrene (Hexadimethrine bromide, Sigma), the supernatant was added to NIN 3T3 cells plated on 10 cm cell culture dishes. Two cycles of infection were repeated, after which the medium was replaced with standard HeLa medium (see above "Cell culture media"). 48 hours after infection, selection of infected cells was performed by adding puromycin

Single clones were isolated and characterized for EGFR surface expression levels and homogeneity by immunofluorescence (IF) (Fig. 41). Briefly, cells were fixed in 4% paraformaldehyde and stained (prior to permeabilization) with the monoclonal anti-EGFR antibody that recognizes the extracellular domain (m108). Cells were then incubated with Alexa488-conjugated secondary antibody (Molecular Probes) and stained with DAPI. Clones expressing different levels of surface EGFR, assessed by saturation binding with ^{125}I-EGF, were selected for further analysis (Figure 8A, right panel).

These clones were subsequently infected with an inducible lentiviral construct carrying cDNA encoding human Cbl (pSLIK-neo vector). This construct was engineered starting from a Cbl cDNA kindly provided by Y. Yarden (Weizmann Institute, Israel). All clones were sequence-

verified. Stable bulk populations were obtained after 10 days of neomycin treatment (200 μg/ml). Upon doxycycline treatment (0.5 μg/ml for 16 h), Cbl was expressed at levels ~80-100–fold greater than those of the endogenous protein and displayed a homogeneous expression level (assessed by IF, Fig. 44). The same clones infected with the empty vector were treated with doxycycline in the same way and used as control.

EGF treatment was for 2 min at 37°C. Under these conditions, EGFR internalization is negligible and the observed phosphorylation and ubiquitination events occur primarily at the PM.

The EGFR mutants were stably expressed in NR6 and CHO cells and were generated as described for NIH 3T3 cells. EGFR surface expression in the transfectants was assessed by saturation binding with ^{125}I-EGF.

Protein production and purification

GST-fusion protein production

List of the GST-proteins used:

- c-Cbl: full length (*Homo sapiens*);
- Grb2: full length (*Homo sapiens*);

Rosetta cells transformed with the indicated GST-fusion construct were picked from individual colonies and, used to inoculate 50 ml LB (containing ampicillin at 25 μg/ml and cloramphenicol at 34 ug/ml). Cultures were grown overnight at 37°C. The 50 ml overnight culture was diluted in 1 litre of LB and was grown at 37°C until it reached an OD of approximately 0.6. Then, 1 mM IPTG was added and the culture was grown either at 37°C for theree hours or at 18°C overnight. Cells were then pelleted at 4000 x g for 10 minutes at 4°C and pellets were resuspended in GST-lysis solution (20 ml/liter of bacteria). Samples were sonicated 5 times for 20 seconds each on ice and were pelleted at 14000 x g for 30 minutes at 4°C. 600 μl – 1ml of Glutathione Sepharose 4B (GE Healthcare) beads (1:1 slurry),

previously washed 3 times with PBS and once with GST-lysis buffer, was added to the supernatants and samples were incubated for either 4 hours or overnight at 4°C with rocking. Beads were washed once in PBS containing 1% triton, followed by 2 times in PBS alone. Beads were finally resuspended in 1:1 volume of GST-maintenance solution and kept at -80°C.

GST-lysis solution

HEPES, pH 7.5	50 mM
NaCl	200 mM
EDTA	1 mM
Glycerol	5%
NP-40	0.1%
Protease Inhibitors (Calbiochem)	1:500

GST-maintenance solution

Tris, pH 7.4	50 mM
NaCl	100 mM
EDTA	1 mM
Glycerol	10%
DTT	1 mM
Protease Inhibitors (Calbiochem)	1:500

Cleavage of GST-fusion proteins

GST-fusion proteins were cleaved with PreScission Protease (provided by the IFOM Biochemistry Unit). 1 unit of enzyme for 100 μg of fusion protein was added to the beads in the presence of GST-maintenance solution and incubated either overnight at 4°C or for 4

hours at room temperature. After cleavage the supernatant containing the cleaved protein was collected.

Protein procedures

Cell lysis

After washing with PBS 1X, cells were lysed in JS or RIPA buffer directly in the cell culture plates using a cell-scraper and clarified by centrifugation at 16,000 *xg* for 20 min at 4ºC using a 5415 R centrifuge. Protein concentration was measured by the Bradford assay (Biorad) following manufacturer's instructions.

For the DELFIA assays and some IP assay, lysis was performed in RIPA buffer containing 1% SDS, followed by clarification for 1 h at 120,000 g and dilution to a final SDS concentration of 0.2%.

1% SDS lysis

For this lysing procedure prepare two independent solutions (see Protein Buffers section):

-RIPA 2% SDS

-RIPA w/o SDS

Wash cells on ice with cold PBS, twice and let cells dry well. Add lysis buffer 2% SDS and ultracentrifuge cells at 45000 rpm at 4'C. Discard pellet, collect the supernatant and dilute 1:5 the lysate in RIPA w/o SDS to reach a concentration of 0.2% SDS. If performing an IP, keep SDS at a concentration of 0.2% also during IP and wash steps.

SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Gels for resolution of proteins were made from a 30%, 37,5:1 mix of acrylamide: bisacrylamide (Sigma). As polymerization catalysts, 10% ammonium persulphate (APS) and TEMED were used.

Separating gel mix

	Gel %		
	6	**8**	**10**
Acrylamide mix (ml)	2	2.7	3.3
1.5 M Tris HCl, pH 8.8 (ml)	2.5	2.5	2.5
ddH_2O (ml)	5.3	4.6	4
10% SDS (ml)	0.1	0.1	0.1
10% APS (ml)	0.1	0.1	0.1
TEMED (ml)	0.01	0.01	0.01
TOTAL (ml)	10	10	10

Stacking gel mix

Acrylamide mix (ml)	1.68
1 M Tris HCl, pH 6.8 (ml)	1.26
ddH_2O (ml)	6.8
10% SDS (ml)	0.1
10% APS (ml)	0.1
TEMED (ml)	0.01
TOTAL (ml)	10

Western Blot (WB)

Desired amounts of proteins were loaded onto 1-1.5 mm thick SDS-PAGE gels for electrophoresis (Biorad). Proteins were transferred in western transfer tanks (Biorad) to nitrocellulose (Schleicher and Schnell) in 1X Western transfer buffer (supplemented with 20% methanol or ethanol) at 30 V overnight or 100 V for 1 hour for small gels and at 30 V overnight or 0.8 A for 2 hours for large gels. Ponceau staining was used to determine the efficiency protein transfer onto the filters. Filters were blocked for 1 hour (or overnight) in 5% milk or BSA in TBS supplemented with 0.1% Tween (TBS-T). After blocking, filters were incubated with the primary antibody, diluted in TBS-T 5% milk or BSA, for 1 hour at room temperature, followed by three washes of ten minutes each in TBS-T. Filters were then incubated with the appropriate horseradish peroxidase-conjugated secondary antibody diluted in TBS-T for 30 min. After the incubation with the secondary antibody, the filter was washed 3 times in TBS-T (10 minutes each) and the bound secondary antibody was revealed using the ECL method (Amersham) or ChemiDoc MP system (Biorad).

Anti-Ub western blot

After SDS-PAGE, proteins were transferred on a PVDF (polyvinylidene fluoride) membrane (Immobilion P, Millipore), previously activated by incubation in 100% MeOH for 5 minutes at room temperature. Ponceau staining was avoided since it might interfere with antibody recognition. After transfer, filters were subjected to a denaturing treatment in a dedicated solution for 30 minutes at 4ºC. This treatment denatures Ub and facilitates the recognition of latent Ub epitopes by anti-Ub antibody resulting in intensification of the anti-Ub signal. After extensive washing in TBS-T buffer, filters were blocked overnight at 4ºC in 5% BSA (dissolved in TBS-T). After blocking, filters were incubated with the antibodies against Ub, diluted in TBS-T 5% BSA, for 1 hour at room temperature, followed by 3 washes of 10 minutes each in TBS-T. Filters were then incubated with the anti-mouse horseradish

peroxidase-conjugated secondary antibody, diluted in TBS-T 3% BSA, for 30 minutes at room temperature. After incubation with the secondary antibody, the filter was washed 3 times in TBS-T (10 minutes each) and the bound secondary antibody was revealed using the ECL method (Amersham).

Denaturing solution

Guanidinium Chloride	6 M
TRIS, pH 7.4	20 mM
PMSF (freshly added)	1 mM
β-Mercaptoethanol (freshly added)	5 mM

Immmunoprecipitation

Lysates prepared in JS (for co-IP) or in RIPA/RIPA 1% SDS (for IP) buffer were incubated in the presence of specific antibodies (about 1-2 μg/mg of lysates) for 2 hours at 4ºC with rocking. Protein G Sepharose beads (Zymed) were then added, and samples were left for an additional hour at 4ºC, rocking. Immunoprecipitates were then washed 4 times in the appropriate lysing buffer. Proteins were eluited in Laemli buffer 2x. IP and were performed starting from 500 μg of lysates, with the appropriate antibodies. Co-IP assays between Cbl and EGFR mutants in NR6 cells were performed starting from 1 mg (for EGFR WT and Y1045/68/86+) or 2 mg of lysate (for the 1045+ mutant), for each condition. Washes were then performed in JS buffer (for EGFR WT and the 1045/68/86+ mutant) or in JS buffer containing a 5-fold reduced amount of Triton X-100 (0.2%) for the 1045+ mutant.

Immunoblotting (IB) was performed as described previously.

Determination of the number of Cbl and Grb2 molecules

To calculate the number of Grb2 or Cbl molecules in HeLa (Fig. 35), NIH (Fig. 41) or CASKI/BT20 (Fig. 42) cell lysates, we compared signal intensities in anti-Grb2 or anti-Cbl IB of increasing amounts of cellular lysate with known amounts of purified Grb2 or Cbl protein. From the signal comparison and taking into account Avogadro's number, we calculated the number of Grb2 and Cbl molecules/μg of lysate. After an additional correction for the number of HeLa or NIH-EGFR cells corresponding to 1 μg of lysate (measured in the same experiment), we obtained the total number of Grb2 and Cbl molecules/cell. To calculate the number of active Cbl molecules (Cbl pY), we IP Tyr phosphorylated Cbl, using the anti-pY antibody, from increasing amounts of lysate (prepared in RIPA buffer containing 1% SDS) from cells stimulated with EGF (100 ng/ml) for 2 min. Immunoprecipitates were then IB with anti-Cbl and the signal intensities of the anti-Cbl bands were compared with the signal intensities of anti-Cbl bands in IB cellular lysates and purified Cbl protein. We corrected the obtained values for the efficiency of IP (estimated >90%), and obtained the % of Cbl in the lysate that is phosphorylated (active) upon stimulation with EGF (100 ng/ml) for 2 min.

Assays with ^{125}I-EGF

Saturation binding assay

Cells were serum starved in binding buffer (MEM or DMEM, BSA 0.1%, Hepes pH 7.4 20 mM) for at least 2 hours. Serum starved cells were then incubated on ice for 6 h in the presence of ^{125}I-EGF (100 ng/ml: 10 ng/ml ^{125}I-EGF plus 90 ng/ml cold EGF) in serum-free medium supplemented with 0.5% BSA. Samples were cooled on ice for 30 minutes and incubated at 4°C with mix containing ^{125}I-EGF. After 6 hours, cells were washed 3 times in PBS, and solubilized in 300 μl 1 M NaOH. After correction for the hot/cold dilution, the

number of receptors on the surface was deduced from the specific activity of the labeled ligand. Non-specific binding was measured in the presence of a 300-fold excess of cold EGF, and was never > 3-10% of the total counts.

Immunofluorescence studies

Cells were plated on glass coverslips pre-incubated with 0.1% gelatin in PBS at 37°C for 30 minutes. Cells were fixed in 4% paraformaldehyde (in 1X PBS) for 10 minutes, washed with PBS and permeabilized in 0.1% Triton X-100, BSA 0.2% in 1X PBS for 8 minutes at room temperature. To prevent non-specific binding of the antibodies, cells were then incubated with 1X PBS in presence of 2% BSA for 30 minutes at room temperature. Next, cells were incubated for 1 hour with primary antibody in 1X PBS in presence of 0.2% BSA (anti-EGFR m108 or anti-c-Cbl BD), washed 3 times with 1X PBS and incubated for 30 minutes with fluorescently labeled secondary antibodies (Amersham). After 3 washes with PBS, nuclei were DAPI-stained for 5 minutes and washed again 3 times with 1X PBS. Coverslips were immediately mounted with moviol and examined under a wide-field fluorescence microscope (Olympus). Images were further processed with the ImageJ software. To detect only surface EGFR (anti-EGFR m108) permeabilization step was avoided.

ELISA assays for EGFR ubiquitination and phosphorylation

For the ELISA-based assays shown in Figure 8 and Supplementary Figure 7, we used the DELFIA (Dissociation Enhanced Lanthanide Fluoroimmunoassay) technology from Perkin Elmer. This technology is based on sandwich-recognition of a target protein by a capture antibody and a detection antibody. The capture antibody is immobilized on a solid surface (microwells) directly through non-covalent bonds. After the addition of the analyte (appropriate cellular lysate), the detection of signals relies on a lanthanide (Europium)-conjugated antibody that is able to produce a fluorescent signal upon enhancement with acidic

enhancement buffer. Plate preparation, analyte incubation and antibody detection were according to the manufacturer's instruction. Briefly, microwells plates were coated with the capturing antibody (see below). Blocking was performed for 2 h with BSA 2% in PBS. 25-50 μg of lysates from HeLa or NR6 cells, stimulated with the indicated concentration of EGF, were incubated overnight at 4ºC. Lysates were prepared in RIPA/1% SDS buffer and diluted to 0.2% SDS before incubation step. After three washes, wells were incubated with primary antibodies, diluted at 1 μg/ml in assay buffer, for 1 h at RT. After three washes, anti-mouse or rabbit Europium-labeled secondary antibodies (1 μg/ml in assay buffer) were added for an additional hour. After three washes and treatment with enhancement solution, fluorescence was measured with EnVision instrument (excitation at 340 nm and emission at 615 nm).

Capturing and detecting antibodies differed depending on whether a forward or reverse approach was performed (see Supplementary Figure 2A-B for a scheme of the two procedures):

	Capturing antibodies	**Detecting antibodies**
Forward ELISA	Home made polyclonal anti-EGFR directed against aa 1172-	Monoclonal antibodies against Ub (FK2), pY (4G10), pY1068, or
Reverse ELISA	Monoclonal antibodies against Ub (FK2, 5 μg/ml), pY (4G10, 5 μg/ml), pY1068 (1 μg/ml),	Home made polyclonal anti-EGFR directed against aa 1172-1186 of human EGFR (1 μg/ml)

Densitometry and statistical analysis

Quantification of blot was performed with Photoshop. Error bars in the plots represent the standard deviation of the mean. All statistical analyses were performed using Excel.

RESULTS

1 - A quantitative assay to follow EGFR ubiquitination and phosphorylation

We have previously shown that, when treated with increasing doses of ligand, cells showed a peculiar behavior: EGFR-Ub increased sharply over a narrow range of EGF concentrations, being minimal at 1 ng/ml and nearly maximal at 10 ng/ml, both in epithelial cells (HeLa) and in fibroblasts (NR6-EGFR cells; **Fig. 8**, page 33). On the other hand, EGFR-pY, which is a readout for receptor activation, displays a gradual and linear increase over the same range of EGF concentrations (1-10 ng/ml EGF). These data were obtained by western blot analysis, which is a powerful technique that allowed us to explore receptor ubiquitination under different conditions (Sigismund, Algisi et al. 2013). However, western blot analysis is a semi-quantitative assay and has several limitations, the major one being its low dynamic range due to sensitivity and saturation problems. Before proceeding with the integration of EGFR-pY/-Ub data into a mathematical formalism, we needed a more sensitive and quantitative high-throughput assay to detect receptor ubiquitination and phosphorylation.

To this aim, we set-up a quantitative ELISA-based immunoassay, DELFIA (Dissociation-Enhanced Lanthanide Fluorescent ImmunoAssay; Perkin Elmer), which employs a fluorescent molecule instead of chemoluminescence. This assay takes advantage of the unique chemical properties of lanthanide-labeled biomarkers, to create an assay that performs well in terms of sensitivity, wide dynamic range and stability. In common with the ELISA technique, DELFIA is based on sandwich-recognition of the target protein using a capture antibody and a detection antibody. The capture antibody is immobilized directly onto a solid surface (e.g. a suitable microwell plate). After the addition of the sample and binding of the target protein to the capture antibody, the detection antibody is added. This antibody binds to an exposed motif

on the target protein and is either directly conjugated to a lanthanide or is, in turn, recognized by a lanthanide-conjugated species-specific secondary tracer antibody. The lanthanide tracer is then detected following enhancement with an acidic enhancement buffer, which results in the production of a fluorescent signal (for a schematic representation see **Fig. 21A**).

Our assay utilizes europium (Eu), a member of the lanthanide chelates. During the assay,

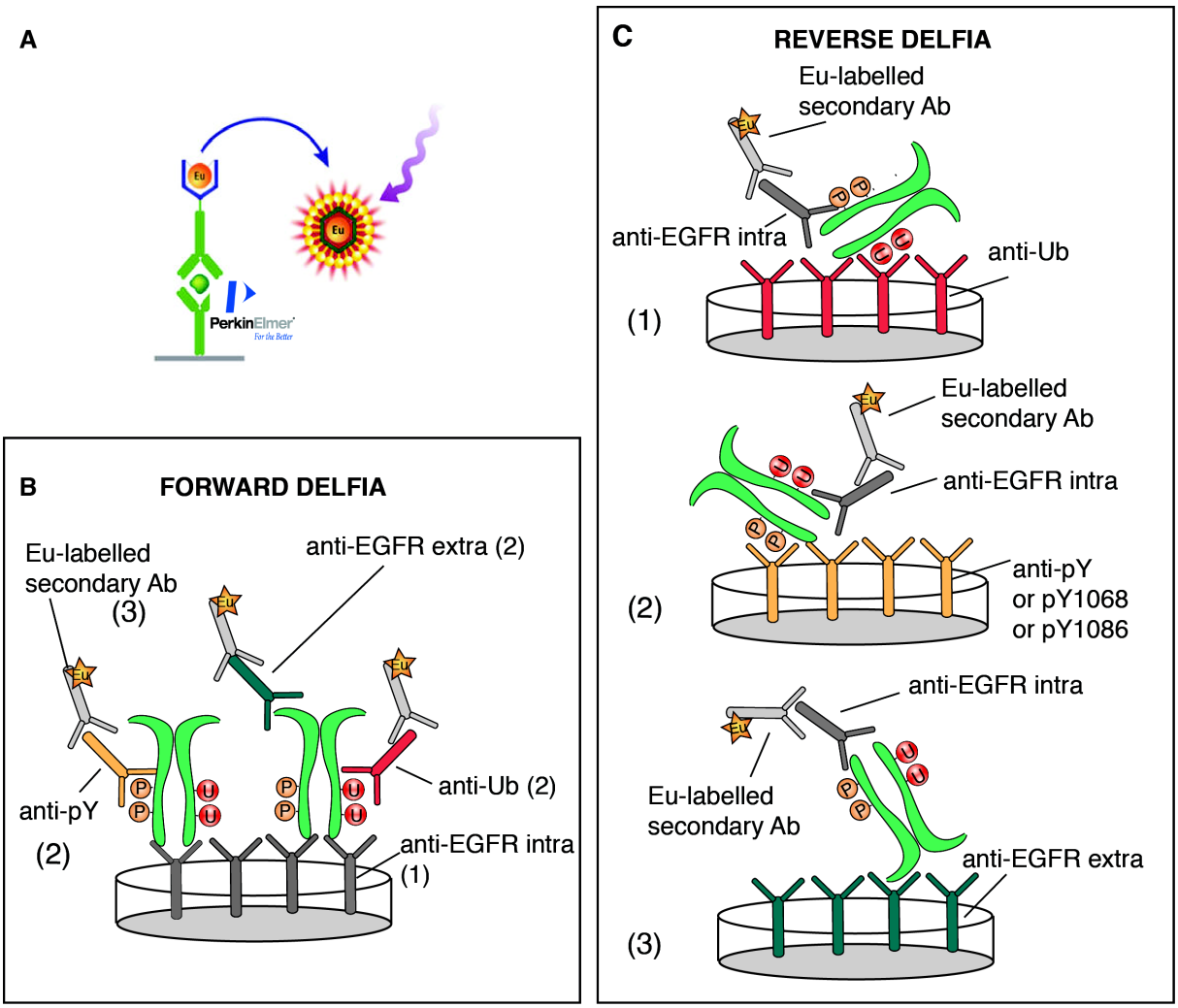


Figure 21: Schematic representation of DELFIA assay. A. DELFIA assay takes advantage of a lanthanide chelate, in this case Europium (Eu) conjugated to a specific antibody. The native lanthanide chelate is essentially non-fluorescent throughout the assay. Upon addition of a specific DELFIA Enhancement Solution, lanthanide ions are released, forming fluorescent chelates that are stably protected inside a micelle. **B.** In the "forward" approach, microwell plates were coated with a polyclonal anti-EGFR antibody (intracellular domain, *anti-EGFR intra*), which captures the receptor from the lysate (1). After incubation of the lysate, the main fraction of the EGFR remains bound onto the plate (green dimeric protein in the scheme). Detection of EGFR-Ub, EGFR-pY or total EGFR was performed with primary monoclonal antibodies directed against Ub (*anti-Ub*: FK2), pY (*anti-pY*: 4G10), or the EGFR extracellular domain (*anti-EGFR extra*: m108) (2), followed by europium labeled secondary antibodies [*(Eu)-labeled secondary Ab*] (3). **C.** In the reverse approach, microwell plates were coated with monoclonal antibodies directed against Ub (*anti-Ub*: FK2), pY (*anti-pY*: 4G10) or the EGFR extracellular domain (*anti-EGFR extra*: m108) that capture EGFR-Ub (1), EGFR-pY (2) or total EGFR (3), respectively. Detection was performed with a polyclonal anti-EGFR antibody recognizing the intracellular domain (*anti-EGFR intra*) followed by a Eu-labeled secondary antibody.

the native lanthanide chelate is essentially non-fluorescent. However, after the binding reaction is complete, lanthanide fluorescence is stimulated by the addition of a specific DELFIA Enhancement Solution: the lanthanide ions are released in solution at low pH, and they rapidly form new, highly stable fluorescent chelates inside a protective micelle with components of the Enhancement Solution. Europium fluorescence is amplified 1-10 million times by this enhancement step and it develops a signal in 5 minutes that is stable for up to 8 hours. Fluorescence can be measured with a suitable detector using the factory-set DELFIA Europium protocol (excitation at 340 nm and emission at 615 nm). Raw data output from the instrument is in the format of a table with a value - representing *counts per seconds* - for each well. This value is a direct measure of the lanthanide fluorescence and is proportional to the amount of tracer antibody that has bound to the well, which, in turn, depends upon the amount of captured target protein.

In the "forward assay" approach (**Fig. 21B**), we immobilized an anti-EGFR antibody that recognizes the intracellular domain of the receptor (anti-EGFR intra) onto the solid surface ("capture antibody"). Plates were then incubated with cell lysate derived from appropriately treated cells (stimulated or not with EGF; see Materials and Methods), in order to capture all EGFR molecules present in the lysate. After extensive washes, plates were incubated with an anti-ubiquitin mouse antibody (anti-Ub) or an anti-phosphotyrosine mouse antibody (anti-pY) ("detection antibodies"), which specifically recognize and bind to epitopes present in the modifications under analysis. The addition of the tracer (Eu-labeled secondary Ab) then allows the amount of ubiquitinated and phosphorylated receptors in the cell to be measured (**Fig. 21B**). In parallel, detection with an anti-EGFR antibody recognizing the extracellular domain of the receptor (anti-EGFR extra), as opposed to the intracellular domain recognized by the capture antibody, permits normalization of results with respect to the total amount of receptor in each well.

A crucial aspect in obtaining a reliable assay was the identification of experimental conditions that result in an optimal signal-to-noise ratio. During the set-up of the assay, we tested different antibodies and different experimental conditions. To optimize the detection of phosphorylated EGFR, we started with a capture antibody (anti-EGFR intra) concentration of 0.1 μg/well. Different quantities of lysate (ranging from 1 to 100 μg/well) derived from HeLa cell that had been serum-starved or EGF-stimulated for 2 minutes, were tested. Phosphorylated EGFR was measured using:

- anti-pY antibody followed by the addition of an anti-mouse Eu-conjugated secondary antibody.
- an anti-pY antibody directly labeled with Eu (pY-Eu).

The anti-EGFR extra antibody was used as an internal control, to normalize the values between serum-starved and stimulated conditions based on the amount of total EGFR in each well.

An example of raw output data from the DELFIA assay is shown in **Table V**. From these data, we can appreciate that both anti-pY antibodies tested returned a significantly higher value (higher counts per second) when lysate from ligand stimulated, rather than serum-starved, cells was analyzed. On the contrary, EGFR levels, as measured by the use of anti EGFR antibodies, gave similar counts, reflecting that roughly the same amount of lysate was used for each cellular condition. Indeed, throughout all our DELFIA assays, we considered EGFR levels as the internal normalization factor. The concentration of lysate that gave the best results in terms of signal-to-noise-ratio was 25 μg/well of lysate (see **Table V** and **Fig. 22A**). As expected, total receptor levels did not change between serum-starved and EGF-stimulated cells, since the receptor after 12 hours of serum starvation followed by 2 minutes of EGF stimulation, does not undergo degradation. In contrast, the levels of activated receptor,

as indicated by the anti-pY signal, is significantly increased following 2 minutes of EGF stimulation.

Starting from these experimental conditions (0.1 µg/well anti-EGFR intra, 25 µg/well lysate), we proceeded with the setting up of the assay by optimizing the detection of ubiquitinated EGFR. We tested three different antibodies that should be able to recognize ubiquitinated proteins: the commercially available FK2 antibody and the in-house antibodies ZTA10 and AQ17 (**Fig. 22B**). The latter two failed to recognize the target protein and were discarded. Only FK2 gave a good signal-to-noise ratio between serum-starved and stimulated cells. In the same experiment, we also considered an additional variable, i.e., the stringency of

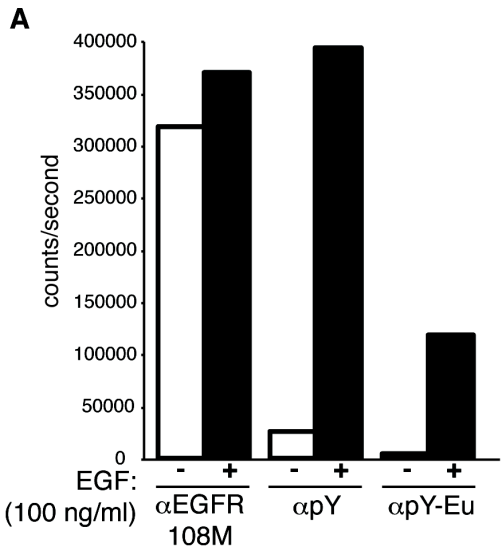
Detection Ab	αpY	αpY	αEGFR-extra	αEGFR-extra	αPY-Eu	αPY-Eu
Secondary Ab	αM-Eu	αM-Eu	αM-Eu	αM-Eu	-	-
Treatment / *Lysate (µg/well)*	SS	EGF 100 ng/ml	SS	EGF 100 ng/ml	SS	EGF 100 ng/ml
100	103353	648641	506806	425978	10935	146729
50	80973	480682	436672	513984	5956	170036
25	25780	394838	319168	370956	4824	119985
12.5	20684	261868	233396	275108	2248	79708
6.25	16941	218828	183023	207080	958	52524
3.125	13850	152866	113938	134632	623	24775
1.56	270	180	289	329	329	250
0	8314	11348	9160	8731	229	16532

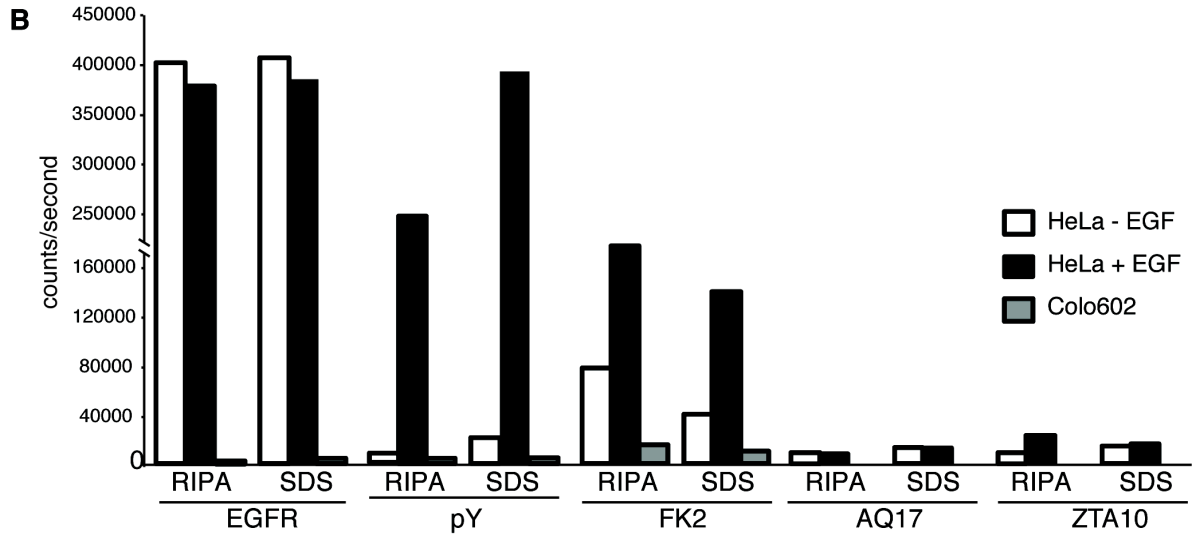
Table V: Typical raw data from a DELFIA assay. The DELFIA assay was performed using cell lysate from HeLa cells that had been serum-starved (SS) or treated with 100 ng/ml EGF for 2 minutes. Cell lysate, at the indicated amounts, was then added to 96 microwell plates coated with 0.1 µg/well of the capture antibody (anti-EGFR intracellular domain). Captured receptor was then detected using different combinations of detection and tracer antibodies: i) the anti-phosphotyrosine (αPY) detection antibody coupled with the europium-conjugated anti-mouse tracer antibody (αM-Eu); ii) the anti-EGFR extracellular domain (αEGFR-extra) detection antibody coupled with the europium-conjugated anti-mouse tracer antibody (αM-Eu); iii) the europium-conjugated anti-phosphotyrosine detection/tracer antibody (αpY-Eu) (see Materials and Methods). Results are expressed as counts per second. The highlighted data obtained with 25 µg/well cell lysate corresponds to data shown in **Fig. 22A**.

our lysis buffer and the concentration of detergents. We tested standard RIPA buffer and RIPA buffer supplemented with 1% SDS (see Material and Methods, section Protein buffers). The rationale for adding additional detergent to the buffer was to prevent the potential co-precipitation of ubiquitinated proteins with EGFR. Furthermore, the FK2 antibody was known to perform better in denaturing conditions.

Figure 22: Optimization of experimental conditions for the DELFIA assay. A - Cell lysate, either serum starved or stimulated with 100 ng/ml, were analyzed by the forward DELFIA assay. Microwell plates were coated with 0.1 µg/well of the capture antibody (anti-EGFR intracellular domain). After lysate addiction and incubation (25 ug of lysate/well), captured receptor was then detected using 0,1 ug/well of the following antibodies: i) the anti-EGFR extracellular domain (αEGFR-extra 108M) detection antibody coupled with the europium-conjugated anti-mouse tracer antibody (αM-Eu); ii) the anti-phosphotyrosine (αPY) detection antibody coupled with the europium-conjugated anti-mouse tracer antibody (αM-Eu); iii) the europium-conjugated anti-phosphotyrosine detection/tracer antibody (αpY-Eu) (see Materials and Methods).

B - The DELFIA assay was performed essentially as described in 'A'. The detection antibodies used were anti-EGFR extra, anti-pY, and various anti-ubiquitin antibodies (FK2, AQ17 and ZTA10, see Materials and Methods). Different lysis conditions were also used: standard RIPA buffer (RIPA) or 1% SDS RIPA buffer (SDS) (see Materials and Methods), as indicated. Colo602 is a cell line that does not express EGFR, used as a negative control.





2 - The EGFR Ubiquitination threshold

We then aimed to validate the EGFR-Ub threshold observed by WB analysis using the DELFIA method. First, we confirmed by WB analysis that the EGFR-Ub threshold was maintained, and occurred over the same EGF concentration range, when cells were lysed in the more stringent RIPA 1% SDS buffer instead of standard RIPA buffer. HeLa cells were plated in duplicate at 50% confluency and stimulated with increasing EGF doses. One sample was lysed in standard RIPA and the other one was lysed in RIPA 1% SDS buffer. Total cell extracts were then immunoprecipitated with an anti-EGFR antibody (see Materials and Methods, section Protein Procedures) and analyzed by WB (**Fig. 23A**). By densitometric analysis of the WB, we observed that the presence of 1% SDS had no effect on the EGFR-Ub threshold, which occurred over the same EGF concentration range under both conditions (**Fig. 23B**), or on phosphorylation, which remained linear with EGF doses.

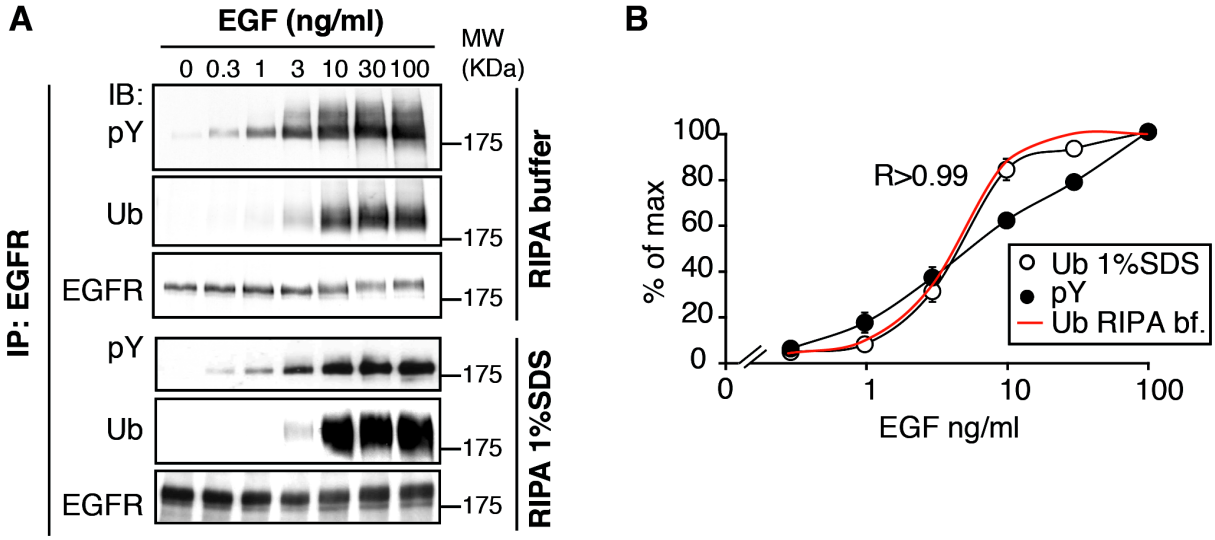


Figure 23: EGFR ubiquitination assay in stringent lysis condition. A - To confirm that the Ub threshold observed by anti-EGFR IP and WB analysis was due exclusively to EGFR ubiquitination and not to co-immunoprecipitating proteins, we repeated the experiment shown in **Fig. 8A** using more stringent lysis conditions (WB relative to RIPA buffer are the same as in Fig.8A). HeLa cells were stimulated with EGF for 2 min at the indicated concentrations and lysates were prepared in 1% SDS RIPA lysis buffer, diluted to 0.2% SDS (see Materials and Methods) and subjected to immunoprecipitation (IP) and IB as shown. Also under these stringent lysis conditions, the Ub threshold was readily detected and indistinguishable from that obtained under standard cell lysis conditions (standard RIPA buffer 0.1% SDS, see Materials and Methods). Blot is representative of 3 repeats. **B -** Quantitation by densitometry analysis of the images relative to 2% SDS blots shown in A as % of max (black lines). The red curve shows quantitation of the anti-Ub signal of the images obtained with lysates prepared with standard RIPA buffer as shown in A and **Fig. 8A**. In all figures, when comparing curves that did not show significant differences it is shown R, the Pearson correlation coefficient.

We thus proceeded with DELFIA "forward" assay (**Fig. 21B**) to analyze EGFR-Ub and -pY. Again HeLa cells were stimulated with increasing EGF doses for 2 minutes, lysed in RIPA 1% SDS and analyzed using the DELFIA assay for the relative levels of EGFR-Ub and EGFR-pY (**Fig. 24A**). With this new approach, we could increase the number of EGF concentrations analyzed, particularly at the transition phase (between 1-10 ng/ml). We confirmed both the sigmoidicity of EGFR ubiquitination and the hyperbolic behavior for phosphorylation (**Fig. 24A**). Importantly, the EGFR-Ub threshold is positioned between 1 and 10 ng/ml of EGF, as observed in WB experiments (**Fig. 24B** and **C**).

Additionally, we performed a "reverse assay" in which plates are coated with anti-

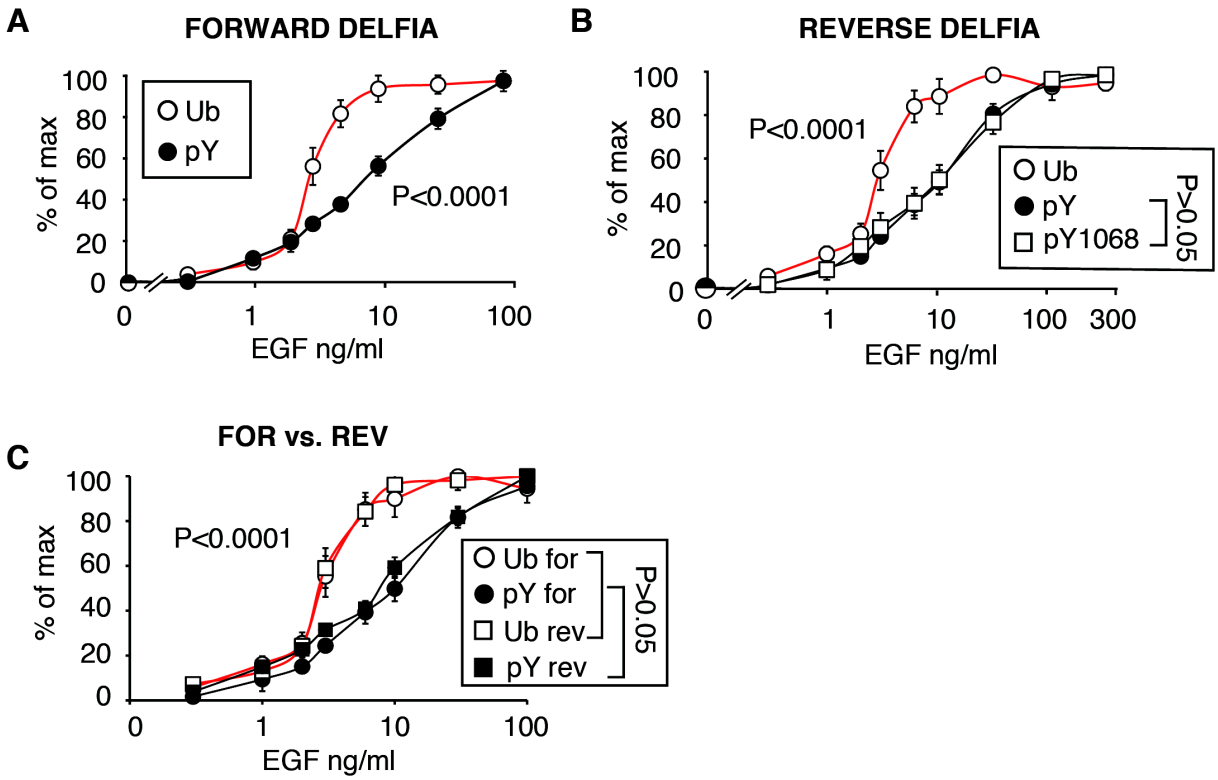


Figure 24: Analysis of EGFR ubiquitination and phosphorylation by DELFIA. A-B - Lysates of HeLa cells stimulated with the indicated doses of EGF for 2 minutes, were subjected to either forward or reverse DELFIA analysis (see **Fig. 21B-C**), using the indicated detection antibodies (Ub, ubiquitination; pY, phosphorylation; pY1068, EGFR 1068 tyrosine specific phosphorylation). Results are shown as % of max. With DELFIA analysis we increased the number of EGF concentrations analyzed with respect to WB analysis. **C -** Comparison of the EGFR ubiquitination and phosphorylation curves obtained by forward and reverse DELFIA analysis of HeLa cell lysates described in C-D. In all panels, results are expressed as the mean ± SD calculated from at least three independent experiments, performed in triplicate. P-values were calculated using two-way ANOVA analysis. In all figures, when comparing curves that showed significant differences it is shown the relative P-values; when comparing curves that did not show significant differences it is shown R, the Pearson correlation coefficient. P-values inside legend boxes (B-C) refer to the groups of data indicated. P-values inset in graphs (A-B-C) refer to Ub vs pY groups of data.

ubiquitin or anti-phosphotyrosine antibodies and captured EGFR is detected using an anti-EGFR intra antibody (as schematized in **Fig. 21C**). The EGFR ubiquitination and phosphorylation curves were verified also with this approach and we obtained curves that were completely superimposable on those obtained with the forward approach, as shown in **Fig. 24C**.

Finally, we wanted to verify if the results obtained with the DELFIA assay were comparable with data obtained from WB analysis. Densitometric analysis of images of digitally acquired blots at different exposures was performed. Despite this procedure being subject to a high degree of variability, the comparison of curves (normalized to max) obtained with the two methods indicated how the two methods gave similar results (**Fig. 25A** and **B).** Ubiquitination behavior is in good agreement between the two methods, although it varies at the 3 ng/ml data point, which is higher in the DELFIA output compared with the WB analysis (**Fig. 25A**). This might be due to a higher sensitivity of the DELFIA signal. In agreement, the standard deviation of DELFIA at this critical point is very low (almost invisible in the chart), indicating higher reliability of results compared with the WB analysis. This variation, however, does not affect the interpretation of the results, since both methods detect a

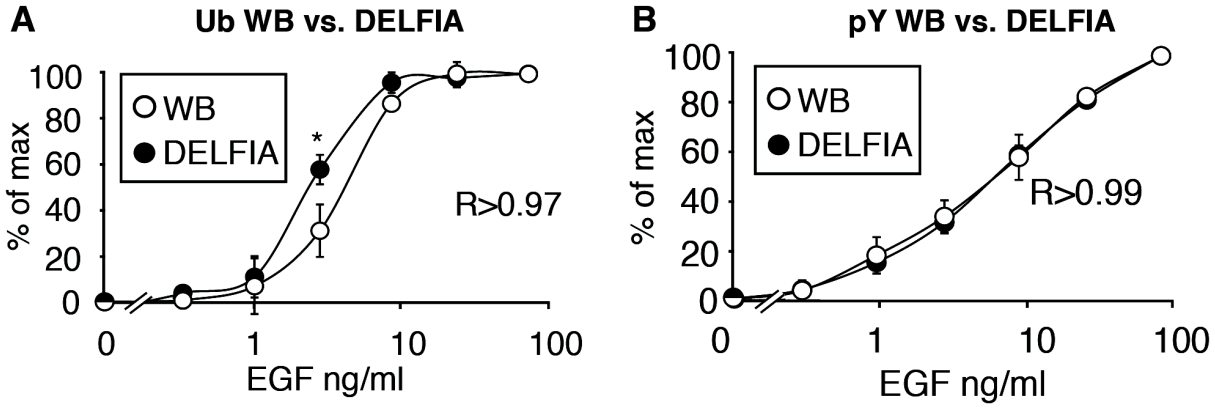


Figure 25: Comparison of EGFR ubiquitination and phosphorylation curves obtained by WB and DELFIA. A - Ubiquitination curves obtained by both WB (from **Fig.23A**) and forward DELFIA analysis (from **Fig.24C**) were superimposed to allow for direct comparison. The asterisk indicates the only point that significantly differed in the two curves (*p = 0.01). **B -** Phosphorylation curves obtained as in **A** are directly compared. In all panels, results expressed as mean ± SD calculated from at least three independent experiments. When curves show significant differences, this is indicated with a p-value calculated using two-way ANOVA analysis. Instead, when curves that did not show significant differences, we calculated the Pearson correlation coefficient (R).

threshold positioned within the same range of EGF concentrations. Moreover, the DELFIA Ub curve exhibited an even sharper threshold effect than the WB curve. The EGFR phosphorylation curves obtained with the two methods were instead identical at all EGF concentrations (**Fig. 25B**). Overall, we concluded that the DELFIA assay is a reliable method for measuring EGFR ubiquitination and phosphorylation.

3 - Dose response curves of EGFR phosphorylation

In collaboration with the Systems Biology group in our Institute, we have designed a mathematical model based on an ordinary differential equations (ODE) that integrates all the overlapping dynamics occurring after two minutes of EGF stimulation. At this time point, the phosphorylation and ubiquitination events take place predominantly at the PM, with a negligible contribution from events at other endomembranes (Sigismund, Algisi et al. 2013). Since EGFR phosphorylation and ubiquitination are causally related, we initially modeled EGFR phosphorylation to obtain a quantitative relationship between these two biochemical modifications.

To guide model construction, we performed experimental dose-response curves in HeLa cells stimulated with increasing EGF concentrations, and analyzed the dynamics of phosphorylation of the individual Tyr residues by WB and DELFIA. The WB analysis was performed using different commercially available phospho-EGFR antibodies that specifically recognize a single phosphotyrosine in the cytoplasmic tail of the stimulated receptor (**Fig. 26A**). Unfortunately, not all the available phospho-specific antibodies worked in the DELFIA assay, with only the pY1068 and pY1086 being functional in this experimental setting (**Fig. 26B**). The data showed that all the individual phosphosites analyzed followed the same gradual increase, with increasing EGF concentration, as global EGFR phosphorylation, suggesting that individual EGFR tyrosine residues can be phosphorylated independently of each other.

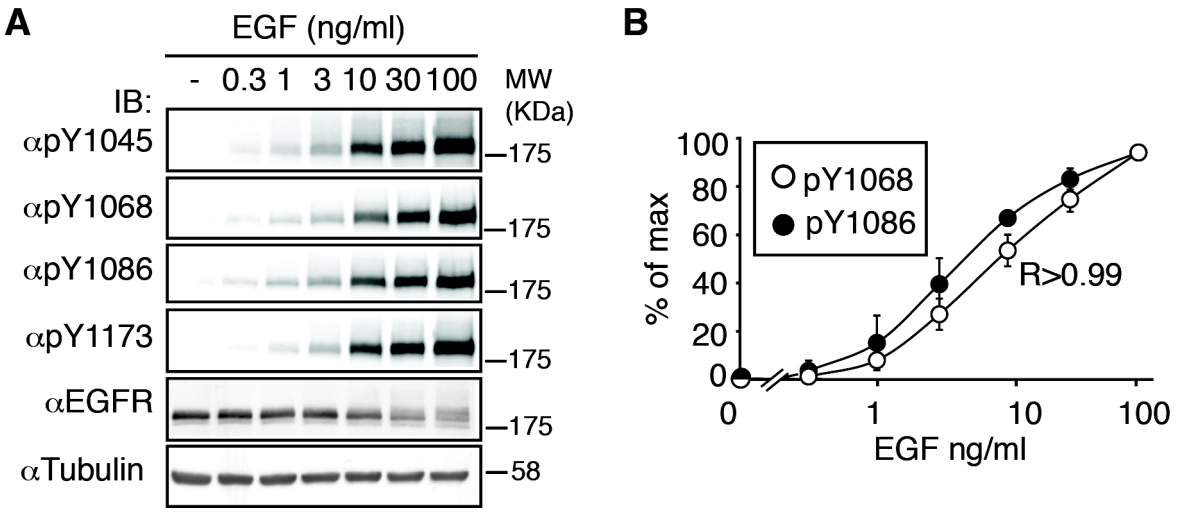
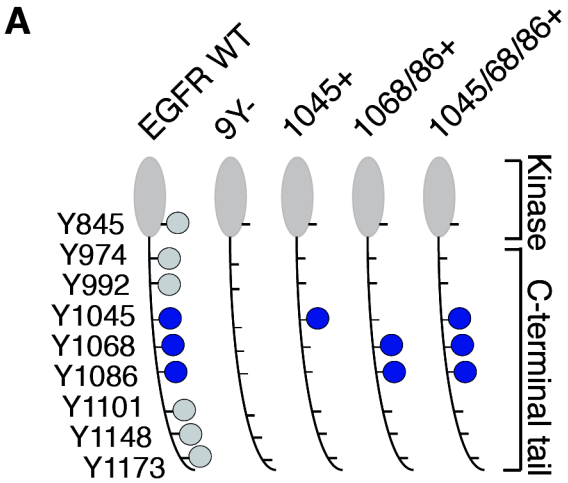


Figure 26 - Analysis of phosphorylation of specific EGFR phosphosites following EGF stimulation of HeLa cells A. HeLa cells were stimulated with different EGF doses, as indicated, and total cell lysates were analyzed by WB using the specified antibodies. pY1045-1068-1086-1173 refer to commercially available antibodies (see Materials and Methods, section Reagents) that recognized the indicated phosphosite of EGFR. αEGFR was used as control for total receptor levels. Tubulin was used as protein loading control. Blots are representative of 3 experimental repeats. Molecular weight markers are shown on the right. **B.** HeLa cells were stimulated with increasing EGF concentrations and total cell lysates were analyzed by DELFIA. Only two specific phosphotyrosine antibodies, pY1068 and 86, were functional in the DELFIA assay. Results are expressed as mean ± SD calculated from at least three independent experiments.

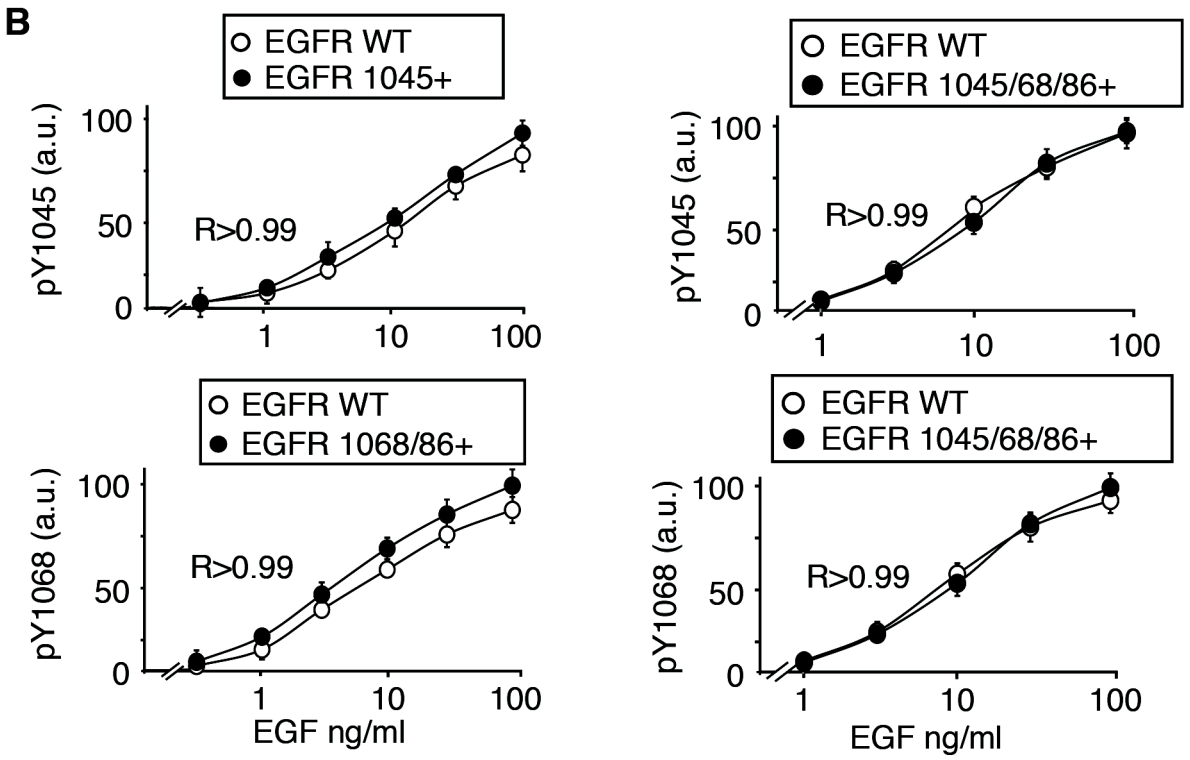
To verify this result, we employed EGFR add-back mutants. EGFR can be phosphorylated on nine tyrosine residues in the intracytoplasmic tail (**Fig. 27A**). We mutagenized all nine tyrosines into phenylalanine (a residue that is structurally similar to tyrosine but cannot accept a phosphate moiety) thus creating a pY-null EGFR backbone (the 9Y- mutant; **Fig. 27A**). We then added back, singularly or in combination, the three tyrosines that are important in the process of EGFR ubiquitination, i.e. Y1045, Y1068, Y1086 to create the add-back mutants EGFR-1045+, EGFR-1068/86+ and EGFR-1045/68/86+ (EGFR-3Y+; **Fig. 27A**). At two minutes of EGF stimulation, the add-back mutants exhibited gradual EGF-dependent tyrosine phosphorylation (**Fig. 27B**). For all mutants, we showed that the dose response curves of pY1045, pY1068 and pY1086 was the same in the context of the add-back mutants EGFR-1045+, EGFR-1068/86+ or EGFR-3Y+, as in EGFR wildtype (WT), and was not influenced by the number of tyrosine residues present in the EGFR tail.

These data suggest a simple model for EGFR early phosphorylation, based on the independent phosphorylation of each single tyrosine.

Figure 27: Add back mutants to study tyrosines important in the process of EGFR ubiquitination A - Schematic representation of human wildtype (WT) EGFR, the EGFR-pY null mutant (9Y-), and the EGFR add-back mutants EGFR-1045+, EGFR-1068/86+ and EGFR-1045/68/86+ (EGFR-3Y+). The intracellular domain (kinase domain and C-terminal tail) of the EGFR is shown, with the position of the relevant residues. Critical tyrosine residues involved in Cbl/Grb2 binding are indicated in blue, while the other tyrosine residues in the EGFR tail are depicted in grey.



B - NR6 mouse cells endogenously devoid of EGFR, were transfected to express the indicated add-back mutants or the human WT EGFR. Cells were then stimulated with increasing EGF doses and analyzed by WB using antibodies directed against the indicated pY sites (pY1045 and pY1068). Quantitation of WB experiments by densitometry was performed to analyze tyrosine phosphorylation of individual phosphosites in the EGFR add-back mutants as a function of EGF dose. Phosphorylation of EGFR-WT is also shown in each graph for comparison. Results are expressed as absolute values in arbitrary units (a.u.) where the max value is WT at 100 ng/ml EGF.



4 - Role of phosphatases

It has been recently shown that phosphatases are active already at two minutes after EGF stimulation (Kleiman, Maiwald et al. 2011). To confirm these data also in our cellular model and experimental setting, we performed experiments with a clinical-grade EGFR kinase inhibitor, Gefitinib, in HeLa cells. This drug is a sub-micromolar inhibitor of EGFR TK activity and is known to inhibit EGFR autophosphorylation by competing with ATP molecules. Gefitinib has been already used in literature (Kleiman, Maiwald et al. 2011) to study the kinetics of receptor dephosphorylation by exploiting its strong inhibitory effect on EGFR phosphorylation. We first tested the optimal concentration of Gefitinib in HeLa cells to inhibit EGFR, in a titration experiment. Cells were either treated for two minutes with 100 ng/ml EGF alone or in combination with the inhibitor at different drug concentrations, and EGFR phosphorylation was assessed by WB analysis (**Fig. 28**). The inhibitory effect on EGFR phosphorylation was already evident at the lowest concentration of Gefitinib tested (1 μM), although a complete inhibition was observed at 5 μM. To be consistent with the published experiment (Kleiman, Maiwald et al. 2011), in which Gefitinib was investigated as a de-phosphorylation agent for EGFR in different cell lines, we chose to stay in line with the experimental set up used and to treat cells with 10 μM Gefitinib.

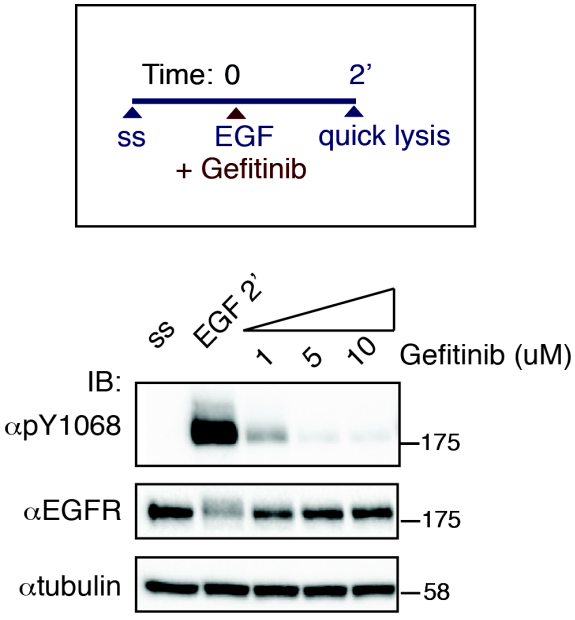


Figure 28: Effect of the different doses of the TK activity inhibitor Gefitinib on EGFR phosphorylation. Top, scheme of the experiment: HeLa cells were serum starved (ss) and then stimulated with 100 ng/ml EGF for 2 minutes in presence or absence of the indicated concentrations of Gefitinib. Cells were then quickly lysed in RIPA buffer at 4°C to avoid further dephosphorylation of the EGFR during washing steps and assessed by WB. Left, scheme of the experiment. Right, WB evaluation of the inhibitory effect of Gefitinib on EGFR phosphorylation using the specific EGFR phosphosite antibody pEGFR-1068. Total EGFR was also detected and tubulin was used as a loading control

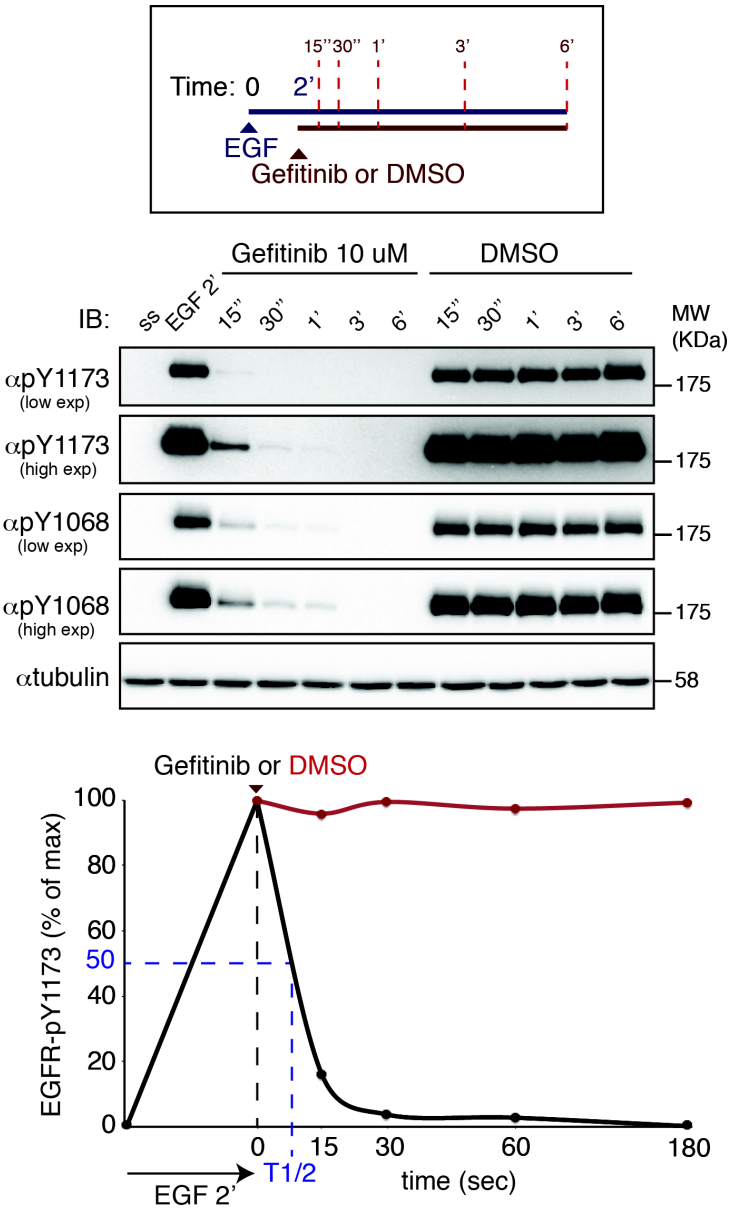


Figure 29: Effect of Gefitinib on EGFR phosphorylation. Top, scheme of the experiment. Serum starved (ss) HeLa cells were stimulated with EGF 100 ng/ml and after 2 minutes were treated with Gefitinib (10 μM) for the indicated time before processing samples for WB analysis. Parallel DMSO treatment was included as a control. Middle panel, lysates were analyzed by WB using the indicated phosphospecific antibodies (anti-pY1068 and anti-pY1173). anti-tubulin was used as a loading control. Different exposures of the blot are shown (low exp and high exp). Bottom panel, quantitation of the pEGFR-1173 WB shown in the middle panel. The half-life of the phosphorylated receptor was calculated with the formula $t_{\frac{1}{2}} = t * \frac{\ln(2)}{\ln(\frac{N_0}{N_t})}$ where t is the time after Gefitinib addition and N is the value of EGFR phosphorylation, reported to its maximum level. The half-life was found to be around 8 seconds.

After having assessed the optimal Gefitinib dose, we stimulated cells with EGF (100 ng/ml) for two minutes to induce EGFR phosphorylation, and then treated stimulated cells with 10 μM Gefitinib or DMSO alone for different lengths of time before processing cells for WB analysis (**Fig. 29**). We confirmed that Gefitinib is able to dramatically decrease EGFR phosphorylation in a very short time interval. We also estimated the half-life for EGFR-pY to be around 8 seconds in our cell model system (**Fig. 29**, lower panel), consistent with previous observations in H1666, MCF10A, HepG2 and HeLa cells (Kleiman, Maiwald et al. 2011).

These data confirm that phosphatases do indeed have a significant role early in the process of EGFR activation, being very active already at 2 minutes of EGF stimulation.

However, to test whether phosphatases are able to counteract the EGFR kinase and influence EGFR phosphorylation behavior at 2 minutes after EGF stimulation, we tested the EGFR dose response curve upon inhibition *in vivo* of tyrosine phosphatases with the chemical compound sodium orthovanadate (Na_3VO_4). First, we verified that Na_3VO_4 was able to effectively block phosphatase activity in the cell. Cells were pre-treated for 2 hours with 200 µM Na_3VO_4, as already performed in the lab, and then were stimulated with high doses of EGF for two minutes. Anti-pY blot on total cellular lysate revealed that total protein phosphorylation is increased when cells were treated with EGF plus Na_3VO_4, compared to

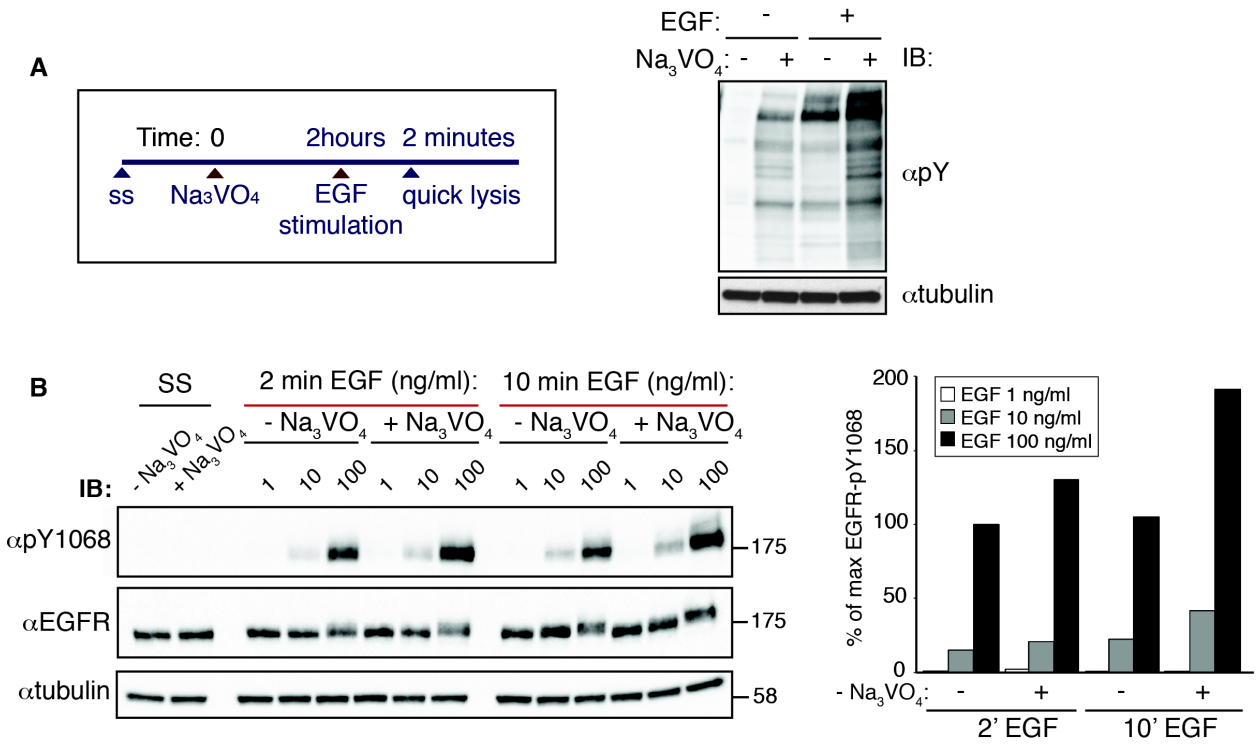


Figure 30: Effect of phosphatases inhibition by orthovanadate treatment on EGFR phosphorylation. A - Left, scheme of the experiment. HeLa cells were serum starved (SS) and treated or not for 2 hours with 200 µM sodium orthovanadate phosphatase inhibitor (Na_3VO_4). DMSO was used as a solvent control. EGF 100 ng/ml stimulation was then performed for 2 minutes and total levels of protein tyrosine phosphorylation were analyzed by WB, using the anti-total pY antibody on total cellular lysate (right panel). Tubulin was used as a loading control. IB is representative of 3 biological replicas. Note that Na_3VO_4 treatment significantly increased the level of total phosphorylated proteins both in the presence or absence of EGF. **B -** HeLa cells were serum starved (SS) and either treated or not with 200 µM Na_3VO_4 as in A. Cells were then stimulated with EGF at 1, 10 or 100 ng/ml for 2 or 10 minutes. EGFR phosphorylation was analyzed by WB using the EGFR-pY1068 antibody (left panel). Total EGFR was also detected and tubulin was used as a loading control. IB is representative of 2 biological replicas. Quantitation of the anti-pY1068 WB was performed using densitometry (right panel).

EGF alone, suggesting that phosphatases are inhibited by the treatment (**Fig. 30A**).

Next, we repeated the experiment, stimulating cells with different EGF doses (i.e. 1, 10, 100 ng/ml) for 2 or 10 minutes. By WB analysis with anti-EGFR-pY1068 antibody, we observed that when phosphatases were inhibited by Na_3VO_4 treatment, EGFR phosphorylation at 2 minutes after stimulation was only slightly increased at 100 ng/ml (**Fig. 30B**). The effect was more clear after 10 minutes of EGF stimulation, at which time even at 10 ng/ml EGF a major effect on the early time-point is evident. At 100 ng/ml EGF, we could score a 1.8-fold increase in the levels of phosphorylated EGFR (**Fig. 30B**, right panel). Our data confirmed a role of phosphatases in EGFR phosphorylation, which becomes more evident at later time points.

Therefore, in the formalization of a mathematical model for EGFR activation, we could not ignore the activity of phosphatases described above. We thus considered a value for EGFR dephosphorylation due to phosphatase activity equal to $\frac{\ln(2)}{t_{1/2}} = 0.016\ s^{-1}$ (as reported in (Kleiman, Maiwald et al. 2011)), in agreement with the experimental observation.

5 - MPM: a simple model for early EGFR phosphorylation

With the aim of obtaining an **Multisite Phosphorylation Model** (**MPM**) that takes into account only the steps of EGFR phosphorylation and dephosphorylation, we considered the enzyme kinetics that EGFR follows. Since kinases respect a standard Michaelis-Menten kinetics, we built a schematic representation of the enzymatic reactions that lead to EGFR phosphorylation in an active dimer (**Fig. 31A**). The TK domain of an EGFR moiety binds reversibly to the C-terminal tail of a partner EGFR molecule - which carries the phosphorylatable tyrosines - with the binding/unbinding rate constants k_{on} and k_{off}, respectively. After addition of the phosphate group with rate constant k_{cat}, the TK dissociates from the substrate. Experimental measurements of these rate constants were obtained from the

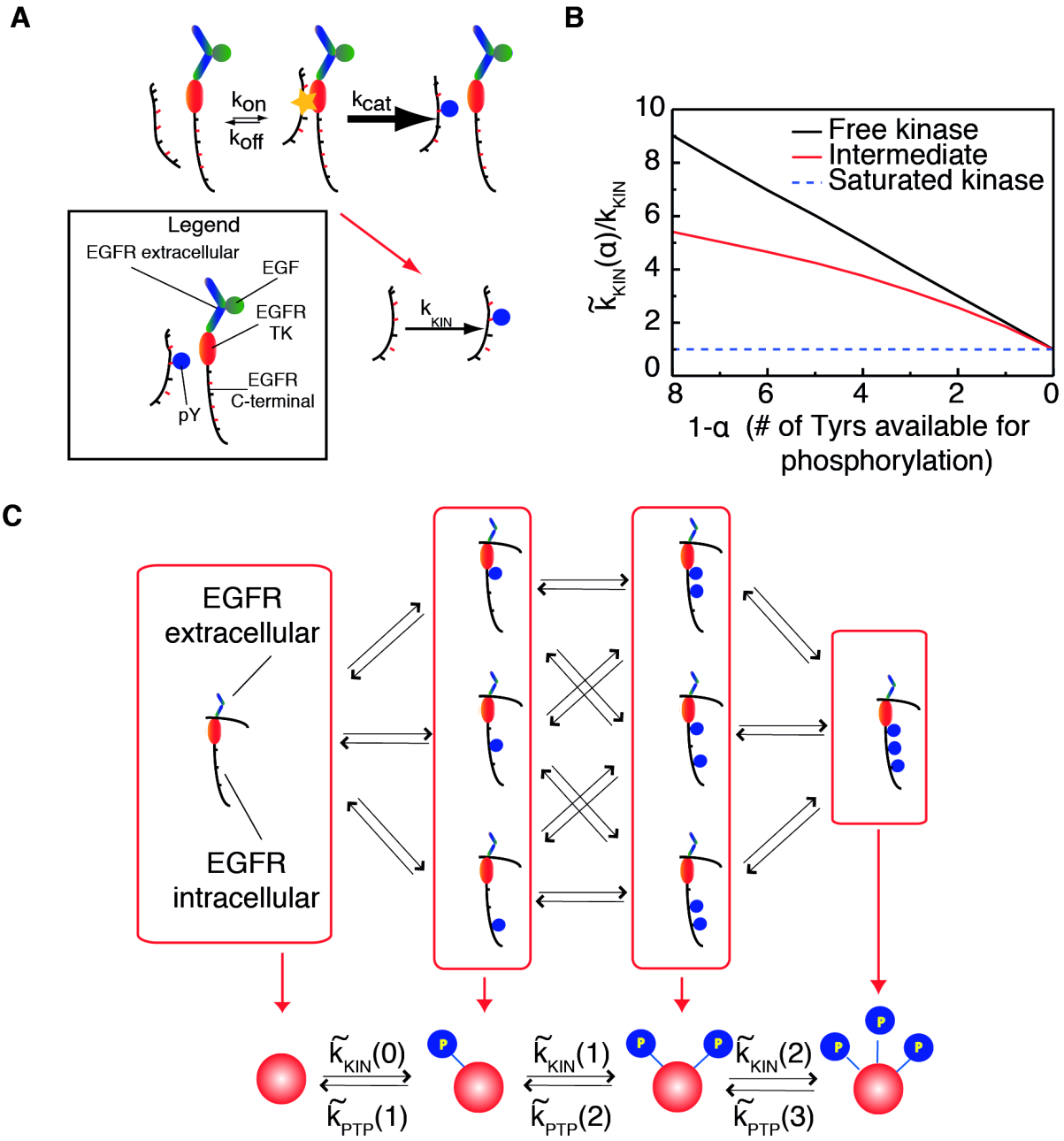
literature (Levkowitz, Waterman et al. 1999; Reynolds, Tischer et al. 2003; Grovdal, Stang et al. 2004; Zhang, Gureasko et al. 2006; Kleiman, Maiwald et al. 2011; Monast, Furcht et al. 2012; Endres, Das et al. 2013; Roda-Navarro and Bastiaens 2014; see Appendix 1). To simplify this two-step reaction, we modeled phosphorylation as a first-order reaction characterized by k_{kin}, the basic rate constant that governs the addition of a phosphoryl group in the absence of competing tyrosine. Similarly, we included the subtraction of one phosphate group by phosphatases considering the k_{PTP} constant.

Two limiting regimes (free and saturated) can be identified for the reaction catalyzed by kinases (and phosphatases), depending on the stability of the complex between the catalytic subunit and its substrate. We wanted to determine which regime the EGFR follows in the first 2 minutes after stimulation.

In collaboration with the Systems Biology group, we ran simulations of a receptor that carries 9 identical tyrosine residues. These simulations highlighted that:

- if kinases (or phosphatases) work under free-enzyme regime, each tyrosine contributes equally to the rate of the overall EGFR phosphorylation, without considering the number of available sites. Therefore, the rate of EGFR phosphorylation decreases linearly with the number of tyrosines accessible for phosphorylation (**Fig. 31B**, black solid line). Moreover, tyrosine residues do not compete for the kinase domain and this can occur only if the kinase:tyrosine complex is unstable.

Figure 31 (on the right): Modeling EGFR phosphorylation. A - Schematic representation of the enzymatic reactions that lead to EGFR phosphorylation in an active dimer. Top: Michaelis-Menten reaction whereby the tyrosine kinase (TK) domain of an EGFR molecule binds reversibly to the C-terminal tail of a partner EGFR molecule with the binding/unbinding rate constants k_{on} and k_{off}, respectively. After adding one phosphate group, the TK domain dissociates from the substrate with the rate constant k_{cat}. Bottom, the simplified reaction scheme for Tyr phosphorylation is characterized by k_{kin}. **B-**Two limiting regimes (free and saturated) can be identified for the reaction catalyzed by kinases, depending on the stability of the complex between the catalytic subunit of EGFR and its tyrosine target substrate. To simulate the reactions schematized in A, k_{cat} and k_{off} were fixed to 10 and 1, respectively, and k_{on} was varied. Simulations were performed for an EGFR molecule with 9 identical tyrosines. This graph shows the dependence of the rate constant $k(\alpha)$ on the number of pYs (α=0,1,…,8) present in the receptor tail, in units of k_{KIN} (the basic rate constant that governs the addition of a phosphoryl group in the absence of competing tyrosine residues). For $(k_{on}+k_{off})/k_{cat}=10^{-2}$ (blue dashed line), k_{cat} controls the rate-limiting step, a condition that we call the saturated-enzyme regime.



For $(k_{on}+k_{off})/k_{cat}=10^4$ (black line), k_{on} controls the rate limiting step, a condition that we dubbed the free-enzyme regime. For intermediate values $(k_{on}+k_{off})/k_{cat}=10$ (red line), there is no single limiting step. (refer also to the main text, Results Section 5). **C** - Wiring diagram of the MPM. Top, phosphorylation of individual Tyr residues (blue circles) occurs independently of the phosphorylation of the other Tyr residues. This results in a branched wiring diagram, in which each phosphorylation event occurs with the same probability. Bottom, EGFR molecules (red circles) with the same total number of phosphoryl groups (blue circles) are grouped together to generate a linear chain of increasingly phosphorylated EGFRs. Only the three Tyr residues relevant for EGFR ubiquitination are shown (i.e., Y1045, Y1068, Y1086). $\tilde{k}_{KIN}(\alpha)$ and $\tilde{k}_{PTP}(\alpha)$ are the rates of addition and subtraction of one phosphoryl group from an EGFR molecule that carries α phosphorylated Tyr residues.

- in an alternative saturated regime in which the kinase (or phosphatase) is predominantly bound to the substrate, the phosphorylation rate of EGFR does not change with the number of non-phosphorylated tyrosine residues (**Fig. 31B**, blue dashed line). In this case, each of the 9 residues are phosphorylated at the same rate as only one residue, indicating that the phosphorylation rate per tyrosine residue decreases with the number of available sites.

As we discussed previously, we observed that individual tyrosine residues are phosphorylated independently of each other (**Fig. 27B**) suggesting that that they do not compete for the kinase domain. This situation supports the scenario in which the rate-limiting step is the kinase:tyrosine binding constant and the kinase is expected to be mostly free (free-enzyme). The same applies for the phosphatases, whose activity at 2 minutes after EGF stimulation has been also implemented in our models. We therefore chose for our model the free-enzyme regime for both kinases and phosphatases.

In the MPM model, the rate constant of EGFR phosphorylation increases with the number of tyrosine residues, and this also increases the number of variables in the model. In order to reduce the number of variables, we introduced a simplification: receptors with an equal number of phosphosites (pYs) [i.e., unphosphorylated EGFR (R_0), 1pY-EGFR (R_1), 2pY-EGFR (R_2), etc.] were grouped together, regardless of the specific identity of the pY. This simplification is justified by the observation that, in the first two minutes of EGF simulation, different sites have similar phosphorylation kinetics (Olsen, Blagoev et al. 2006). We schematically envisioned our system with a wiring diagram of the MPM (**Fig. 31C**), in which each phosphorylation event occurs with the same probability: EGFR molecules (red circles) with the same total number of phosphoryl groups (blue circles) are grouped together to generate a linear chain of increasingly phosphorylated EGFRs (**Fig. 31C**, bottom line). Only the three tyrosine residues relevant for EGFR ubiquitination are shown (i.e., Y1045, Y1068, Y1086).

To compare the pattern of EGFR phosphorylation reproduced by the MPM with experimental data, we required a relationship between EGFR phosphorylation and EGF concentration. To keep this model simple, we initially ignored the numerous reactions that contribute to EGFR activation - that will be accounted for in a more complex model (the Early Activation Model, see Results Section 7) - and considered only a phenomenological law to link EGFR phosphorylation and EGF concentration (the "Hill function", see next Section 5.1).

5.1 - The Hill function

The properties and mode of action of a wide range of enzymes can be studied by the classical Michaelis-Menten equation. However, some enzymes display kinetic properties that cannot be fully described by this equation. We have recently shown that the dose-response curves for EGFR phosphorylation and ubiquitination display different degrees of sigmoidicity (Sigismund, Algisi et al. 2013) and both of them can be best approximated by the Hill function.

The Hill equation was first introduced by A.V. Hill in 1913 to describe the equilibrium between oxygen tension and the saturation of hemoglobin (Hill 1913). Since then, the Hill equation has been widely used to analyze quantitative drug or ligand–receptor relationships. Many biochemical models have exploited the potential of Hill equation to describe non-linear dose response relationships.

The general Hill equation is represented by:

$$\theta = \frac{[L]^n}{[C]^n + [L]^n}$$

where:

- [L] is the free ligand concentration
- C is the ligand concentration occupying half of the binding sites

- n is the Hill coefficient and describes biochemical properties, depending on the context in which the Hill equation is being used.

The Hill functions that best reproduce the experimental EGFR phosphorylation and ubiquitination curves have Hill coefficients (n_H) of 1 and 3, respectively.

More precisely, to model EGFR phosphorylation, we used the Hill function:

$$k_{kin} = k_{kin}^{max} \frac{[EGF]^n}{[EGF]^n + J^n}$$

where:

- [EGF] is the EGF concentration
- J is the EGF concentration where k_{kin} is half-maximal

n is the Hill coefficient and represents how steeply k_{kin} increases with EGF.

This equation defines a sigmoidal curve in an interval between 0 and k_{KIN}^{MAX}, which is the maximal kinetic activity of EGFRs.

From the fitting of the experimental dose-response curves of EGFR-WT and the EGFR-3Y+ add-back mutant, we identified n and J. Experimental points were obtained both by WB densitometric analysis and the DELFIA assay on stimulated HeLa cell lysates. After fitting these parameters, the MPM reproduced closely the experimental dose-response curve of EGFR-pY (**Fig. 32A**).

Since we used the free-regime for EGFR phosphorylation, the result is independent of the number of tyrosine residues when the curves are normalized to their maximal value (**Fig. 32B**). Moreover, the model correctly attributed a 3-fold increase to the total Tyr phosphorylation of EGFR-WT compared to the total Tyr phosphorylation of the EGFR-3Y+ mutant, for every value of EGF (**Fig. 32C**), as expected from a substrate carrying only 3 out of 9 accessible catalytic residues.

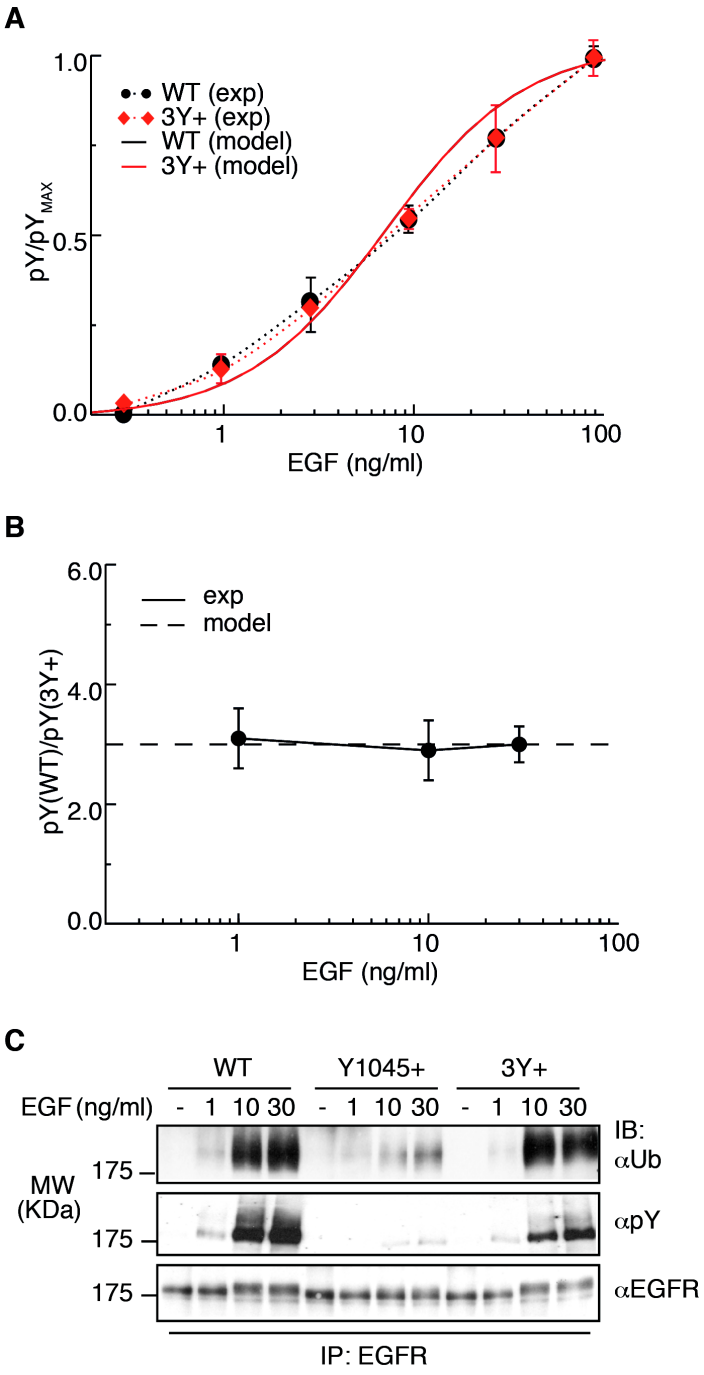


Figure 32: Modeling EGFR phosphorylation: EGFR-WT *vs.* EGFR-3Y+. A - A comparison between EGFR-WT and -3Y+ phosphorylation, computed by the MPM in the free-kinase/free-phosphatase regime as a function of EGF (model, solid lines), or determined experimentally (dashed lines: experimental data taken from Fig.14). EGFR phosphorylation was normalized to the maximum pY value (pY/pYmax). **B -** A comparison of the ratio of total pY of EGFR-WT and EGFR-3Y+, as a function of EGF concentration (at 2 min), calculated by the MPM (dashed line) or determined experimentally (exp; solid line, data taken from blot in C). **C -** Dose response analysis of Ub and pY of EGFR-WT, and the Y1045+ and 3Y+ add-back mutants. NR6 cells stably expressing the indicated mutants were stimulated for 2 min with increasing concentrations of EGF, as indicated. Lysates were subjected to IP and IB as shown. The ratio of total pY of EGFR-WT *vs.* EGFR-3Y+ was determined by densitometry analysis and is shown in B.

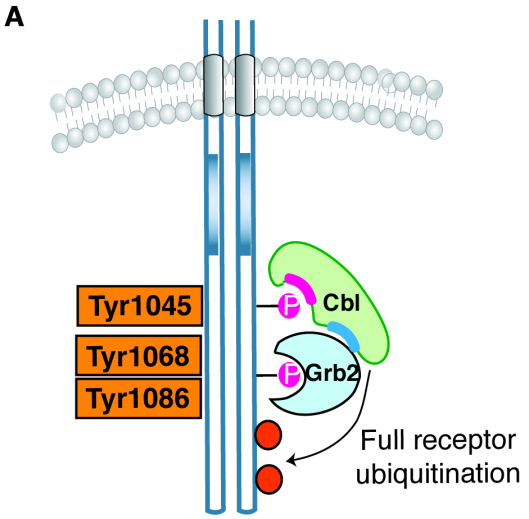
In conclusion, we have developed a model, the MPM, which faithfully reproduces the distribution of EGFR phosphorylated species, as a function of EGF concentration, at two minutes of stimulation.

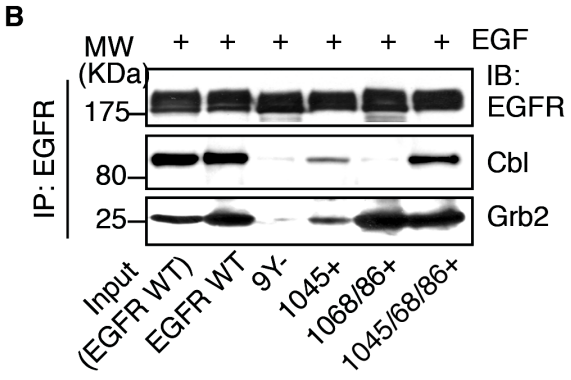
6 - Modeling EGFR ubiquitination

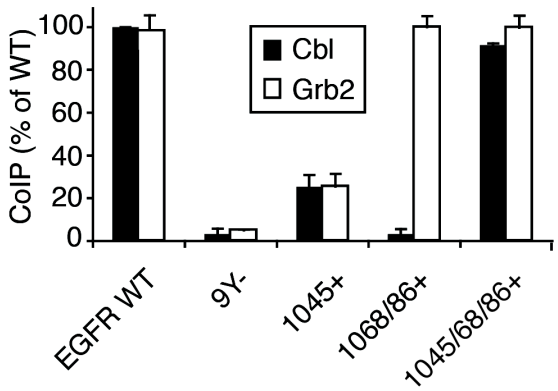
We then progressed to add to the MPM the processes leading to EGFR ubiquitination. This covalent modification is catalyzed by the E3 ligase Cbl, which can bind to the EGFR directly through pY1045, or indirectly through the adaptor protein Grb2 that binds pY1068/86 (see also Introduction Section 3.2, **Fig. 33A**). As previously elucidated (page 31), Cbl family proteins consist of three homologues: c-Cbl, Cbl-b and Cbl-c. By q-PCR analysis, we showed that in HeLa cells only c-Cbl and Cbl-b are expressed, the latter being expressed at levels ~10% of the level of the former (Pascolutti et al, unpublished). With an RNAi approach, we demonstrated that c-Cbl is the major E3 ligase responsible for EGFR ubiquitination in HeLa cells (Sigismund, Algisi et al. 2013), since the KD of Cbl-b is only able to slightly decrease EGFR-Ub levels, whereas KD of c-Cbl is able to abrogate it (Sigismund, Algisi et al. 2013). Thus, we described only c-Cbl (hereafter referred to as Cbl) in the model.

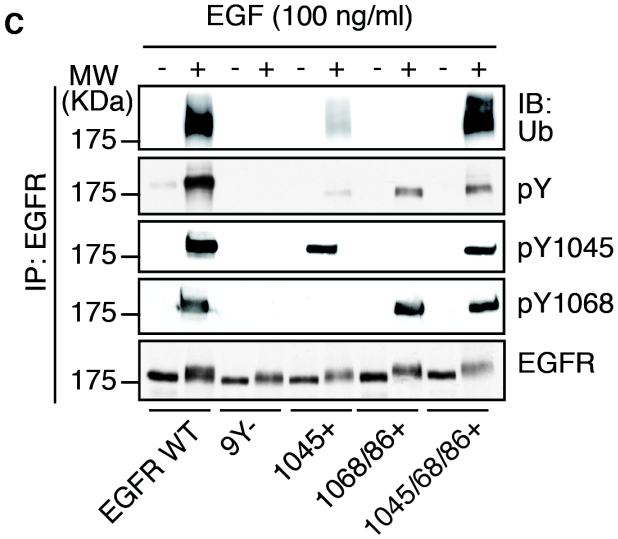
By the use of EGFR phospho-mutants, we have also shown that the binding of Cbl to its phosphosites is necessary and sufficient for EGFR ubiquitination, and that the EGF dose-dependent Cbl:EGFR interaction *in vivo* displays a threshold-like profile, analogous to that of EGFR ubiquitination (Sigismund, Algisi et al. 2013). Based on this evidence, we further simplified the model by assuming that ubiquitination is simply proportional to the amount of Cbl-bound EGFR. Thus, we modeled EGFR ubiquitination as simply Cbl binding to the receptor.

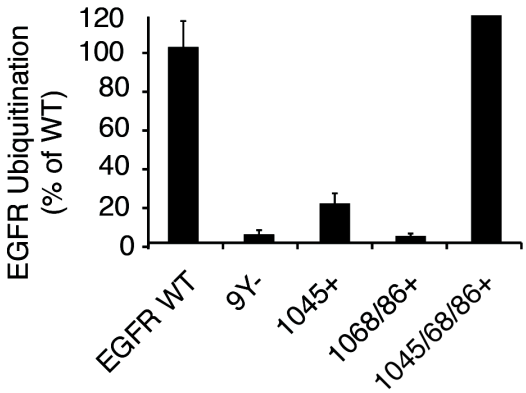
Figure 33: Cooperativity between the EGFR phosphosites pY1045 and pY1068/86 in the binding of Cbl to the receptor and in EGFR ubiquitination. A - Schematic representation of EGFR ubiquitination, as in **Fig.7**. The intracellular domain (kinase domain, shaded in blue, and C-terminal tail, unshaded) of the EGFR is represented as a long tail, where the three fundamental tyrosines (Y1045, Y1068, Y1086) for EGFR ubiquitination and Cbl binding are depicted. The E3-ligase Cbl can interact with the receptor directly through the phosphorylated-Y1045 or indirectly through the interaction with Grb2 that binds to pY1068 or pY1086. The concomitant binding of Cbl through both sites generates stable Cbl/Grb2 binding to the EGFR and increases EGFR-Ub. **B -** Left: CHO cells expressing the EGFR-WT, -9Y- mutant and the indicated add-back mutants (Fig.27A), were stimulated with EGF (100 ng/ml) for 2 minutes and lysates were subjected to IP and WB as indicated, to investigate the binding of Cbl and Grb2 to the receptor. 10% of the IP lysate was loaded alone as internal control (input). Right: quantitation of the amount of Cbl or Grb2 co-immunoprecipitating (CoIP) with EGFR based on densitometry analysis of blots, as shown on the left.**C -** NR6 cells expressing the EGFR-WT, -9Y- mutant and the indicated add-back mutants (described in Fig.27A) were stimulated or not with EGF. Lysates were subjected to IP and IB as indicated, to explore the level of ubiquitination of EGFR mutants. Right: quantitation of the level of ubiquitination of the different EGFR mutants relative to EGFR-WT based on densitometry analysis of blots, as shown on the left. In B and C, results were normalized to the amount of immunoprecipitated EGFR and are expressed as % of the values obtained in EGFR-WT cells. Results are expressed as the mean ± SD of three independent experiments.











6.1 - Cooperativity

The experimental analysis of Cbl binding to the EGFR suggested that the interaction between EGFR and Cbl is *cooperative* with the binding of Grb2 to the receptor (Sigismund, Algisi et al. 2013). Indeed, if the binding of Cbl to the EGFR were to take place with the same affinity regardless of the binding of Grb2, then EGFR-WT and EGFR-Y1045+ should be ubiquitinated to the same level and bind to a similar amount of Cbl. We tested experimentally this hypothesis by measuring the levels of EGF-induced ubiquitination in the add-back EGFR mutants. Add-back mutants were previously introduced (Results Section 3, **Fig. 27A**); briefly, EGFR gene was mutagenized in the 9 tyrosine residues in the intracytoplasmic tail, to obtain a pY-null backbone. From this backbone, single or combinations or residues were back-mutagenized to WT. Through this experimental approach, we established that binding of Cbl to the EGFR was reduced by about the 80% in the EGFR-Y1045+ mutant compared with EGFR-WT, and absent or negligible in the EGFR-Y1068/86+ mutant (**Fig. 33B**). In contrast, Cbl binding was almost indistinguishable in EGFR-WT and in the EGFR-3Y+ mutant. Clearly, we could also score a cooperative effect between pY1045 and pY1068/86 towards the association with Cbl, since the binding of the EGFR-3Y+ mutant to Cbl is far stronger than the simple sum of the Cbl interaction with pY1045 and with pY1068/86 (see quantitation in **Fig. 33B, right panel**).

Next, we experimentally determined the ability of the EGFR mutants to be ubiquitinated by Cbl. As expected, EGFR ubiquitination was completely abolished when all three tyrosines were mutagenized together in the 9Y- mutant (**Fig. 33C**). Moreover, the presence of pY1068/86 was not sufficient to allow EGFR ubiquitination. Ubiquitination of the receptor was, however, partially rescued by the presence of pY1045, but only the concomitant restoration of the three sites (3Y+) allowed for full EGFR ubiquitination (**Fig. 33C**). Again, there was a clear cooperative effect between the three Cbl/Grb2 binding sites, which allowed the EGFR to be fully ubiquitinated (see quantitation in **Fig. 33C, right panel**). Importantly,

the ubiquitination levels of the EGFR mutants reflected the levels of the *in vivo* binding of Cbl to the EGFR (compare **Fig. 33B** and **33C**).

Based on these observations, we also used the add-back mutants to investigate whether the sigmoidal behavior of the interaction between Cbl and the EGFR is maintained when the relevant pY sites (i.e. pY1045/68/86) are lost. To this aim, we performed an EGF dose-response experiment and analyzed the Cbl:EGFR interaction by co-immunoprecipitation (co-IP) in NR6 cells expressing EGFR-1045+, EGFR-3Y+ and EGFR-WT. The EGF dose-response curve for the Cbl:EGFR interaction in cells expressing the EGFR-3Y+ mutant displayed a threshold behavior similar to that of cells expressing EGFR-WT, whereas the EGFR-1045+ mutant showed a non-threshold gradual response over a range of EGF concentrations (**Fig. 34A**). We then aimed to understand whether the sigmoidal behavior of EGFR ubiquitination could be re-established going from the EGFR-1045+ mutant to the EGFR-3Y+. We compared EGF dose-response curves for ubiquitination of these mutants with that of the EGFR-WT by DELFIA analysis (**Fig. 34B**). We observed that EGFR-WT and the EGFR-3Y+ mutant behaved similarly, showing a threshold effect for receptor ubiquitination, while the EGFR-1045+ mutant displayed a non-threshold gradual response curve, consistent with the presence of one single site for Cbl binding.

Together, these data point to a scenario in which the EGFR ubiquitination threshold is due to a non-linear recruitment of Cbl-Grb2 complex to the EGFR, which in turn is dependent on cooperativity between the key phosphosites. A plausible explanation for cooperativity is that when Cbl and Grb2 are simultaneously bound to EGFR via their docking phosphosites (pY1045 and pY1068/86, respectively), they are in a state of enforced-proximity, which increases the chance of the three species (EGFR, Cbl and Grb2) binding to each other in a mass action kinetic. The fully bound trimer corresponds *in vivo* to the fully ubiquitinated EGFR. Instead, if Cbl only binds EGFR directly through pY1045, their interaction is not stabilized by Grb2 and the EGFR is only partially ubiquitinated.

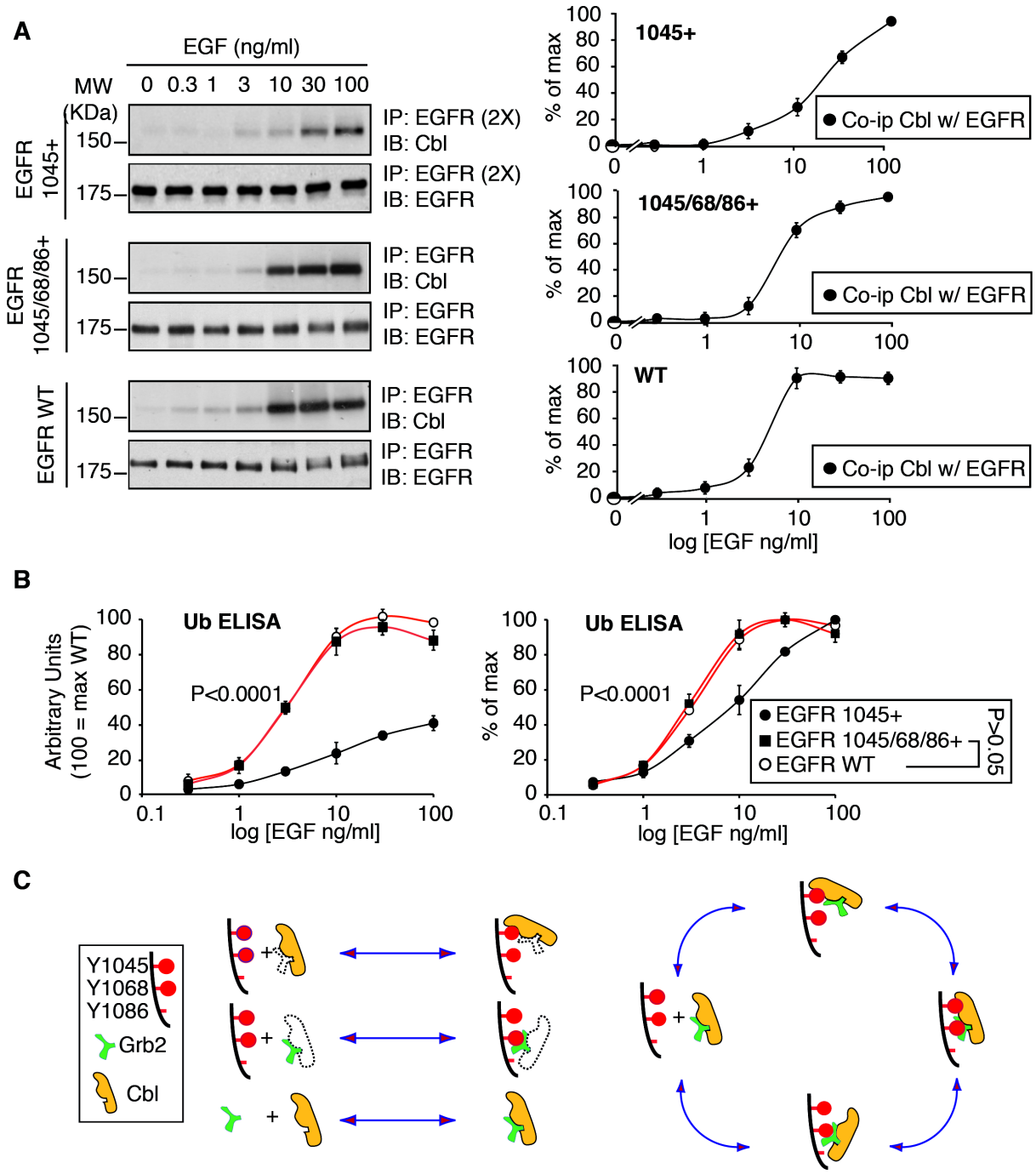


Figure 34: Comparison of the EGF dose-response behavior of Cbl-binding to the receptor and receptor ubiquitination in the EGFR mutants. A - Analysis of the ability of EGFR-Y1045+, -3Y (Y1045/68/86+) and -WT to bind Cbl. NR6 cells stably expressing the Y1045+ mutant, the 3Y+ mutant, or EGFR-WT, were stimulated with EGF for 2 minutes at the indicated EGF doses. Lysates were subjected to IP and IB as shown. Note that for the 1045+ mutant, we used 2-fold (2x) more lysate because total EGFR-Ub is impaired in this mutant. For each blot, quantitation by densitometry is shown on the right, expressed as a percentage of max. **B -** The same samples described in 'A' were subjected to the DELFIA assay (forward approach) to assess ubiquitination the EGFR mutants, using anti-EGFR intra as the capture Ab and anti-Ub (FK2) as the detection Ab. Results are expressed as a percentage of EGFR-WT ubiquitination (arbitrary units, 100 = max WT, left panel) or as a percentage of the maximal ubiquitination of each EGFR construct (% of max, right panel). Error bars indicate the SD calculated from at least three independent experiments. All p-values were calculated using the two-way ANOVA analysis. P-value inset in graphs refer to wt vs phospho-mutants. **C -** Schematized reactions involved in EGFR ubiquitination. Upper reaction: Cbl binds to pY1045 irrespective of whether it is free or in a complex with Grb2 (Grb2 is thus represented with a dotted line); middle reaction: Grb2 binds to pY1068 (or pY1086, not shown) irrespective of whether it is free or in a complex with Cbl (represented with a dotted line); bottom reaction: Cbl and Grb2 bind to each other. On the right: simplification of the consecutive reactions that lead to a stable interaction of the complex Cbl/Grb2 to the EGFR.

In our model, we interpreted the presence of cooperativity as follows: Cbl and Grb2 can form a complex that binds to the corresponding pYs on EGFR, but they can also bind EGFR individually or bind each other forming a dimer. What is fundamental in the system is that the Cbl and Grb2 binding rates (to each other and to the EGFR) are increased in the multimeric complex. All the possible combinations accounted in the model are depicted in **Fig. 34C**.

6.2 - Cbl, but not Grb2, is limiting in the system.

Assuming that enforced proximity drives cooperativity, we postulated that the binding affinity of EGFR and Cbl increases proportionally with the local increase in the availability of Grb2 and Cbl on the receptor. Under this scenario, Cbl and Grb2 concentrations become important. For what concerns Cbl, it is known that it acts as an E3 ligase in a number of cellular complexes and is highly expressed in cells. However, consistent with the notion that Cbl binds to >100 different proteins (Schmidt and Dikic 2005; Dikic and Schmidt 2007), we have verified that the majority of Cbl is engaged in stable complexes. Performing a size exclusion chromatography, Cbl mainly elutes in high molecular weight fractions, with only a minor portion of the protein eluting as the free protein (**Fig. 35A**).

We also experimentally determined total Cbl levels through a semi-quantitative WB analysis. By comparing known amounts of cell lysates with increasing amounts of bacteria-purified Cbl, we determined that Cbl is expressed in HeLa cells at ~150,000 molecules/cell (**Fig. 35B**, upper panel, see also Table 2 on page 17). This result combined with the size exclusion chromatography analysis, strongly suggests that only a minor fraction of Cbl (in the free form or engaged in Cbl:Grb2 complexes) is available for direct binding with EGFR. In order to estimate the amount of Cbl available for binding to the EGFR, we used the phosphorylated form of Cbl (Cbl-pY) as a proxy for the max amount of available Cbl because

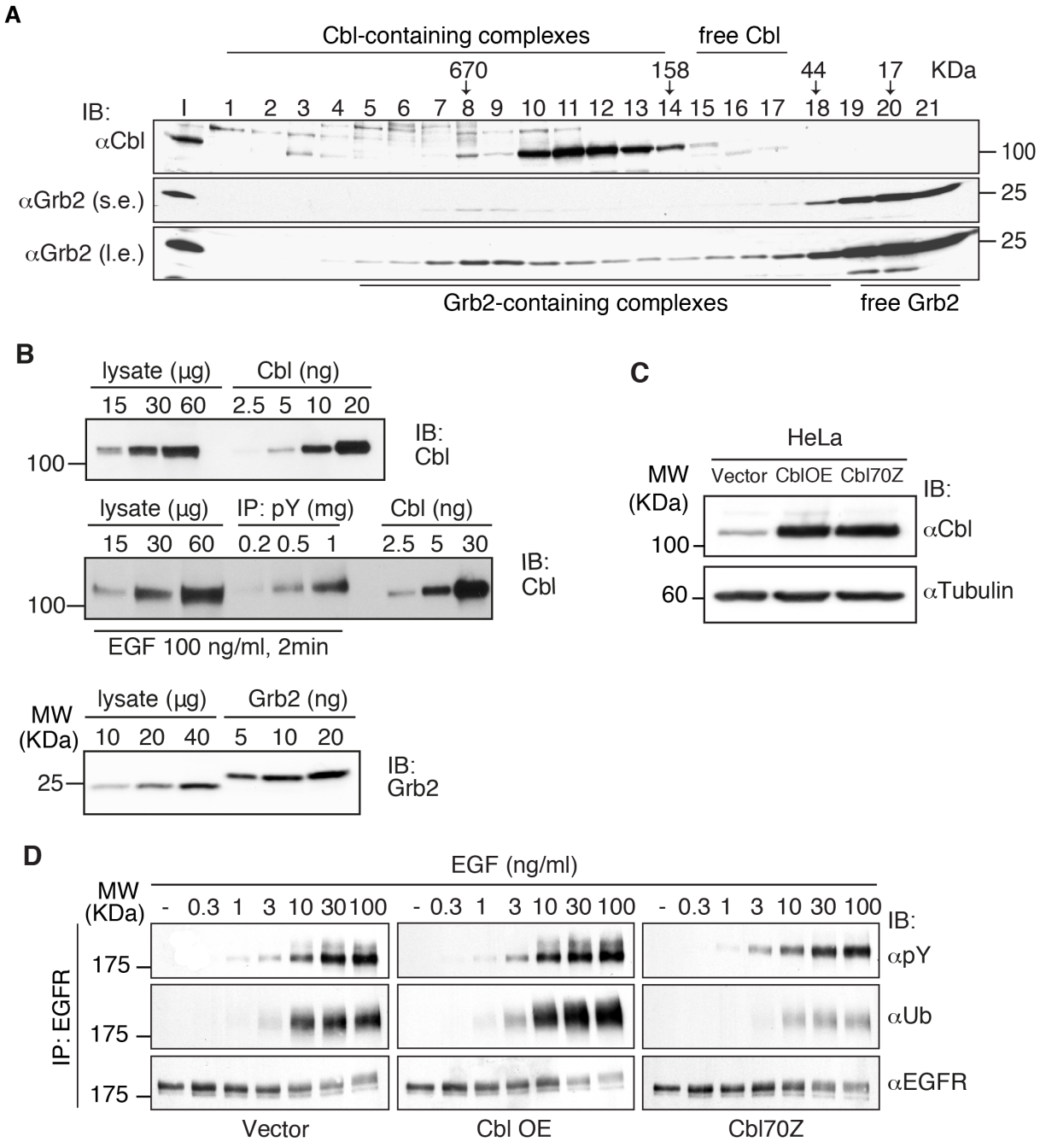


Figure 35 – The E3 ligase Cbl is limiting in the system. A. Analysis of Cbl and Grb2 elution profiles by size exclusion chromatography, performed as described (Penengo, Mapelli et al. 2006, see also Materials and Methods). Starting from a cellular lysate of HeLa cells loaded on a size exclusion column, consecutive fractions eluting from the column according to their molecular weight were collected and subjected to WB analysis. I, input (50 μg). In the Grb2 panels: s.e. and l.e., represent short and long exposure, respectively. **B.** Quantitation of Cbl and Grb2 molecules in HeLa cells. Increasing amounts of HeLa cell lysate were subjected to IB and IP as indicated and compared to increasing amounts of *in vitro* purified Cbl (upper and middle panels) or Grb2 (lower panel) proteins. For further information about the estimation of Cbl and Grb2 molecules, refer to Table 1 and Matherials and Methods, Section Protein Procedures. **C.** HeLa cells were stably transfected with Cbl-WT (Cbl overexpression, OE), Cbl dominant negative (Cbl70Z), or empty vector (Vector). Lysates were analyzed by WB as indicated. Tubulin was used as a loading control. Densitometry analysis revealed an ~7-fold increase in Cbl expression levels in Cbl-OE and Cbl70Z transfectants compared with control cells. **D.** HeLa cells transfected as in **C**, were treated with EGF at the indicated concentrations for 2 minutes. Lysates were subjected to IP and IB analysis with the indicated antibodies.

I) Cbl is phosphorylated upon binding to the EGFR; and II) only phosphorylated Cbl is competent for EGFR ubiquitination (Levkowitz, Waterman et al. 1999; Zhang, Gureasko et al. 2006). We reasoned therefore that the levels of Cbl-pY at the maximal EGF concentration might represent a reasonable approximation of the maximal amount of Cbl available for binding to the EGFR. After stimulating with a high concentration of EGF (100 ng/ml) for two minutes, the amount of Cbl-pY that was experimentally determined was ~5,000 molecules/cell (**Fig. 35B**, middle panel). This result confirmed that the amount of Cbl that is available for EGFR binding in the first 2 minutes after EGF stimulation is a minimal fraction of total Cbl molecules/cell (around 4%) and thus could be limiting.

To investigate in more detail the correlation between Cbl levels and EGFR ubiquitination in our system, we took advantage of a dominant negative form of Cbl (Cbl70Z) (Yokouchi, Kondo et al. 1999), which mimics Cbl downmodulation. We transfected HeLa cells with an empty vector as a negative control, with Cbl-WT as a control for Cbl overexpression, or with Cbl70Z, and achieved an ~7-fold overexpression of the two Cbl constructs (**Fig. 35C**). We then treated these cells with increasing EGF doses for 2 min. We observed that Cbl overexpression or downmodulation (expression of the dominant negative mutant) induced either an increase or decrease, respectively, in the levels of EGFR ubiquitination (**Fig. 35D**). This observation corroborates the hypothesis that Cbl is rate-limiting in the EGFR ubiquitination process, since the modulation of its total levels affects the net EGFR-Ub output.

Using an analogous approach to that used for Cbl, we also estimated a total number of ~1,000,000 Grb2 molecules/cell through the comparison of total cell lysate with bacteria-purified Grb2 by WB analysis (**Fig. 35B** - lower panel). Moreover, by size exclusion chromatography, we showed that Grb2 elutes as a single monomeric species (**Fig. 35A**), indicating that the majority of the Grb2 molecules are free or form very unstable complexes.

The *in vivo* association of Grb2 with the EGFR add-back mutants (**Fig. 33B**) supports the idea that, in living cells, the majority of Grb2 exists in a free form, while Cbl is

predominantly present in complexes that also contain Grb2. The EGFR-1045+ mutant, which binds directly Cbl but not Grb2, co-immunoprecipitated with similar amounts of Cbl and Grb2 (**Fig.33B**). In contrast, the EGFR-1068/86+ mutant, which binds directly to Grb2 but not to Cbl, co-immunoprecipitated with Grb2 as efficiently as EGFR-WT or the 1045/1068/1086+ mutant, while it did not co-immunoprecipitate with Cbl.

Therefore, in our model, we assumed that all measured Grb2 is available for binding to the EGFR and/or to Cbl. The above results indicate that Cbl is rate-limiting in the EGFR ubiquitination reaction, compared with EGFR and Grb2.

6.3 - A probabilistic hypothesis

Cooperativity is necessary, but not sufficient, to generate the threshold. Indeed, the concomitant presence of two key phosphorylated tyrosines is also required, and this occurs via a purely probabilistic process. Since tyrosine residues on EGFR are phosphorylated independently of each other (**Fig. 27**), this implies that while the probability of individual tyrosyne phosphorylation increases gradually with the concentration of ligand, the probability of having two, three, or more phosphorylated tyrosine residues on the same EGFR moiety, does not. Instead it increases non-linearly. Thus, at low doses of ligand the receptor is scarcely phosphorylated so the probability of having two out of three key tyrosines phosphorylated is negligible, causing an unstable binding of the Cbl/Grb2/receptor complex. Instead, as the concentration of EGF increases, this probability follows an exponential growth, allowing a stable (cooperative) binding of the complex and full ubiquitination of the receptor (**Fig. 36**). Summarizing, we have identified three main features to integrate into the model:

1. Cooperativity between Cbl:Grb2 and the EGFR;
2. "Competent" Cbl is present at limiting concentrations;
3. The non-linear availability of EGFR molecules carrying at the same time the two key pY (1045 and 1086/86) sites required for Cbl/Grb2 binding, as EGF concentrations increase.

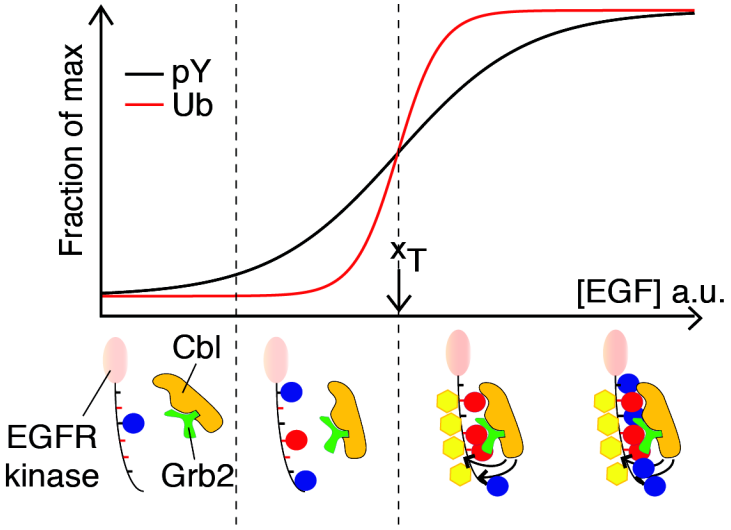


Figure 36: The probabilistic hypothesis that contributes to the generation of the EGFR-Ub threshold. At low EGF, the EGFR is scarcely phosphorylated. Increasing the EGF concentration results in an exponential increase in the probability of having two residues phosphorylated on the same moiety, therefore, allowing a stable interaction of the complex Cbl/Grb2 to the receptor and permitting full EGFR ubiquitination.

6.4 - MPM-B: a model of EGFR ubiquitination

Before proceeding with the construction of a complete model for EGFR activation, we implemented the MPM and used it as an input for the formalization of Cbl and Grb2 binding. We obtained a model of EGFR ubiquitination designated as **MPM plus Binding** (**MPM-B**). We computed the model assuming that all species are at steady state in the absence of EGF; at time zero, we simulated the addition of EGF into the system and computed the values of EGFR ubiquitination and phosphorylation after 2 minutes. MPM-B accurately reproduced the behavior of EGFR-Ub/-pY that was observed *in vivo*. Moreover, it also reproduced in quantitative detail the increase in steepness (**Fig. 37A**) of the ubiquitination curve *vs*. the phosphorylation curve for EGFR-WT, as shown by the normalized curves. MPM-B also faithfully reproduced the experimentally determined dose-response curves for ubiquitination in conditions of Cbl overexpression or downmodulation (**Fig. 37B-C**), both when curves were normalized to their max (**Fig. 37B**) or to WT ubiquitination (**Fig. 37C**). In this latter case, however, simulated and experimental curves were superimposable only if we assumed a 3-fold increase in Cbl overexpression, instead of the 7-fold increase measured experimentally (**Fig. 35C)**. Thus, the model suggests that the 7-fold increase in Cbl expression that was obtained experimentally, does not lead to a real 7-fold increase in Cbl available for EGFR binding. Rather, the model suggests that only ~1/3 of the overexpressed Cbl molecules are capable of binding the receptor, supporting the notion that not all Cbl is available for EGFR ubiquitination.

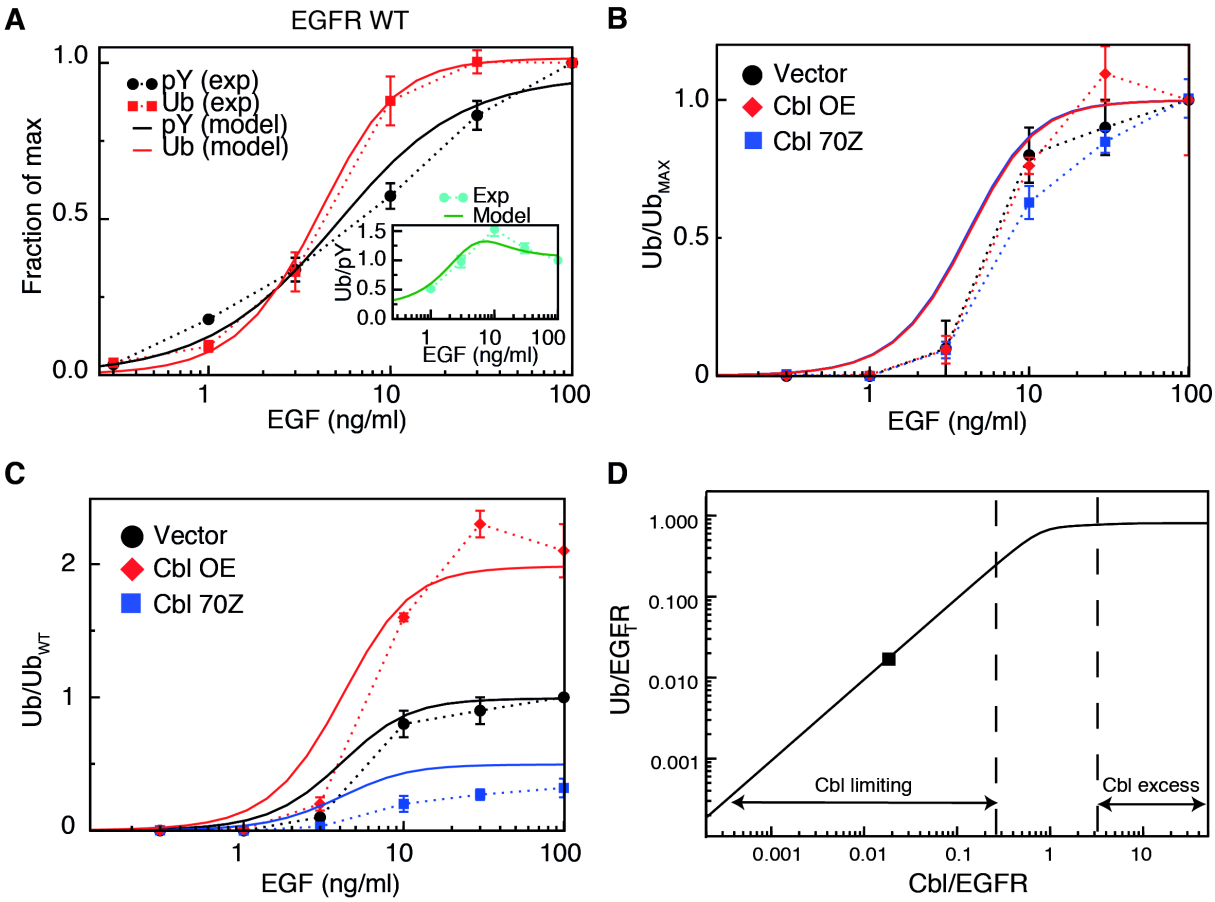


Figure 37: The MPM-B reproduces the EGFR phosphorylation and ubiquitination dose-response curves. A - Comparison of experimental (exp; dashed lines) and MPM-B modeled (solid lines) phosphorylation (pY) and ubiquitination (Ub) dose-response curves for EGFR-WT. Experimental data are expressed as normalized to the maximum value for both EGFR-Ub or -pY. Inset shows the ratio of ubiquitination to phosphorylation as a function of EGF concentration for experimental and modeled data. **B-C** - Comparison of experimental (dashed lines) and modeled (solid lines) EGFR ubiquitination dose-response curves under conditions of Cbl overexpression (OE) or downmodulation (Cbl70Z). Experimental data were taken from **Fig. 21D** in which HeLa cells were transfected with empty vector (vector), Cbl-WT for overexpression, or the dominant negative Cbl70Z for downmodulation. Data were obtained through quantitation of EGFR ubiquitination by densitometric analysis of WBs shown in Fig. 21D. EGFR ubiquitination is expressed, for each condition, as normalized to their own maximum (Ub/Ub_{MAX}, **B**) or to the maximum value obtained in the empty vector control (Ub/Ub_{WT}, **C**). **D** - Dependence of EGFR ubiquitination [expressed as normalized to the total EGFR concentration (Ub/EGFR)] on the total Cbl concentration [expressed as normalized to the total EGFR concentration (Cbl/EGFR)]. The black square indicates the amount of total active Cbl that we have experimentally measured (see Section 6.2). Calculations were performed using MPM-B and the parameters given in Appendix 1 (with the exception of Cbl).

Interestingly, if we analyze the experimental ubiquitination curves normalized to the same maximum (**Fig. 37B**), modulation of Cbl levels does not affect the ubiquitination threshold. In parallel, the model also showed that alteration of Cbl levels (within a limit), despite affecting the absolute amount of EGFR ubiquitination (**Fig. 37C**), does not affect the ubiquitination threshold (**Fig. 37B**). This lack of effect on the EGFR-Ub threshold is

not immediately obvious, given the sizable changes in total EGFR ubiquitination upon Cbl modulation. Importantly, this condition was verified by the model only when the ratio between Cbl available for EGFR binding (estimated as Cbl-pY) and the EGFR surface levels was less than 0.2 (**Fig. 37D**). Notably, this number is in agreement with the number of Cbl-pY molecules/cell that we estimated experimentally (**Fig. 33D**).

7 - The Early Activation Model

In order to investigate how the EGFR ubiquitination threshold can be modulated, we used the *MPM-B* to identify the parameters that can alter the position of the threshold. To this aim, a normalized "sensitivity parameter" for the EGFR ubiquitination threshold x_T (the position of the Ub-threshold on the x-axis - or the EGF concentration where the Ub or pY curves reach their half-maximal value) was derived by dividing or multiplying, by one order of magnitude, all the parameters in the model. A parameter was considered as a good candidate for experimental verification when it was greater than 1. This analysis identified only one parameter, the EGFR phosphorylation rate k_{KIN}, the experimental validation of which would be challenging because it represents a sum of reactions that lead to EGFR intracytoplasmic tyrosine phosphorylation.

To identify alternative parameters for experimental validation, we developed the **Early Activation Model** (**EAM**), which includes the molecular details of EGFR activation. The *EAM* includes the features that were formalized in the *MPM-B*, as summarized in Section 6: I) cooperativity, II) Cbl limiting, III) probabilistic hypothesis. Moreover, in the *EAM,* we replaced the Hill function used to couple EGF concentration and EGFR phosphorylation (Result Section 5.1) with more detailed biological rules, such as EGFR opening and closing, EGF binding to EGFR, conformational changes in the extracellular domain and receptor dimerization. To address these additional reactions, we started from previously proposed models (Klein, Mattoon et al. 2004; Macdonald and Pike 2008). Obviously, the EAM carries

more parameters than *MPM-B*; for many of these, we defined a reasonable range of values, thanks to the vast amount of experimental and modeling data available in literature (Appendix 1). After implementation of the new model, we fitted the dose-response curves for EGFR phosphorylation and ubiquitination with the EAM. Importantly, we obtained results very similar to those obtained with the MPM-B: also the EAM quantitatively accounts for the ubiquitination threshold observed *in vivo* at 2 minutes of EGF stimulation (**Fig. 38**), reproducing the molecular circuitry underlying EGFR-Ub *in silico*.

7.1 - The EGFR-Ub threshold is sensitive to surface EGFR levels

The EAM was explored to highlight parameters whose alteration might affect the positioning of the threshold over the range of EGF concentrations used. The identification *in silico* of those parameters would be advantageous in designing experiments aimed at understanding how cells control their EGF-mediated responses and how this control might be subverted in pathological conditions.

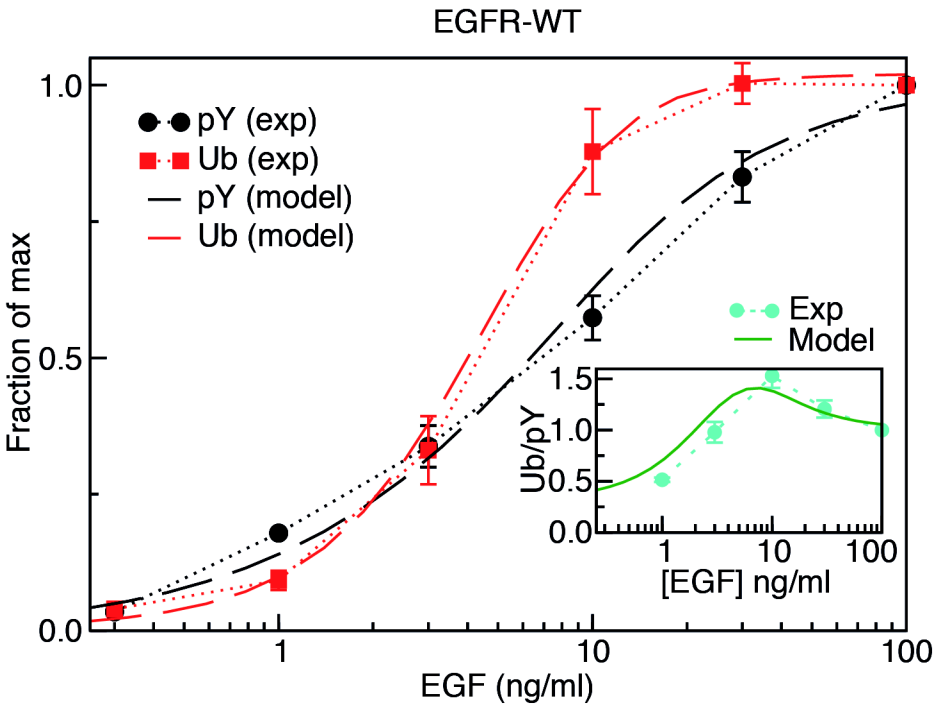


Figure 38: The EAM reproduces the EGFR phosphorylation and ubiquitination curves. Comparison of experimental data (exp; dashed lines) and modeled (solid lines) phosphorylation and ubiquitination dose-response curves for EGFR-WT. Inset shows the ratio of ubiquitination to phosphorylation as a function of EGF concentration for experimental and modeled data. Experimental data as in **Fig.37.**

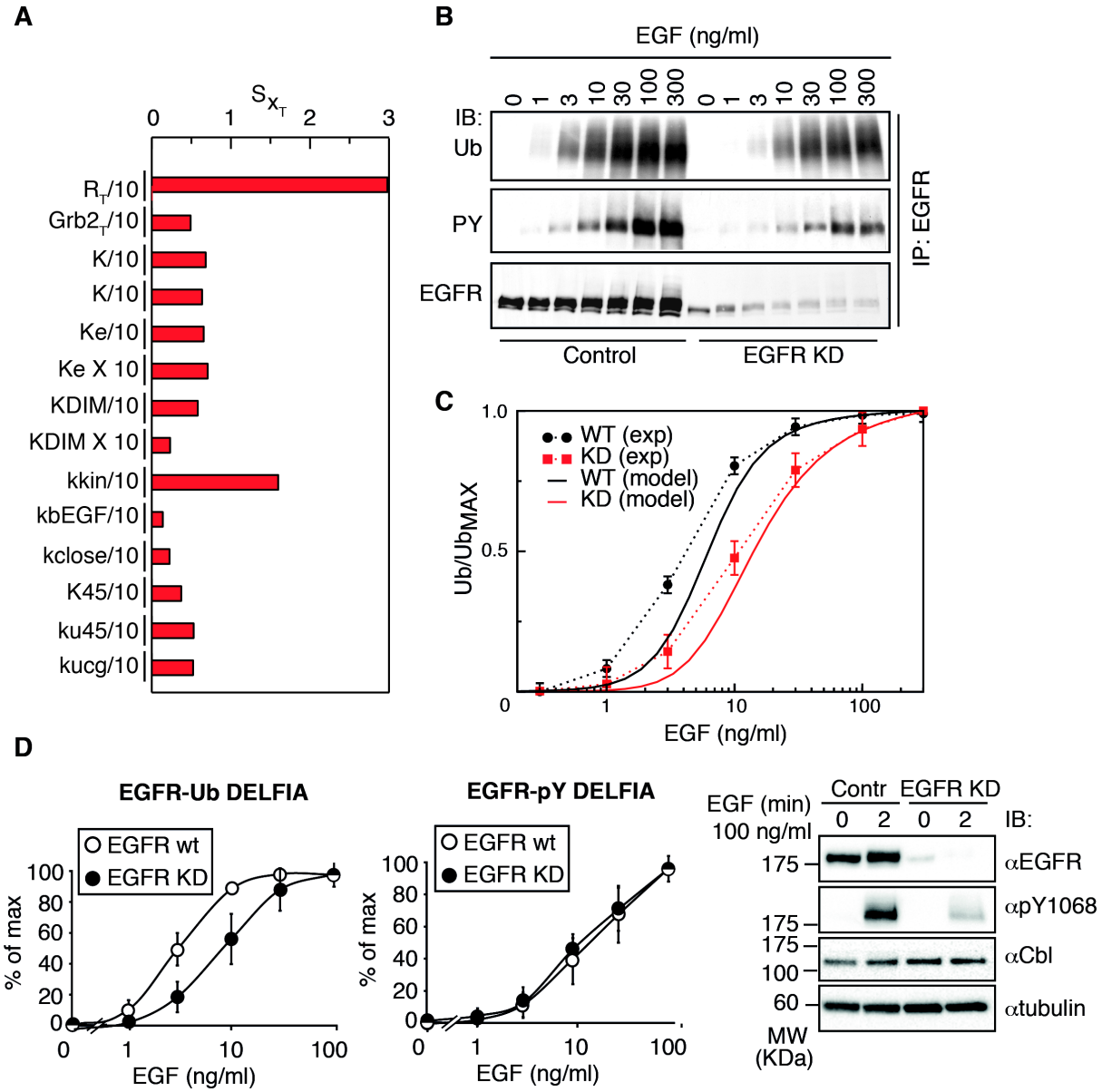


Figure 39: Downmodulation of EGFR levels shifts the ubiquitination dose-response curve. A - EAM parameters were varied 10-fold and the sensitivity coefficient S for the EGFR ubiquitination threshold position x_T was computed. Only parameters whose variation resulted in a sensitivity coefficient of at least 0.1 are reported. **B -** HeLa cells were subjected to EGFR knockdown (KD) with an anti-EGFR siRNA oligo, whereas control cells were transfected with mismatched oligos. Cells were treated for 2 minutes with EGF as indicated and IP and IB were performed as shown. Quantitation of EGFR ubiquitination by densitometry is shown in C. Blots are representative of at least 3 independent experiments. **C -** Model predictions and experimental data (based on blots in **B**) of the ubiquitination threshold in EGFR-KD (KD) or control (WT) HeLa cells. In the simulation of the EGFR-KD, the decrease in EGFR levels was assumed to be 4.2-fold, as determined in ^{125}I-EGF saturation binding assays performed on the KD cells (not shown). EGFR ubiquitination is expressed, for each condition, as normalized to the maximum value obtained in that condition (Ub/Ub_{max}). Results are expressed as the mean ± SD calculated from at least 3 experiments **D -** Left and middle panels, EGFR-Ub and -pY dose-response curves upon downmodulation of EGFR measured by DELFIA. HeLa cells were treated an in B. Again, results are expressed as the mean ± SD calculated from at least 3 experiments. Right panel, WB analysis was performed using the indicated antibodies to verify efficiency of EGFR KD. Note that Cbl levels do not change upon EGFR KD. Tubulin was used as a loading control. This blot control is representative of 3 independent experiments.

The position of the threshold was revealed to be robust, with only two parameters identified that could affect the positioning of the threshold (**Fig. 39A**). One candidate parameter whose modification affected significantly the positioning of the Ub threshold was again the EGFR phosphorylation rate, k_{KIN}, in agreement with MPM-B. Additionally, the total number of EGFRs, R_T, emerged as a good candidate. Compared to k_{KIN}, R_T has the significant advantage of being easy to manipulate. Thus, we validated the predictive power of the model by attenuating the expression of EGFR in HeLa cells via incomplete KD (**Fig. 39B**), achieving an ~4-fold decrease in EGFR levels (from ~3×10^5 to ~7×10^4 EGFRs/cell, measured by ^{125}I-EGF saturation binding assay). Stimulation with increasing EGF doses for 2 minutes resulted in a shift of the ubiquitination curve to a position that was remarkably similar to that shown in simulations in which the EGFR levels were reduced 4-fold (**Fig. 39C**), shifting the curve towards higher EGF concentrations. HeLa EGFR-KD cells were also assessed by DELFIA assay, achieving again a shift of the EGFR-Ub threshold towards higher EGF concentrations (**Fig. 39D**, left panel), while no effect was observed on the EGFR-pY (**Fig. 39D**, center panel).

In conclusion, guided by the EAM, we were able to establish that the total number of surface EGFRs is a key parameter in the control of the position of the ubiquitination threshold. We also confirmed this prediction experimentally, thereby endorsing the predictive power of the model. This prediction might have important implications on how the Ub threshold - and the switch between CME and NCE - controls biological output.

8 - Challenging the model: changing EGFR and EGF levels

Having highlighted the importance of the number of EGFRs per cell, we wanted to investigate in more detail the behavior of EGFR-Ub/-pY over a range of different receptor levels - spanning from physiological levels (<10^5 EGFRs/cell) to pathological levels detected in human tumors (>10^6 EGFRs/cell) - and over different EGF concentrations. This issue is

highly relevant to cancer since overexpression of EGFR in human tumors is frequently accompanied with the overproduction of its ligands, which can act on the receptor in an autocrine or paracrine fashion (Arteaga 2002; Zandi, Larsen et al. 2007).

The EAM was interrogated about the behavior of EGFR-Ub and -pY upon variation of EGFR levels per cell and at different EGF concentrations, generating an advanced prediction of EGFR-Ub/-pY patterns (**Fig. 40**). At all EGF doses, the average EGFR ubiquitination per receptor (red line, normalized to the maximum ubiquitination obtained at 100 ng/ml EGF) displayed a bell-shaped curve, indicating that ubiquitination increases as a function of EGFR surface expression, until a maximum is reached, after which it decreases. Interestingly, the peak is progressively shifted towards lower EGFR surface levels, as the EGF concentration increases. In contrast, the average phosphorylation per receptor (black line, normalized to the maximum phosphorylation obtained at 100 ng/ml EGF) behaved differently, with the phosphorylation peak progressively shifting towards higher EGFR surface levels, as a function of EGF concentration. This behavior has important consequences: in the physiological range of EGFR levels, the system responds to increasing EGFR activation (i.e. pY) with a congruent increase in EGFR-Ub levels, at all EGF doses. However, when EGFR levels exceeded the physiological range, and for EGF concentrations above 1 ng/ml - two conditions frequently associated with cancer – EGFR-Ub decreases much more rapidly than its pY and the Ub/pY curves became *uncoupled*.

We experimentally tested this prediction by measuring EGFR ubiquitination and phosphorylation dose-response curves in cell lines expressing increasing amounts of EGFR. We followed two parallel approaches to verify this prediction: first, in order to eliminate biological variability, we took advantage of an isogenic background, i.e. NIH cells, in which we reintroduced the EGFR at different levels (Result Section 8.1); in parallel, we validated the prediction also in normal *vs*. cancer cells with increasing surface receptor levels (Result Section 8.2).

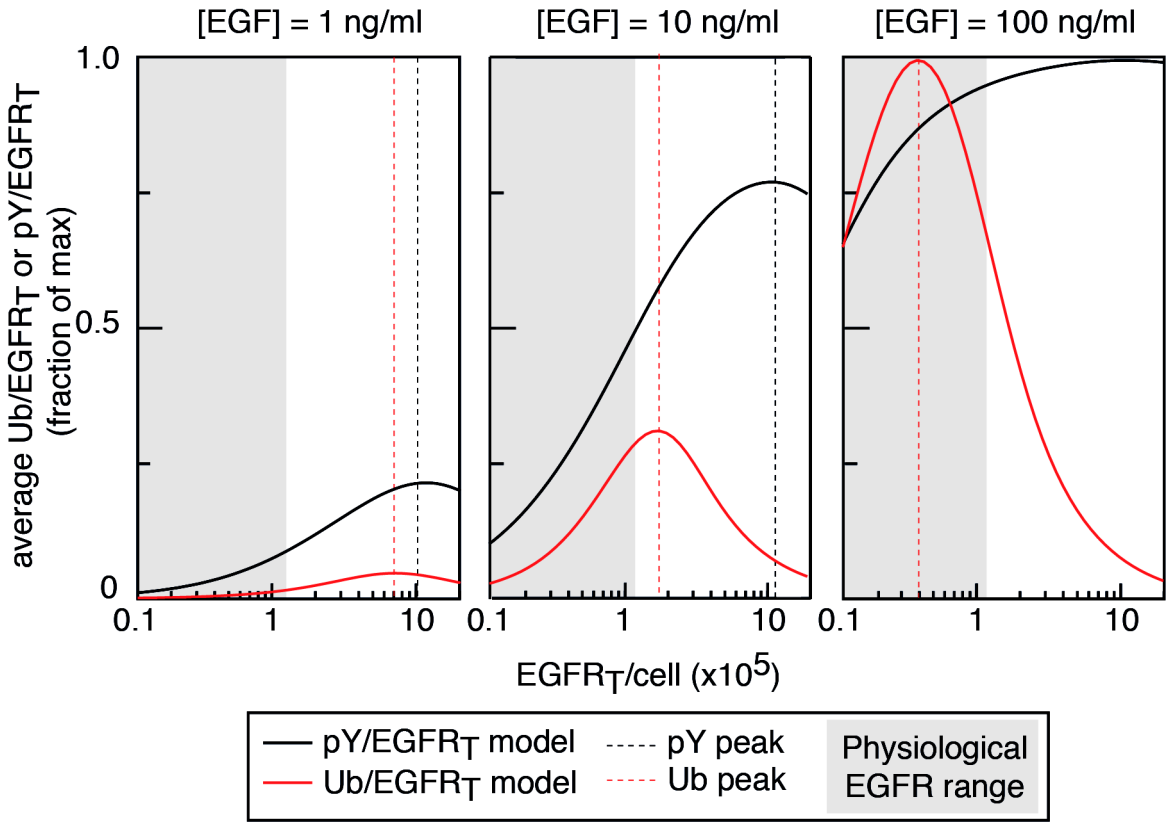


Figure 40 - Advanced prediction for EGFR-Ub and -pY as a function of EGF concentration and surface EGFR number. Relative phosphorylation ($pY_T/EGFR_T$, black lines) and ubiquitination ($Ub_T/EGFR_T$, red lines) levels of the EGFR, as given by the EAM, for the indicated EGF concentrations. The gray area represents the physiological range of EGFR levels. Dashed lines indicate the maximum levels of phosphorylation and ubiquitination. Curves are plotted as normalized to the individual maximum of either phosphorylation or ubiquitination obtained at 100 ng/ml of EGF.

8.1 - Isogenic background approach to validate the EGFR-Ub/pY uncoupling prediction

To perform experiments in a homogeneous genetic background, we used NIH-3T3 fibroblasts, which express low levels of endogenous EGFR (~10^4 EGFRs/cell) transfected with a retroviral expression vector (pBABE) encoding human EGFR-WT. We then selected three NIH-EGFR clones expressing different surface levels of EGFR (as assessed by saturation ^{125}I-EGF binding assay, **Fig. 41A**) representative of:

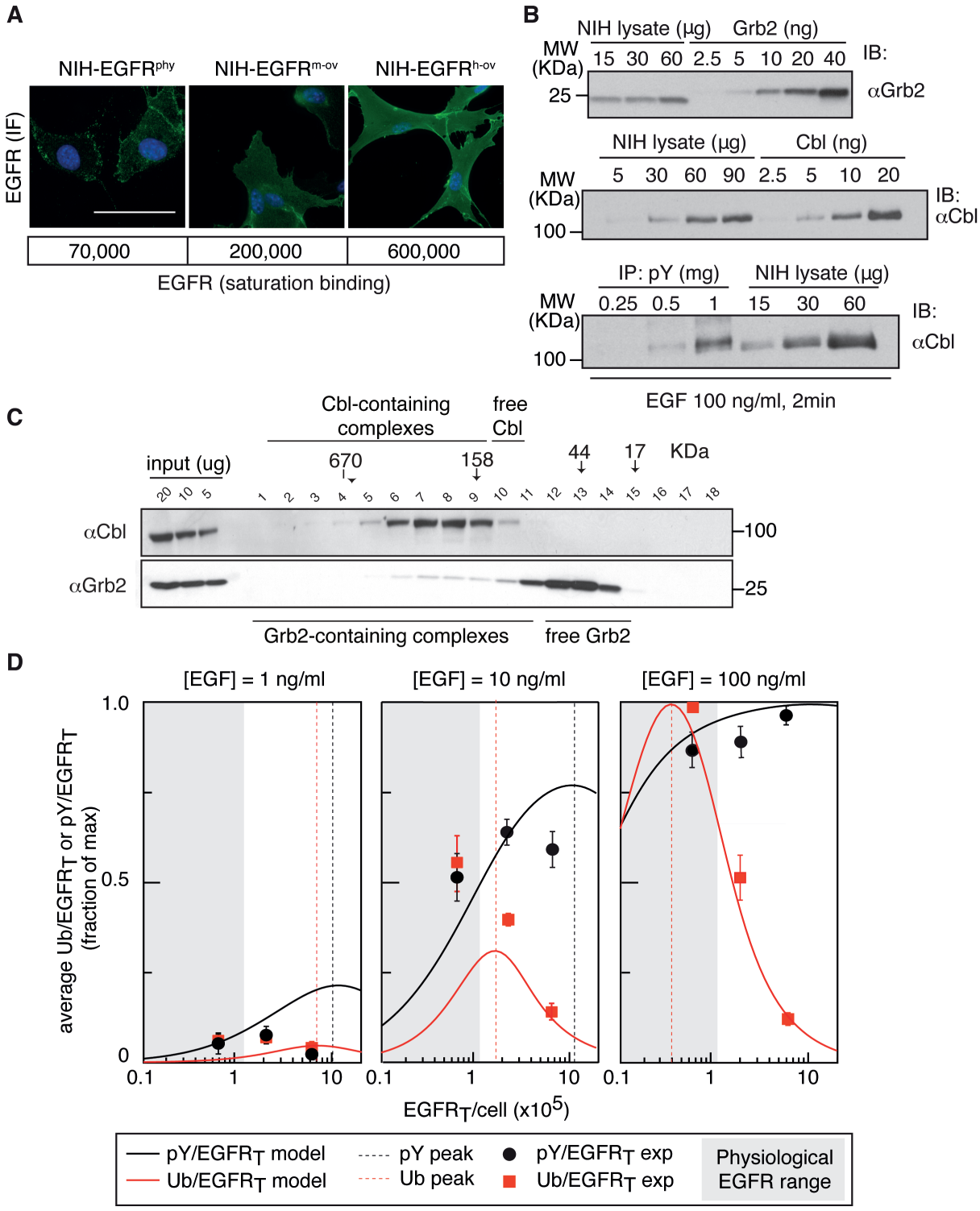
- the physiological condition (~7 x 10^4 EGFRs/cell, $EGFR^{phy}$ cells);
- moderate overexpression (~2 x 10^5 EGFRs/cell, $EGFR^{m\text{-}ov}$ cells);
- high overexpression (~6 x 10^5 EGFRs/cell, $EGFR^{h\text{-}ov}$ cells).

The three clones displayed homogenous expression of the receptor at the single cell level (**Fig. 41A**).

To better characterize this cell line in terms of absolute amounts of the key players in the system - Cbl and Grb2 - we calculated the number of total Cbl, Cbl-pY and Grb2 in NIH-EGFR$^{h\text{-}ov}$ cells in condition of high EGFR levels, in which the ratio between Cbl and receptors/cell could be critical. The analysis showed that NIH cells express similar levels of the key players of the Ub reaction as HeLa cells: 800,000 Grb2 molecules/cell and 130,000 Cbl molecules/cell. Considering only the active form of Cbl (Cbl-pY), its level decreased to 6,000 molecules/cell, thus Cbl remained limiting in the Ub reaction under these conditions compared to the number of EGFRs/cell (**Fig. 41B**, **Table VI**). We also verified, by size exclusion chromatography, whether Cbl, in NIH cells, exists mostly as a free monomer, or as part of macromolecular complexes. Again, we observed that the majority of Cbl fractionates at high MW, indicating that it is mostly engaged in macromolecular complexes, and that it is not commonly found as a monomer in cells (**Fig. 41C**).

We then proceeded to test the model predictions (**Fig. 40**). To this aim, the selected NIH clones were stimulated for 2 minutes al 37°C with 1, 10, 100 ng/ml of EGF, and their Ub and pY levels were measured by DELFIA assay. This quantitative assay allowed us to measure EGFR ubiquitination and phosphorylation in a relatively high number of samples in the same experiment, in contrast to WB analysis. Importantly, the NIH clones were able to recapitulate the Ub/pY bell-shaped curves reported by the model prediction, as reported in **Fig. 41D,** where the experimental data were superimposed onto the model simulation. In particular, the experimental data confirms the progressive uncoupling of the phosphorylation and ubiquitination dose-response curves with increasing EGFR surface levels.

Figure 41 (next page): Experimental validation of the EAM prediction reveals how EGFR-Ub and -pY are uncoupled at high EGFR number. A - NIH-3T3 cells were stably transfected with a vector carrying WT EGFR. NIH clones were selected expressing different numbers of EGFRs/cell (from lowest to highest: NIH-EGFR^{phy}, physiological EGFR; NIH-EGFR$^{m\text{-}ov}$, medium overexpression; NIH-EGFR$^{h\text{-}ov}$, high overexpression). Representative images of immunofluorescence (IF) analysis of EGFR surface levels (see Materials and Methods) in the indicated NIH-EGFR clones is shown. *Continues in the next page*...



Nuclei were stained with DAPI (blue). Bar 18 µm. Bottom, number of surface EGFRs/cell, as measured by 125I-EGF saturation binding assay (see Materials and Methods). **B** - Quantitation of the number of Grb2 and Cbl molecules in NIH3T3 cells overexpressing EGFR. Increasing amounts of NIH-EGFR$^{h\text{-}ov}$ (NIH3T3 clone with the strongest EGFR overexpression) cell lysate were subjected to IP/IB as indicated and compared with increasing amounts of in vitro purified Grb2 (top panel) or Cbl (middle panel) proteins. To estimate the ratio between the total and the phosphporylated form of Cbl, we IP increasing amount of NIH-EGFR$^{h\text{-}ov}$ cells (as indicated) and compared them to increasing amount of total lysate. Additional information about the estimation of Cbl and Grb2 molecules can be found in **Table VI** and Materials and Methods. **C** - NIH-EGFR$^{h\text{-}ov}$ cell lysate was subjected to size exclusion chromatography (see Materials and Methods). Sequential fractions eluting from the column were collected and analysed by WB to identify fractions containing Cbl and Grb2. **D** - Advanced model for EGFR-Ub and -pY as in **Fig.40**. Red squares and black circles represent experimental measurements of EGFR-Ub and -pY behavior, respectively, obtained by DELFIA assay on NIH-EGFR clones with increasing numbers of EGFRs/cell and stimulated with the indicated EGF doses. Experimental data are reported as mean ± SD from at least three independent experiments.

	N. of molecules / cell (x 10^3)	
	HeLa WT	**NIH-EGFR$^{h\text{-}ov}$**
EGFR	250 (± 50)	200 (± 50)
Grb2 tot	1000 (± 260)	800 (± 120)
Cbl tot	150 (± 25)	130 (± 8)
Cbl-pY	5.2 (± 0.75)	6 (± 1.5)

Table VI - Amount of critical players involved in the EGFR ubiquitination reaction in HeLa and NIH-EGFR$^{h\text{-}ov}$ cells. The number of surface EGFR molecules in HeLa cells and NIH3T3 cells overexpressing EGFR at high levels (NIH-EGFR$^{h\text{-}ov}$) was measured by ^{125}I-EGF saturation binding (see Materials and Methods). Data are expressed as number of surface EGFRs/cell, results are expressed as the mean ± SD calculated from at least three independent experiments. To calculate the total number of Grb2 and Cbl molecules (Grb2 tot and Cbl tot, respectively), we compared signal intensities of anti-Grb2 or anti-Cbl bands in IB of increasing amounts of cellular lysate with known amounts of purified Grb2 or Cbl protein (see **Fig. 35B** and **41B**). To calculate the number of active Cbl molecules (Cbl-pY), we IP Tyr phosphorylated Cbl, using an anti-pY antibody, from increasing amounts of cellular lysate derived from cells stimulated with EGF (100 ng/ml) for 2 min (see **Fig. 35B** and **41B**). Immunoprecipitated proteins were then IB with anti-Cbl. The amount of Cbl in the IPs was then compared with the amount in total cellular lysates and with known amounts of purified Cbl protein (see Materials and Methods). The IBs in **Fig. 35B** and **41B** from which these calculations derived, are representative of at least three independent experiments. Results are expressed as mean ± SD calculated from at least three independent experiments.

8.2 - Non-isogenic background approach to validate the EGFR-Ub/pY uncoupling prediction

We extended our analysis and tested the non-obvious prediction generated by the EAM in normal *vs*. cancer cells. We performed a screening of different cell lines, taking into consideration the EGFR levels/cell along with the expression of other proteins directly involved in the EGFR ubiquitination process (i.e. Cbl and Grb2; **Fig. 42A**). We selected for further analysis a number of tumor cell lines displaying increasing amounts of EGFR (HeLa, CASKI, BT20; in this latter line EGFR is amplified), compared to a normal fibroblast cell line (WI38; **Table VII**), while having similar levels of Cbl and Grb2 (**Fig. 42A, B)**.

Cells were stimulated for 2 minutes al 37°C with 100 ng/ml of EGF, and their Ub and pY levels were measured by the DELFIA assay. We also included in the experiment the three NIH clones with increasing EGFR levels as an internal reference. Data obtained by DELFIA analysis were plotted normalized to the NIH-EGFR^{phy} (**Fig. 42C**). Importantly, we could confirm the dramatic reduction in EGFR ubiquitination at high EGF concentrations, going from normal to cancer cell lines with increasing numbers of EGFRs/cell, while EGFR phosphorylation remained constant.

Overall, uncoupling of EGFR-Ub and -pY curves was observed by two independent approaches - using an isogenic background or a set of different cell lines - validating the EAM prediction and leaving us with the open question as to why this uncoupling is observed.

It is important to highlight that in our model, designed for HeLa cells, the edge where Cbl becomes limiting is when the ratio between Cbl-pY and surface EGFRs is less than 0.2 (**Fig. 37D**). We verified that this was the case both in CASKI and BT20. To this aim, we performed a quantitative WB analysis to calculate the number of total Cbl and Cbl-pY per cell (**Fig. 42D**, see Methods). Results are reported in **Table VII** and show that the levels of total Cbl are similar in all cell lines; more importantly, the fact that Cbl-pY is limiting is verified both in CASKI and BT20 cells (Cbl-pY/EGFRs <0.2).

Cell line	Origin	EGFRs/cell (x 10^3)	Cbl tot/cell (x 10^3)	Cbl-pY/cell (x 10^3)	Cbl-pY/ EGFRs
WI38	*Normal lung fibroblasts*	80	n.a.	n.a.	n.a.
HeLa	*Cervix adenocarcinoma*	250	150	5.2	0.02
CASKI	*Cervix epidermoid carcinoma*	400	190	2.7	0.0067
BT20	*Breast carcinoma*	1200	210	15	0.0125

Table VII – Origin of the cell lines used in the non-isogenic approach for the EAM validation. The indicated cell lines were characterized for the number of surface EGFRs/cell by saturation ^{125}I-EGF binding assay (see Materials and Methods). For the high EGFR expressing cell lines (CASKI and BT20), where Cbl levels may be critical for EGFR-Ub, we calculated the number of total Cbl (tot Cbl) and active Cbl (Cbl-pY) molecules/cell by comparing the signal intensities of the anti-Cbl bands in IBs of total lysates and anti-pY IPs, with the signal of the purified Cbl protein as in **Fig. 35B** and **41B**, as previously described (see also Table VI and Materials and Methods). Last column indicates the ratio of Cbl-pY molecules to EGFR molecules/cell. Note that, in our model, tailored for HeLa cells, Cbl becomes limiting when Cbl/EGFRs/cell reaches ≈0.2. Calculation of Cbl and pY-Cbl molecules are from blot in **Fig. 42D**.

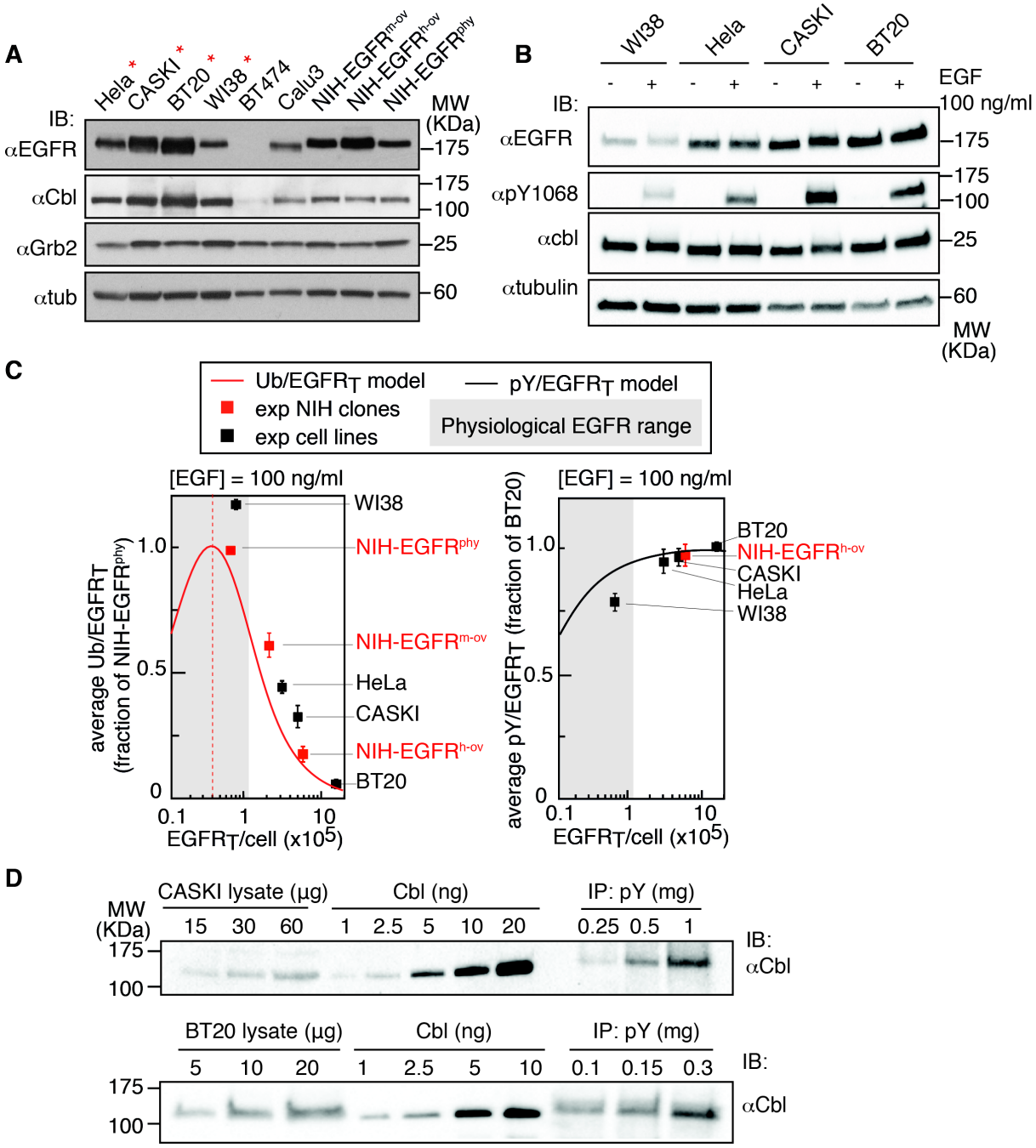


Figure 42: Uncoupling of EGFR-Ub/-pY at high EGFRs/cell was confirmed by a non-isogenic strategy. A - A panel of cell lines was investigated by IB to assess the levels of the key players in the EGFR ubiquitination machinery. Lysates from the indicated cell lines were IB as shown. Asterisks indicate cell lines with levels of Cbl/Grb2 comparable to HeLa cells, but with different EGFR levels (see **Table VII**), which were selected for the Ub/pY analysis performed in C. **B -** EGFR and Cbl levels were evaluated by IB in the indicated cell lines at steady state or upon stimulation with 100 ng/ml EGF for 2 minutes. Tubulin was used as loading control to evaluate that Cbl levels are unchanged in the different samples. **C -** Relative EGFR ubiquitination (left, $Ub/EGFR_T$, red line) or phosphorylation (right, $pY/EGFR_T$, black line), as given by the EAM, for 100 ng/ml EGF, normalized to their respective maximum. The gray area represents the physiological range of EGFR levels. Red dashed line indicates the maximum level of ubiquitination. Squares represent experimental measurements of EGFR ubiquitination (left) or phosphorylation (right), obtained by the DELFIA assay in the indicated NIH-EGFR clones (red) or cell lines (black). Note that cell lines used display increasing amounts of EGFR at the cell surface, going from physiological receptor levels to strong EGFR overexpression. Experimental data are reported as mean ± SD from at least three independent experiments. **D -** Increasing amounts of CASKI (upper) or BT20 (middle) cell lysate were subjected to IP/IB as indicated, and compared with increasing amounts of *in vitro* purified Cbl protein to obtain an estimate of the number of Cbl molecules per cell (see **Table VII** and Materials and Methods)

9 - Recoupling of EGFR phosphorylation and ubiquitination

The presence of bell-shaped curves in the model prediction can be explained by the saturation of the reactions leading to both phosphorylation and ubiquitination when the number of EGFRs increases and EGF concentration remains constant. When EGFRs increase above physiological levels, low doses of EGF are diluted among the receptors and the average phosphorylation per receptor decreases, possibly due to impairment of dimer formation at the PM due to competition with the many unbound inactive EGFRs. Similarly, as phosphorylated EGFRs decrease, also the fraction of ubiquitinated EGFRs decreases. However, at high EGF concentrations and especially at non-physiological levels of EGFR, the ligand is no longer diluted and the average phosphorylation per receptor does not decrease. Ubiquitination, however, decreases also in this case because it requires Cbl, which we have shown to be rate-limiting compared to EGFR and Grb2 and which becomes saturated by the many activated EGFRs. Therefore, despite the increase in potentially "ubiquitinatable" EGFRs, the fraction of actual ubiquitinated receptor decreases at high EGF under conditions of EGFR overexpression, while EGFR-pY is maintained.

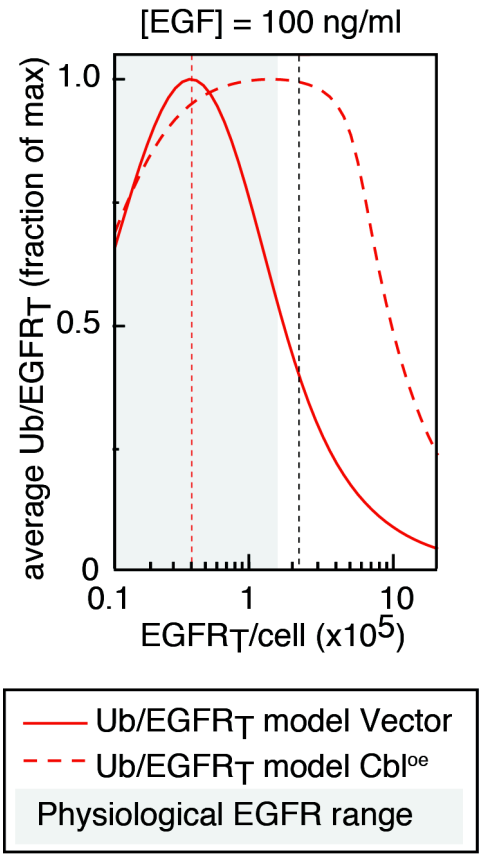


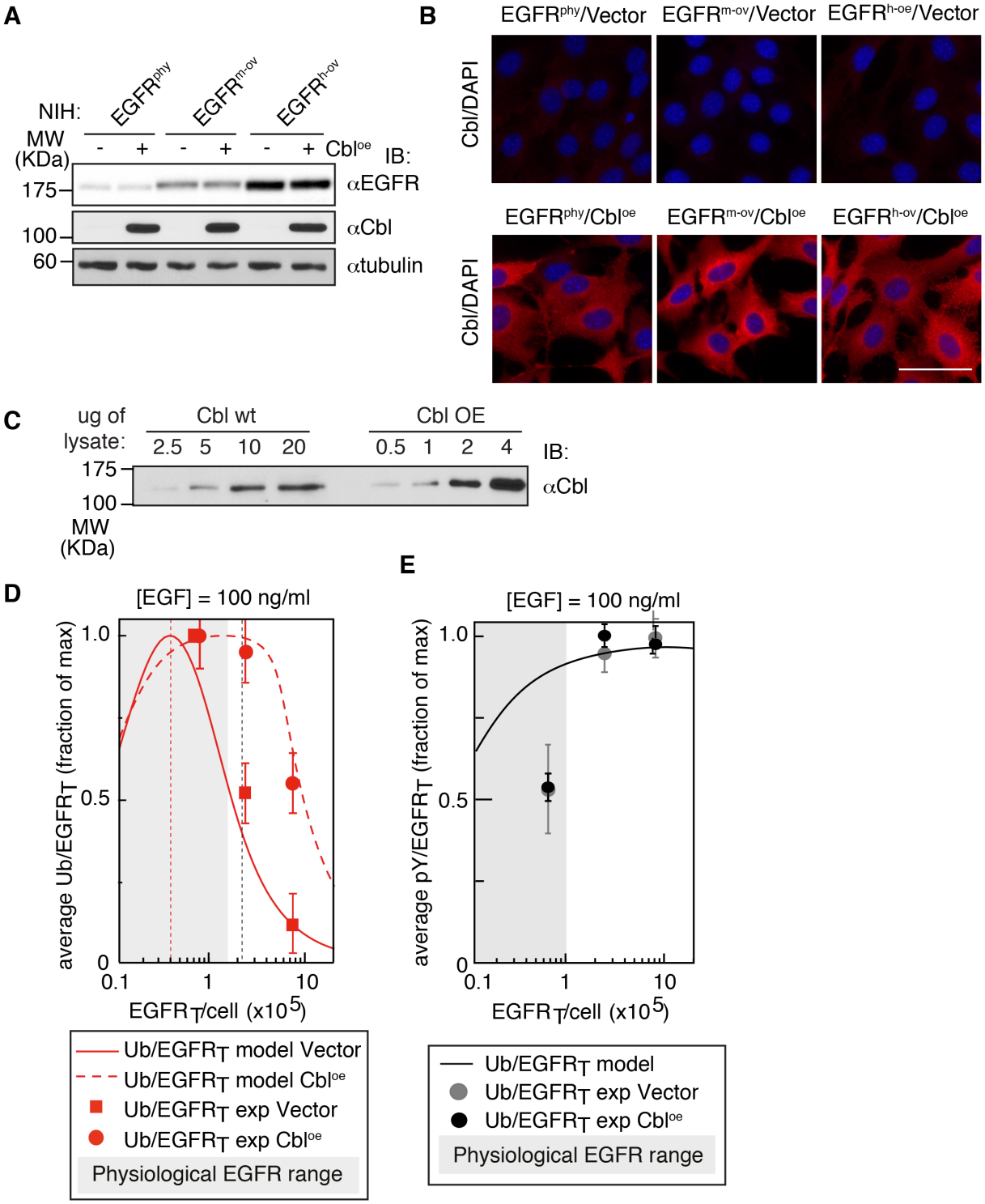
Figure 43 - Modeling EGFR-Ub under conditions of Cbl overexpression. Relative EGFR ubiquitination levels ($Ub/EGFR_T$), as given by the EAM, at 100 ng/ml of EGF under control (red line) or Cbl overexpressing (Cbl^{oe}; red dashed line) conditions. The gray area represents the physiological range of EGFR levels. Cbl overexpression was modeled as a 100X increase. Data were normalized to the maximum ubiquitination in each condition.

In support of this interpretation, we attempted to re-establish ubiquitination under conditions of

EGFR overexpression *in silico*, by increasing the levels of Cbl of 100 times. Cbl overexpression resulted in a partial restoration of EGFR ubiquitination (**Fig. 43**).

To experimentally test the prediction of the uncoupling between EGF-Ub and -pY behaviors, we transfected the NIH clones with a lentiviral inducible vector (pSLIK) that drives the overexpression of Cbl upon doxycycline treatment. We checked clones for Cbl expression levels by: i) WB to assess that overexpression was comparable between different NIH-EGFR clones (**Fig. 44A**) and ii) single cell IF analysis to control the homogeneous expression of the proteins among cells (**Fig. 44B**). An overnight treatment with 0.5 μg/ml of doxycycline was sufficient to induce an ~80-100 times increase in Cbl expression compared to endogenous levels, as calculated by quantifying different amounts of cell lysate by densitometric analysis (**Fig. 44C**). NIH clones -/+ Cbl were stimulated for 2 minutes with 100 ng/ml of EGF, and their Ub and pY levels were measured by the DELFIA assay. EGF treatment of NIH-Cbl overexpressing clones resulted in the complete restoration of EGFR ubiquitination in the $EGFR^{m\text{-}ov}$ cells and in a 50% increase in ubiquitination in the $EGFR^{h\text{-}ov}$ cells, in perfect agreement with model prediction (**Fig. 44D**). We conclude that the levels of Cbl, in conditions of EGFR overexpression, are critical to guarantee the correct coupling of EGFR phosphorylation and ubiquitination.

Figure 44 (next page): EGFR-Ub and -pY upon Cbl overexpression. A - NIH clones ($EGFR^{phy}$, $EGFR^{m\text{-}ov}$ and $EGFR^{h\text{-}ov}$) infected with empty vector or with a lentiviral vector driving inducible Cbl overexpression, were treated with doxocycline (0.5 μg/ml for 16 h), as indicated above each lane, to induce expression of the transgene. Lysates were analysed by IB as shown. **B -** Immunofluorescence staining of Cbl in the NIH clones infected with empty vector (Vector) or upon Cbl overexpression (Cbl^{oe}); comparable exposures are shown. Note that endogenous Cbl is not detected at this exposure, while it gives a punctuate pattern at higher exposures (data not shown). Clones with Cbl^{oe} showed homogeneous Cbl expression level in >95% of the cells (Cbl, red; nuclei, DAPI blue). **C -** Increasing amounts of total cellular lysate from NIH-$EGFR^{h\text{-}ov}$ cells expressing endogenous Cbl (Cbl wt) or overexpressing Cbl (Cbl^{oe}), were examined by IB using an anti-Cbl Ab, followed by densitometric quantitation. Cbl overexpression was calculated to reach ~100 times the expression of endogenous Cbl by densitometric analysis. This result is a mean of at least three independent experiments. **D -** Relative EGFR ubiquitination in WT conditions or upon Cbl overexpression as presented in **Fig.43**. Squares represent experimental measurements (obtained by DELFIA assay) of EGFR-Ub NIH clones (phy, m-ov and h-ov) infected with empty vector (Vector, red) or with inducible Cbl overexpression vector (Cbl^{oe}, black). Results are expressed as the mean ± SD calculated from at least three independent experiments. **E.** Relative EGFR phosphorylation in WT conditions or upon Cbl overexpression as given by the EAM. EGFR-pY is not predicted to be altered upon Cbl overexpression (gray line represents the WT condition in model curves, Cbl OE is not reported because it is coincident with EGFR-pY in the WT background). Gray circles represent experimental measurements (obtained by DELFIA assay) of EGFR-pY NIH clones empty vector, whereas black circles represent experimental data upon Cbl overexpression.


A
NIH:
EGFRphy
EGFRm-ov
EGFRh-ov
MW (KDa)
- + - + - + Cbloe
IB:
175
αEGFR
100
αCbl
60
αtubulin
B
EGFRphy/Vector
EGFRm-ov/Vector
EGFRh-oe/Vector
Cbl/DAPI
EGFRphy/Cbloe
EGFRm-ov/Cbloe
EGFRh-ov/Cbloe
Cbl/DAPI
C
ug of lysate:
Cbl wt
2.5 5 10 20
Cbl OE
0.5 1 2 4
IB:
175
100
MW (KDa)
αCbl
D
[EGF] = 100 ng/ml
average Ub/EGFRT (fraction of max)
1.0
0.5
0
0.1
1
10
EGFRT/cell (x10^5)
Ub/EGFRT model Vector
Ub/EGFRT model Cbloe
Ub/EGFRT exp Vector
Ub/EGFRT exp Cbloe
Physiological EGFR range
E
[EGF] = 100 ng/ml
average pY/EGFRT (fraction of max)
1.0
0.5
0
0.1
1
10
EGFRT/cell (x10^5)
Ub/EGFRT model
Ub/EGFRT exp Vector
Ub/EGFRT exp Cbloe
Physiological EGFR range

Discussion

In this study, we started from our previous observation that when HeLa cells are stimulated with increasing EGF doses, ubiquitination of the EGFR follows a threshold behavior, being minimal at low doses, maximal at high doses and presenting a rapid increase in an EGF interval centered between 1 and 10 ng/ml of EGF. Importantly, in this same interval of EGF concentrations, NCE is activated, and experimental evidence indicated that ubiquitination is essential for the internalization of the EGFR through this pathway (see Introduction Section 3.2). NCE targets the majority of EGFRs to degradation, thus downregulating EGFR signaling under conditions of high ligand concentration. Given the biological importance of the activation of this internalization route, in this project, we aimed to understand the molecular mechanisms underlying threshold generation and NCE activation. To this end, in collaboration with the Andrea Ciliberto's group at IFOM, we expressed as mathematical formalism all the reactions leading to EGFR activation and ubiquitination, attempting to describe at the molecular level the input layer of our system. Each step in the development of the model was supported either by the literature or by direct experimental validation. Thanks to this side-by-side use of modeling and experiments, we have uncovered a novel mode of EGFR-dependent signal transduction that does not follow a simple linear input->output relationship. This finding provides a mechanistic explanation of how the ubiquitination threshold is generated as a function of ligand concentration. The model, designed as such, was able to predict non-obvious behaviors of normal *vs.* cancer cells that have been experimentally validated.

1 - Qualitative to quantitative analysis of EGFR ubiquitination and phosphorylation

To build a mathematical model of EGFR activation and ubiquitination, we needed to obtain reliable quantitative data of EGFR phosphorylation and ubiquitination following stimulation with ligand. Quantitative data better deal with a mathematical formalism.

Initially, the EGFR ubiquitination and phosphorylation patterns were characterized by WB analysis. Qualitative experimental data were translated into quantitative data by image processing and densitometric analysis of the WB images. However, this procedure did not guarantee perfect reproducibility in data acquisition. Moreover, exposure times and signal saturation introduced additional levels of variability in the data. For instance, depending on the exposure time, some signals might reach saturation while others are still in the linear phase; this results in over-estimation of the latter signals. By analyzing dose-response curves from different sets of WB data, we found that the highest variability was, as expected, in the transition phase (i.e., inside the 1-10 ng/ml EGF interval; for a significant example see **Fig.25**. During this phase, the system is instable and highly sensitive to small experimental alterations (e.g. small variations in EGF concentration, incubation time, and signal saturation). For these reasons, although we obtained WB data that was in good agreement with the threshold model, we decided to develop a quantitative assay to measure EGFR phosphorylation and ubiquitination in order to reduce experimental variability.

As described in the Results Section 1, we developed a reliable quantitative EGFR ubiquitination/phosphorylation assay. This assay is based on the ELISA method and was developed using the DELFIA technique (Perkin Elmer). We now routinely use this assay in our laboratory to measure ubiquitination and phosphorylation levels of EGFR upon different perturbations. DELFIA was found to be a specific assay, as determined by the absence of signal in cells devoid of endogenous EGFR (colo602 cells, see **Fig. 22B**), and very

reproducible (from both technical and biological replicas, see error bars in **Fig. 23** and **24**). Indeed, the standard deviations (SD) derived from DELFIA outputs were much lower with respect to WB outputs for both ubiquitination (**Fig. 25A**) and phosphorylation (**Fig. 25B**) of the receptor. Importantly, we confirmed the existence of a threshold for receptor ubiquitination, which is generated between 1 and 10 ng/ml of EGF. The two Ub curves obtained by WB and DELFIA analysis display a similar pattern. However, with DELFIA the Ub curve was slightly shifted to the left (to lower EGF concentrations) relative to the curve obtained by WB. This difference is due almost exclusively to differences in the 3 ng/ml EGF data point, which is higher in the DELFIA output compared to the WB analysis. A likely explanation for the different results is the higher sensitivity and higher dynamic range of the DELFIA signal. In addition, the SD of DELFIA at this critical point is very low (almost invisible in the chart), indicating higher reliability. Of note, the phosphorylation curves obtained with the two assays were almost identical, and we were also able to show that phosphorylation of pY1068 follows similar hyperbolic behavior in DELFIA as the one observed by WB (**Fig. 24B**).

Thus, we developed a sensitive and reliable quantitative assay based on DELFIA technology to measure EGFR ubiquitination and phosphorylation. Using this assay we generated reliable data for inclusion in our models.

2 - Role of phosphatases in counteracting EGFR kinase activity at early vs. late time points

Prior to the modeling of EGFR ubiquitination, we focused our efforts on reproducing the behavior of EGFR-pY observed *in vivo*, developing a simple model to account for EGFR phosphorylation at two minutes of EGF stimulation. We considered EGFR phosphorylation kinetics as a fast enzymatic reaction (Olsen, Blagoev et al. 2006; Kleiman, Maiwald et al.

2011), in which the catalytic domain of the one receptor moiety trans-autophosphorylates the cytoplasmic tail of the second receptor moiety in the EGFR homodimer. To take into account the potential contribution of phosphatases in the system, we investigated experimentally their impact on EGFR phosphorylation.

It had already been reported in literature that phosphatases have an important role in determining the overall EGFR phosphorylation levels (Monast, Furcht et al. 2012). In a panel of human tumor lines, including H1666, HeLa and HepG2 cells along with non-transformed MCF-10A cells, Kleiman and colleagues showed that active EGFR presents a phosphorylation peak 10 minutes after EGF stimulation, which progressively decreases to background levels within 2 hours (Kleiman, Maiwald et al. 2011). Total EGFR levels also decreased with time after EGF treatment, indicative of EGF-dependent receptor degradation. To study the role of phosphatases in the same panel of cell lines, the authors blocked EGFR kinase activity using the clinical-grade kinase inhibitor Gefitinib (see Results Section 4), and observed a rapid decrease in total EGFR phosphorylation, along with a decrease in phosphosite-specific receptor phosphorylation. The half-life of the phosphorylated receptor was estimated to be approximately 10-15 seconds, depending on the cell line used, indicating a very fast turnover for EGFR phosphorylation. Thus, the authors concluded that total EGFR phosphorylation levels is the result of a balance between phosphorylating and dephosphorylating events, and that receptors cycle between EGFR and EGFR-pY forms.

Considering these findings, we investigated the action of phosphatases in our cell model system. The aim of our experiments was to assess the activity of phosphatases at 2 minutes after EGF stimulation, since the published work focused mostly on the 10 minutes time point (Kleiman, Maiwald et al. 2011). Using the inhibitor Gefitinib, we observed that phosphatases are very active already at 2 minutes after EGF stimulation (Results Section 4); abrogation of the EGFR kinase activity post-EGF treatment, caused a rapid and complete impairment of EGFR phosphorylation. We calculated the dephosphorylation constant k_{PTP} of the receptor at

2 minutes after EGF stimulation, which is in agreement with previously published data (Kleiman, Maiwald et al. 2011 in Supplementary Material). To fully understand whether phosphatase activity has a critical role in counteracting kinase activity at 2 minutes after EGF stimulation, we used the chemical compound sodium orthovanadate to inhibit phosphatase activity *in viv*o, while leaving the EGFR kinase activity unaltered. We found that phosphatase inhibition caused a 10-20% increase in EGFR phosphorylation levels at 2 min, while at later time points (i.e. 10 minutes) it had a bigger effect. Thus, kinase activity is much stronger than phosphatase activity at 2 minutes, compatible with efficient and fast EGFR phosphorylation at the PM. This point is important for the generation of the EGFR ubiquitination threshold. Indeed, EGFR ubiquitination follows a threshold behavior only when the ratio between k_{KIN}/k_{PTP} does not exceed a certain value, otherwise the system would be unbalanced towards phosphatase activity. Such an imbalance would translate into an insufficient level of EGFR phosphorylation that, especially in presence of high ligand concentration, would result in insufficient ubiquitination of the receptor.

It is also interesting to speculate what may happens at later time-points, when phosphatases appear to be more active. A recent study has demonstrated that mean EGFR-pY content is constant among endosomes, a compartment that is reached by active EGFR once it has been internalized (Villasenor, Nonaka et al. 2015). The authors explained this observation by suggesting that, although increasing EGF concentration results in increased levels of total phosphorylated EGFR, there is a balanced partition of EGFR-pY molecules between cellular endosomes, regardless the EGF dose. Increasing the number of endosomes, rather than varying their content, is one way to achieve this balanced partitioning. Indeed, when endosome dynamics and number were altered by knocking down components of the endosome tethering and fusion machinery, an alteration in the intracellular signaling events due to an unbalanced partitioning of EGFR-pY among endosomes was observed. It has been proposed that such a finely-tuned compartmentalization of EGFR-pY in endosomes is

achieved via the activity of phosphatases in the endosomes. A simple mechanism through which this could happen is that the EGFR dephosphorylation rate increases as the EGFR-pY content per endosome increases. In such a scenario, phosphatase activity might be activated by the kinase itself in an autoregulatory loop.

Our data are in agreement with such a scenario where phosphatase activity increases after the initial steps of EGFR activation, with the specific aim of regulating downstream events such as signaling. Although phosphatases are accounted for in our early model, they do not substantially contribute to the generation of the EGFR-Ub threshold. However, their more prominent role at the endosomal level must be taken into account when generating a time-resolved model of EGFR activation (see Ongoing work and future directions), which will include EGFR endocytosis and signaling along with data formalized in previous models (Kleiman, Maiwald et al. 2011; Villasenor, Nonaka et al. 2015).

3 - Mechanisms underlying the ubiquitination threshold

The ubiquitination threshold is essentially determined through a mechanism of "coincidence detection" of two Tyr residues in the EGFR by the Grb2:Cbl complex. In this mechanism, two components are essential: 1) *cooperativity* and 2) *probability*.

1) ***Cooperativity*** necessitates that a combination of multiple phosphorylated tyrosines is required for Cbl recruitment. Cbl is the E3 ligase responsible for EGFR ubiquitination and it operates in collaboration with the adaptor protein Grb2 (see Results Section 6). Indeed, Cbl can bind directly to the receptor through the pY1045 or indirectly through Grb2 that interacts with the pY1068/86 on the receptor. Only the cooperative binding of the Cbl:Grb2 complex through both phosphorylated sites allows for full EGFR ubiquitination (see Results Section 6.1). This scenario has been verified through different lines of evidence:

I. the interaction of Cbl to the EGFR-pY is dependent upon the presence of its binding site, pY1045, and on Grb2 binding sites, pY1068/86. Indeed, Cbl binding to the receptor is

impaired in mutants that carry only pY1045 or pY1068/86 (**Fig. 33**): their abrogation (singularly or in combination) strongly reduces the Cbl/EGFR interaction (**Fig. 33B**);

II. we also observed that Cbl is recruited to the activated EGFR in a threshold-controlled fashion, when all the three key pYs are present (**Fig. 34A**, middle and lower panels), whereas if only the Y1045 site is present, the Cbl:EGFR complex forms with a linear pattern over increasing EGF concentrations (**Fig. 34A**, upper panel);

III. the interaction between Cbl and the receptor is reflected in the pattern of EGFR ubiquitination (**Fig. 33C**). The abrogation of pY1045 and pY1068/86 results in total impairment of EGFR-Ub, while the presence of pY1045 alone determines a strong reduction in EGFR ubiquitination, which gradually increases with increasing EGF concentrations. In contrast, the presence of all the three binding sites is able to fully rescue the EGFR-Ub threshold.

2) The ***probability*** component represents a mechanism for generating a non-linear output (threshold for EGFR-Ub) from a linear input (EGF concentration). As already mentioned, phosphorylation of two key pYs is necessary and sufficient for full EGFR ubiquitination and - since Tyr residues on EGFR are phosphorylated independently of each other - the concomitant phosphorylation of two Tyr residues occurs via a purely probabilistic process. This implies that while the probability of individual Tyr residues undergoing phosphorylation increases gradually with the concentration of ligand, the probability of having two, three, or more phosphorylated Tyr residues does not. Indeed, multiple phosphorylation events will be insignificant at low EGF concentrations and will increase abruptly after reaching a critical concentration of EGF (**Fig. 36**).

In further support of the probabilistic hypothesis, we already published direct experimental proof that the concomitant presence of pY1045 and pY1068 on the same receptor moiety increases sharply and non-gradually as a function of ligand (Sigismund, Algisi et al. 2013). Indeed, the probability of having two given sites (pY1045 and either one

of Y1068 or Y1086) phosphorylated on the same EGFR molecule increases non-linearly as a function of ligand concentration, and becomes significant only when a critical EGF concentration is reached. Therefore, we could envision a situation in which – at high doses of ligand – the simultaneous presence of the two binding sites (pY1045 and pY1068/pY1086) allows for efficient recruitment of the Cbl:Grb2 complex.

We translated this phenomenon into the model, formalizing the probabilistic hypothesis whereby the concomitant presence of the two key phosphorylated tyrosine residues (for the binding of Cbl:Grb2 complex) increases non-linearly as a function of the EGF concentration (see Results Section 6.3). This probabilistic mechanism guarantees that such a combination arises naturally only at high EGF concentrations.

Cooperativity and probability were both implemented in the early model, to obtain the *MPM-B* consisting of two modules: phosphorylation (*MPM*) and Cbl:Grb2 binding (*B*). *MPM-B* faithfully reproduced the EGFR-Ub and -pY dose response curves observed *in vivo* (**Fig. 37A**).

4 - Additional mechanisms that could contribute to the ubiquitination threshold

In developing the MPM-B, we used standard mathematical formulations to represent the chemical reactions that lead to EGFR activation. For example, phosphorylation was described as a standard multisite chain of reactions, Grb2 and Cbl binding was represented as simple mass action, and cooperativity/probability were also formalized. In order to decrease model complexity and simulation time, we were forced to introduce some simplifications, such as the fact that in the free regime of phosphorylation of Tyr residues, each residue can be phosphorylated independently of each other; this allowed us to restrict the analysis to the

three critical Tyr residues. For the EGFR ubiquitination reaction, the model traces back only Cbl binding to the EGFR.

For what concerns the choice of parameters, many of them were either well constrained or were experimentally estimated. EGFR activation was modeled by introducing a phenomenological function based on the Hill equation, which avoided the introduction of many parameters that describe the series of events in EGFR activation. Phosphorylation was characterized by fitting the dose-response curves. The binding of the Cbl:Grb2 complex to the EGFR relied on protein concentrations that were carefully measured and binding affinities that were indirectly determined by fitting ubiquitination dose-response curves under various experimental conditions, within well-defined experimental constrains taken from previous published studies (Appendix 1).

Given this series of simplifications, it is not unexpected that the minimalism of the model came at a price. Despite the overall agreement, there are areas in which we noticed some discrepancies between modeled and experimental data. First, the experimental EGFR ubiquitination curves always showed a higher degree of sigmoidicity than the *in silico* simulations, suggesting that the analogical-to-digital conversion operating in real-life is even more efficient than in the modeled predictions. Thus, the model requires further optimization, and this will be the focus of future investigations.

Second, the overexpression of Cbl does not reproduce model predictions in exact quantitative terms. In HeLa cells, we achieved an overexpression of Cbl to levels approximately 6-8 times those of endogenous Cbl, but in the model we used an overexpression of 2-fold. In this case, model predictions and experiments can be made to agree quantitatively only if we assume that only 1/3 of the overexpressed Cbl molecules are capable of binding to the activated EGFR. In the case of Cbl overexpression in NIH cells, we were able to reach a higher level of overexpression (around 100 times), but this resulted in a partial re-coupling of EGFR-Ub and pY.

A possible explanation of the discrepancy between the model and experimental data with regard to Cbl, might be that in our model we lack some parameters of the Cbl regulatory network. For example, Cbl phosphorylation upon EGF stimulation is the event that triggers Cbl activation as a functional E3 ligase (Dou, Buetow et al. 2012). In our model, we did not consider this feedback loop, whereby EGFR activation is responsible for Cbl activation. We can speculate that the lack of this level of information may subtract sharpness from the modeled ubiquitination curve. Moreover, Cbl can homodimerize and heterodimerize with other Cbl family members, such as Cbl-b (Peschard, Kozlov et al. 2007), and is also subjected to multiple levels of regulation within the cell. Those different levels of regulation include for instance, lysosomal and proteasomal degradation as well as interaction with other proteins that bind Cbl and subtract it from the pool able to ubiquitinate the EGFR (Ryan, Davies et al. 2006). This was partially confirmed in our study by size exclusion chromatography experiments, in which Cbl was observed mainly engaged in large complexes (**Fig. 35A**). Finally, the activity of DUBs might counteract the activity of the Cbl E3 ligase at the PM. Indeed, it is known that some DUBs are active at the level of the endosome (Clague, Liu et al. 2012), although no DUBs have thus far been shown to operate on EGFR at the PM. If PM-active DUBs were to be identified, their activity could be implemented in our model of early EGFR activation.

Cbl phosphorylation, dimerization and counteraction by DUB activity, although not essential for the generation of the EGFR-Ub threshold, might help to refine the mathematical abstraction of the system in a future more comprehensive model.

5 - The ubiquitination threshold is robust but sensitive to variations in the number of EGFRs

The threshold effect predicted by the MPM-B was robust to changes in parameter values, a property that was corroborated by experiments in living cells. In particular, Cbl overexpression and the expression of a Cbl dominant negative mutant (Cbl70Z, modeled as down-regulation, **Fig. 35D**) did not displace the threshold, both in the modeled response and in experimental settings. Of note, this condition is verified only if Cbl is overexpressed within a certain level, in particular when the ratio between the amount of active Cbl and the number of EGFR molecules does not exceed 0.2.

To identify those parameters whose alteration might change the position of the EGFR-Ub threshold, we had to expand the MPM-B to create the EAM. The EAM includes the molecular details of EGFR activation (i.e. EGF binding, EGFR conformational changes, dimerization, and trans-phosphorylation), which were hidden in the MPM-B behind the phenomenological relationship between EGF and EGFR activation, given by the Hill equation (see Results Section 5.1). The addition of new biological laws that connect EGF concentration to EGFR activation, led to a higher degree of complexity in the model. Strikingly, EAM proficiently reproduced the patterns of EGFR-Ub and -pY (**Fig. 38**). The EAM also displayed a significant robustness to changes in the parameters and the increased complexity of the model permitted the generation of an interesting prediction: the position of the EGFR-Ub threshold can be shifted to higher EGF concentrations by decreasing the number of surface EGFRs. We also experimentally verified this prediction in HeLa cells by EGFR KD, achieving a shift of the Ub threshold towards higher EGF concentrations (**Fig. 39B-C**). Thus, the surface number of EGFRs is a key parameter in the control of the position of the ubiquitination threshold. We hypothesized that this happens because the reduction of receptors at the PM decreases the capability of EGFRs to dimerize and to be activated.

Results from a study by Mellman and colleagues, in which they investigated the time for dimer formation upon EGF stimulation support our hypothesis (Chung, Akita et al. 2010). Unliganded EGFR continuously fluctuate between monomer and dimer states. Pre-formed dimers can be made through the transient interaction of two monomers, but only the binding with the ligand is able to stabilize the dimer form and activate the trans-phosphorylation of the two intracellular tails. This means that in the absence of the EGF, the k_{off} of the dimer dominates over the k_{kin} of the kinase domain. Moreover, the EGF-induced EGFR dimerization time, which contributes to the generation of the k_{on} parameter, is dependent upon the local concentration of EGFR pre-formed dimers, at least in EGFR overexpressing cells (Chung, Akita et al. 2010). At the single cell level, there are differences in the dimer time formation between the periphery, where EGFR dimers are formed in a shorter time interval, and the center of the cell, where the time for dimer formation is longer. This observation explains the higher probability of finding a dimer at the periphery of the cell with respect to the center (Chung, Akita et al. 2010). However, in certain EGFR cell lines that do not overexpress the receptor or where EGFR expression is knocked down, the difference in dimer time formation between the periphery and center of the cell was lost and the overall time for dimer formation turned out to be significantly longer with respect to overexpressing cells (Chung, Akita et al. 2010). The hypothesis drawn from these data was that cells with increased surface EGFRs form dimers faster than those which have decreased EGFR surface levels, probably due to simple mass action dynamics. Moreover, even in presence of EGF, dimers are formed and disrupted continuously, and the rate of dimer formation is dependent upon EGFR surface number (Chung, Akita et al. 2010).

These findings are in agreement with our observations: we saw that by decreasing the EGFR number, the ubiquitination threshold is shifted towards higher EGF concentrations. This shift has important implications to how EGFR ubiquitination and its threshold pattern control biological output by governing the switch between CME and NCE. The EGFR-Ub

threshold shift is perhaps due to the fact that the probability of EGF-bound dimers to form is reduced under conditions of decreased EGFR, and only by increasing EGF concentration can we restore a scenario in which more stable dimers are formed, resulting in phosphorylation of the three key Tyr residues and thus allowing EGFR ubiquitination.

An implied effect of the EGFR-Ub threshold shift is that as EGFRs decrease on the surface because of downmodulation and degradation, and under conditions of high EGF concentration, the threshold for NCE internalization is progressively shifted towards higher EGF doses. This directs an increasing proportion of EGFRs to CME and thus towards recycling rather than degradation. In turn, this would lead to a final equilibrium in which a certain amount of EGFR is still preserved at the PM, thereby allowing the persistence of signaling required for generating a biological output. This mechanism therefore might allow for proper signal maintenance over a wide range of ligand concentrations, which is achieved by a balance between EGF-dependent signaling and degradation.

6 - An important non-obvious prediction: uncoupling of EGFR-Ub/pY

Given the importance of surface EGFR at the PM in determining the extent of receptor ubiquitination, we challenged the EAM to predict the behavior of both EGFR-Ub and -pY upon the variation of surface receptor levels, but also EGF doses. The prediction that was generated by the model displays a complex behavior (**Fig. 40**, see also Results Section 8) and appears to have evolved in such a way that the maximal response is in the physiological range of EGFR levels. At all EGF concentrations, the average EGFR ubiquitination per receptor showed a bell-shaped curve, indicating that ubiquitination increases as a function of EGFR surface expression until a peak is reached, after which it decreases. Indeed, the EGFR-Ub peak is progressively shifted towards lower EGFRs, as EGF concentration increases. Thus, at low EGF concentrations (1 ng/ml), the peak values for EGFR ubiquitination and phosphorylation per receptor were essentially coincident (**Fig. 40,** left panel), while at higher

EGF concentrations (10-100 ng/ml), ubiquitination saturated at lower EGFR levels compared with phosphorylation (**Fig. 40,** middle and right panel). Since EGFR ubiquitination is coupled to the internalization of the receptor through the NCE pathway (Sigismund, Argenzio et al. 2008; Sigismund, Algisi et al. 2013), this means that over the wide physiological range of EGFR levels, which spans one order of magnitude, the cell is well equipped to react to increased signaling with increased receptor degradation, to protect itself from overstimulation. Instead, at supra-physiological EGFR levels and for EGF concentrations above 1 ng/ml - two conditions frequently associated with cancer - the ubiquitination of the receptor decreases much more rapidly than receptor phosphorylation and the two curves are *uncoupled*. This would translate into a faster attenuation of NCE-mediated degradation with respect to signaling.

This prediction was experimentally confirmed in cell lines through two parallel approaches: the first approach used an isogenic background model system, while the second employed a panel of cell lines. In the first approach, we managed to express human EGFR at different levels in the murine NIH-3T3 cell line, in order to achieve a homogeneous tool in which to study EGFR-Ub and –pY (see Results Section 8.1). By exploiting the NIH clones, we validated the non-obvious prediction generated by the EAM, and confirmed the uncoupling of the ubiquitination and phosphorylation curves (**Fig. 41D**). We also confirmed the prediction in the panel of normal and tumor cell lines that express increasing amounts of surface EGFR, mimicking the progression towards a more aggressive phenotype (Rimawi, Shetty et al. 2010). Thus, our results might harbor relevant implications for tumors displaying overexpression of the EGFR.

7 - Modeling EGFR ubiquitination in cancer

Our findings are highly relevant to human cancers since the EGFR is frequently overexpressed and/or amplified in tumors, either alone or in combination with its ligands (for a recent review see (Arteaga and Engelman 2014). EGFR has been shown to be overexpressed in a wide variety of cancers and, for some cancers, its overexpression is suggested to be a risk factor for poor prognosis and is associated with a more aggressive clinical progression, such as in lung, breast, ovarian, bladder, esophageal, cervical, and head and neck cancers (Spano, Lagorce et al. 2005 and references therein).

In many cancer types, overexpression of EGFR is accompanied by aberrant autocrine production of EGFR ligands, including EGF, TGFα, AREG, BTC, EREG, EPGN and HB-EGF (P, Rhys-Evans et al. 2000; Schneider and Wolf 2009, see also Introduction Section 2.2). Interestingly, also risk factors such as tobacco smoking, which is one of the main factors contributing to the development of many types of cancer, increases the expression of some EGFR ligands, for example TGFα and AREG in oral epithelial cells.

In NSCLC, it has been found that EGFR overexpression is often accompanied by overexpression of TGFα ligand, constituting an autocrine loop able to start signaling cascades. This distinctive trait is frequent in early stage NSCLC, but is not associated with differences in survival, suggesting a role for the EGFR/TGFα loop in tumor formation, rather than in tumor progression (Rusch, Klimstra et al. 1997). In addition, head and neck squamous cell carcinoma (HNSCC) frequently displays overexpression of EGFR and its ligand, especially EREG (Jedlinski, Ansell et al. 2013), and high EGFR expression levels have been associated with a more aggressive phenotype, poor prognosis, and resistance to cancer therapy (Ang, Berkey et al. 2002).

Multiple anti-EGFR antibodies and small molecule tyrosine kinase inhibitors targeting the EGFR have been developed; up to now, the main EGFR-targeted therapy in use comprises the

two tyrosine kinase inhibitors Erlotinib and Gefitinib for treatment of NSCLC and the monoclonal antibody Cetuximab, approved for the treatment of NSCLC (Yang, Liu et al. 2014), advanced HNSCC (Loeffler-Ragg, Schwentner et al. 2008) and colorectal cancer.

For what concerns treatment response based on ligand expression, data are divergent: in HNSCC, for example, EREG mRNA expression was found to be low in cell lines resistant to Cetuximab treatment, suggesting it could be a predictive marker of Cetuximab response. This finding can be explained by the hypothesis that cells overexpressing both EGFR and its ligands depend on signaling through this RTK for maintenance of viability and/or growth and will therefore respond well to EGFR-targeted therapy. In the same direction, in colorectal cancer, high ligand expression of AREG and EREG identified a subgroup of KRAS WT patients who had a high probability of responding to Cetuximab compared with KRAS WT patients with low ligand expression. Conversely, in some cases, ligand overexpression is associated with a lack of treatment response. For example, AREG overexpression correlates with Gefitinib resistance in NSCLC (Busser, Sancey et al. 2010) and overexpression of HB-EGF has been proposed as a possible mechanism of Cetuximab resistance in HNSCC (Hatakeyama, Cheng et al. 2010). The rationale behind these data might be that high levels of ligand may compete with the antagonist at the binding site and displace it.

No matter what the downstream effects are of EGFR and ligand overproduction, these data are in agreement with our proposed scenario. Indeed, high EGFR levels, along with ligand overexpression, result in a situation in which cells are not able to counteract receptor activation and stimulation of signaling cascades with Ub-mediated EGFR downmodulation. This determines a set of events that potentially facilitates the onset of cancer.

Various strategies involving small-molecule inhibitors have been developed to target the EGFR system, such as Gefitinib and Erlotinib: two drugs that specifically target the tyrosine kinase activity of EGFR (see Results Section 4). The observation that sensitivity to Gefitinib and Erlotinib correlated very strongly with a newly discovered class of dangerous somatic

activating mutations in the EGFR kinase domain, highlighted the importance of mutationally activated kinases as anticancer drug targets (Haber, Bell et al. 2005). Indeed, although WT EGFR is able to determine cancer transformation by simple overexpression, another mechanism by which a cell might transform into cancer cell is through the mutation of the EGFR kinase domain. Kinase domain mutations in EGFR are generally referred to as activating mutations, as they result in the increased kinase activity of the receptor, leading to hyperactivation of downstream pro-survival pathways (for a review see Zhang, Stiegler et al. 2010). EGFR kinase domain mutations target four exons (18 – 21), which encode part of the tyrosine kinase domain and are clustered around the ATP-binding pocket of the enzyme (Kobayashi, Ji et al. 2005; Pao, Miller et al. 2005). Recent studies revealed that these mutants are oncogenic in both cell culture and transgenic mouse studies. However, not all the mutated EGFRs are necessarily constitutively or fully active, since their degree of ligand independence might not be total.

This is the case of mutation L834R (or L858R if numbering includes the signal peptide) that is a partial activating mutant of EGFR, which can be rendered fully ligand-independent - and therefore constitutively active - by second site substitutions in EGFR, such as a specific mutation in exon 20 (Kobayashi, Ji et al. 2005). *In vitro* biochemical studies using purified recombinant cytoplasmic domains of WT and L834R EGFR have shown that this mutant has increased k_{cat} values and an increased K_m for ATP, along with decreased K_D for homodimer formation (Sharma, Bell et al. 2007). Additionally, EGF binds more avidly to mutant L834R EGFR expressing cells, and its rate of dissociation from the receptor is retarded (Shtiegman, Kochupurakkal et al. 2007). Interestingly, the L834R EGFR mutant shows impaired association with Cbl, and consequently compromised ubiquitination (Shtiegman, Kochupurakkal et al. 2007). The most direct mechanism for abrogating c-Cbl-mediated ubiquitination of L858R would be abrogation of Y1045 phosphorylation in the EGFR, the only direct docking site for Cbl (Waterman, Levkowitz et al. 1999). However, data present in

literature indicate that Y1045 in the L858R EGFR mutant is still found strongly phosphorylated (Shtiegman, Kochupurakkal et al. 2007). The reason why this mutant cannot efficiently recruit Cbl is still unclear. One hypothesis that has been put forward is that L834R EGFR displays an enhanced propensity to heterodimerize with ErbB2, which would lead to the recruitment of several signaling molecules, but not Cbl (Shtiegman, Kochupurakkal et al. 2007). The reason for the decreased affinity of the mutant for Cbl binding might reside in a different conformation of the heterodimers compared to EGFR homodimers, possibly masking the Cbl binding site on the EGFR. Further studies are required to clarify this point.

According to these observations, we challenged our model to predict the behavior of the L834R EGFR mutant, which was shown to display increased phosphorylation together with decreased Cbl binding and ubiquitination. Indeed, Cbl binding to the L834R EGFR mutant corresponded to ~20% of that to WT EGFR at 20 ng/ml EGF, as experimentally shown (Shtiegman, Kochupurakkal et al. 2007). Based on parameters present in literature (Shtiegman, Kochupurakkal et al. 2007), we introduced the mutant and WT kinetic values into our EAM model and produced the simulation for EGFR ubiquitination and phosphorylation. According to our model, EGFR is hyperphosphorylated already at low EGF doses, whereas EGFR-Ub remains overall at very low values (**Fig. 45**). In this setting, Ub and pY curves are uncoupled already at physiological levels of EGFR (**Fig. 45**), providing a possible explanation of why this mutant is tumorigenic in the absence of receptor overexpression and/or ligand overproduction.

8 - Cbl as the limiting factor

In the model, both the robustness of the EGFR ubiquitination threshold and the uncoupling of activation (that is represented by EGFR-pY) from attenuation (i.e. EGFR-Ub) is due to the fact that Cbl is limiting. In all the cell lines used in this work, the total amount of Cbl is relatively abundant; however, experiments confirmed that the number of Cbl molecules

available for EGFR binding is limited (see **Tables VI** and **VII**), in agreement with the fact that Cbl is engaged in many different macromolecular complexes in the cell (**Fig. 35A** and **Fig. 41C**).

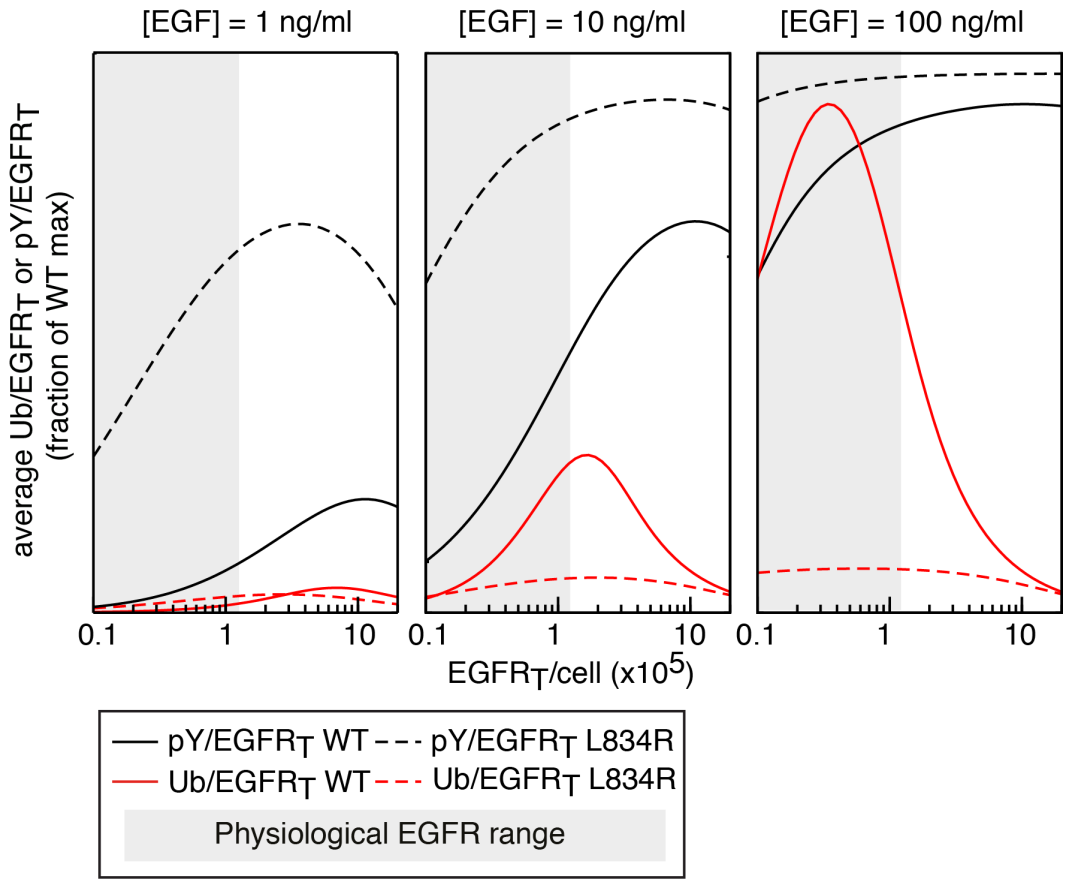


Figure 45: Predicted behavior of EGFR-Ub and -pY for the L834R mutant, as a function of EGF concentration and EGFR number. Relative L834R EGFR phosphorylation ($pY/EGFR_T$, black lines) and ubiquitination ($Ub/EGFR_T$, red lines) levels, as given by the EAM, are shown for the indicated EGF concentrations. The gray area represents the physiological range of EGFR levels. Data are normalized to the maximum phosphorylation/ubiquitination of EGFR WT obtained at 100 ng/ml EGF. Dashed lines represent mutant curves; continuous lines represent WT curves (the same as in **Fig. 27D**).

Indeed, Cbl is responsible for ubiquitination of a variety of proteins. It is known that more than 150 proteins are substrates of Cbl ubiquitination, including receptors with or without intrinsic kinase activity, cytosolic kinases, phosphatases, ubiquitin ligases, adaptors and many more proteins, representing a cross section through the signal transduction proteome (reviewed in Schmidt and Dikic 2005). This makes it difficult to separate Cbl E3 ligase activity and adaptor function, even when focusing only on EGFR endocytosis. In our lab, it was observed that Cbl is necessary for EGFR endocytosis through both clathrin-dependent

and -independent pathways. For the latter case, we widely discussed the need of EGFR ubiquitination for NCE internalization and receptor degradation. For what concerns the CME pathway the situation is still unclear, but recent evidence points to an essential role of Cbl in the ubiquitination of protein adaptors involved in this pathway (Pascolutti et al, unpublished). This finding highlights another important aspect of Cbl availability in EGFR ubiquitination; at high doses of ligand, when NCE is turned on, CME is still active and those Cbl moieties that act on this internalization route might be subtracted or compete with the ones directly involved in EGFR ubiquitination and NCE. This would further reduce the pool of available active Cbl moieties able to ubiquitinate the receptor.

Importantly, one critical element that contributes to the robustness of the ubiquitination threshold is the rate-limiting availability of Cbl: increase in Cbl leads to the same proportional increase in EGFR ubiquitination for all EGF concentrations. This holds true until the ratio between active Cbl/EGFRs is less than ≈0.2. Based on this finding, we can estimate that in HeLa cells, the number of active Cbl molecules per cell could reach the level of 50,000 without affecting the Ub threshold. Paradoxically, this feature also determines the intrinsic instability of the balance between EGFR phosphorylation and ubiquitination, which could be unmasked at supra-physiological levels of EGFR. NIH cells were again useful to test this complex behavior; through Cbl overexpression to levels ~100 times that of endogenous Cbl, we managed to partially re-couple EGFR-Ub and –pY as predicted by the model (see also Discussion Chapter 4). Thus, the model has been critical in revealing yet another fundamental property of the system: i.e., the uncoupling of EGFR ubiquitination and phosphorylation depends on the rate limiting availability of Cbl.

9 - Could this mechanism work for other RTKs?

The EGFR regulation circuit that we highlighted in this work is peculiar and finely tuned and it determines an important biological outcome, especially with respect to cancer-related

pathologies. One interesting question that raised is whether other receptors are similarly regulated. The MET receptor might be one candidate, since Cbl can also be recruited to this receptor following ligand-dependent activation (Peschard, Fournier et al. 2001).

The MET receptor is another member of the RTK family, important in the control of tissue homeostasis under normal physiological conditions. Similarly to the EGFR, MET receptor has also been found to be aberrantly activated in human cancers through mutations, amplifications or overexpression. For instance, aberrant expression of the MET receptor and its ligand, the HGF, has been observed in tumor biopsies of many solid tumors and MET signaling has been documented in a wide range of human malignancies (Comoglio and Trusolino 2002; Birchmeier, Birchmeier et al. 2003; Peruzzi and Bottaro 2006).

Structurally, the MET receptor is produced by the proteolytic processing of a common precursor into a a/b heterodimer (**Fig. 46**) (Trusolino and Comoglio 2002). The extracellular portion of the MET receptor is responsible for ligand binding and dimerization, whereas the TM and juxtamembrane regions link the extracellular side of the MET receptor to the intracellular TK domain. This domain contains the catalytic tyrosines that positively modulate receptor activity by activating downstream responses common to many RTKs, such as the MAPK and AKT pathways (Abella, Peschard et al. 2005). Following ligand binding, MET undergoes internalization through clathrin mediated endocytosis and degradation in a

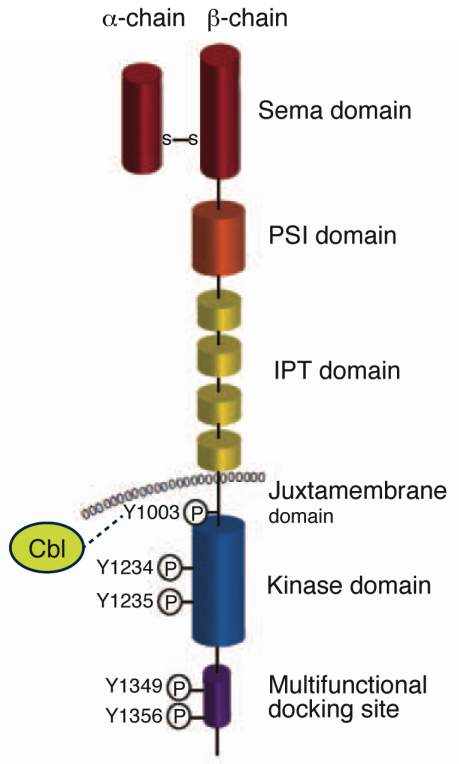


Figure 46: Domain structure of MET receptor. The receptor is formed by proteolytic processing of a common precursor into a single-pass, disulphide-linked a/b heterodimer. The MET receptor contains an intracellular tyrosine kinase catalytic domain flanked by distinctive juxtamembrane and C-term sequences. This portion of c-MET contains the catalytic tyrosines Y1234 and Y1235, which positively modulate enzyme activity, while the juxtamembrane tyrosine 1003 negatively regulates MET by recruiting Cbl. The multifunctional docking site in the C-term tail contains tyrosines Y1349 and Y1356, which recruit several transducers and adaptors. (Adapted from Organ and Tsao 2011)

dynamin-dependent poorly-carachterized mode (Hammond, Urbe et al. 2001).

It is interesting to note that the juxtamembrane domain of MET carries the Tyr 1003, whose phosphorylation negatively regulates the receptor by recruiting the ubiquitin ligase Cbl, that binds to MET receptor either through its direct binding site (pY1003), or indirectly through Grb2 (Peschard, Kozlov et al. 2007), destining the receptor to lysosomal degradation (Hammond, Urbe et al. 2001; Abella, Peschard et al. 2005). Indeed, it was observed a dependency of MET degradation on its ubiquitination, suggesting that the degradative pathway might require receptor ubiquitination, whereas it is not required for its internalization (Abella, Peschard et al. 2005). Moreover, MET receptor lacking the Cbl binding site was found to be tumorigenic, showing enhanced transforming activity in Rat1 fibroblasts with respect to the WT counterpart, likely due to reduced MET degradation and sustained phosphorylation (Mak, Peschard et al. 2007).

Even though MET internalization, in particular clathrin independent pathways, show a much lower resolution with respect to EGFR internalization pathways, a recent work by McMahon lab was published, proposing a role of the FEME pathway (see Introduction Section 4.3) for MET internalization. They showed that at high doses of HGF, stimulation of the MET receptor induced in the cell lines under scrutiny the formation of morphological structures that are linked with the newly endocytic clathrin-independent dynamin-dependent pathway therein described (Boucrot, Ferreira et al. 2015). We can thus speculate the existence of multiple entry pathways for the MET receptor, that might depend on ligand concentration and receptor ubiquitination.

MET might be one candidate to extend the study of receptors ubiquitination. Since stimulation with increasing doses of ligand seems to determine also in this case the engagement of different cellular fates, it would be interesting to elucidate this mechanism.

10 - Concluding Remarks

The EAM represents the input layer and the starting point for the construction of a 'time-resolved' model that will be later generated, taking into account events downstream of EGFR activation that take place in the signaling processing unit, such as integration of CME and NCE, sorting and intracellular signaling. The 'time-resolved' model will allow us to predict the most likely mechanisms of selection of a particular endocytic route, the evolution of the population of inactive, phosphorylated and ubiquitinated receptors, as well as the intensity of the signals transmitted to the downstream pathway regulated by the EGFR. The predictions generated by the time-resolved model will have - if validated - important implications for the future development of novel cancer diagnostic and therapeutic tools based on the Cbl/EGFR balance in cells.

Ongoing work and future directions

1 - Construction of a quantitative and time-resolved advanced model

We expect that the EAM is able to explain the EGFR-Ub and -pY early responses, but not the downstream biological readouts, such as EGFR endocytosis and biological responses, such as signaling and proliferation. To gain a deeper understanding of the biological events following EGFR activation, we need to extend our predictions to later time points.

Receptors residing at the PM represent the first-line sensors of extracellular signals. Not surprisingly, therefore, the regulation of their surface levels achieved by internalization has an immediate impact on the response of a cell to stimulus from the external environment. There are several mechanisms through which endocytosis controls receptor signaling specifically at the cell surface, such as the *regulation of receptor availability* at the PM, by which cells control the number of the surface receptors. NCE ligand-mediated endocytosis determines a rapid internalization of the EGF:EGFR complex, thus providing a mechanism that may directly limit the magnitude and duration of signaling from the PM. We thus hypothesize that continuous stimulation with high EGF doses, which targets ~50% of engaged receptors to NCE for degradation, is the cause of a long-term reduction in the number of cell surface receptors. This negative feedback loop, activated only at high EGF, could be essential to prevent excessive signaling and to maintain homeostasis. Loss of this safety feedback loop could contribute to pathological conditions, such as cancer, where overexpression of both ligands and receptors frequently leads to malignant transformation.

In line with this, the *integration of different endocytic pathways,* CME and NCE, is predicted to be critical to control the net biological output. Indeed, CME (active at all EGF doses) is the recycling/signaling-competent route, while NCE (active only at high EGF dose concomitant to CME) is a degradative pathway (Sigismund, Woelk et al. 2005). In

pathological conditions, such as cancer, one could predict that alterations in this balance (e.g. CME up-regulation or NCE down-regulation) might contribute to aberrant EGFR signaling.

Another mechanism by which endocytosis controls receptor signaling is through the assembly of *endosome-specific platforms* that result in a differential distribution of signaling effectors between the PM and the endosomal compartment. It is widely recognized that signaling is not restricted to the PM; as internalization proceeds, activated transmembrane molecules, with their tails exposed toward the cytoplasm, are confined and enriched within endomembrane organelles, from where they continue to signal. Such structures are *bona fide* signaling platforms that influence the duration, amplitude, and specificity of the downstream signals (Gould and Lippincott-Schwartz 2009). Consistent with this view, a growing number of signal transduction pathways are reported to require active endocytic machinery, or to originate exclusively from different endosomes. The term *signaling endosome* symbolizes this concept and indicates that endosomes (and MVBs) are signaling compartments, which confer time- and space-resolution to signals that would otherwise be only partially informative, and which add specificity to signaling through a variety of molecular mechanisms.

A number of specific features of endosomes make them ideal for both signal propagation and specificity. Endosomes are characterized by:

1) a limited volume that may favor the interaction between receptor and ligand when the two are internalized in the same vesicle by a mass action kinetic, further sustaining receptor activity;

2) a relatively longer resident time of activated receptors, with respect to the PM from which active receptors are rapidly removed as a consequence of internalization;

3) a scaffold-promoting microenvironment, thanks to their enrichment in particular lipids or proteins, such as the PI3P, that are able to assemble specific signaling complexes;

4) a chemically defined microenvironment characterized by low pH that favors specific reactions, such as proteolysis of signaling molecules, deubiquitination of ubiquitinated

receptors by DUBs, or dephosphorylation by phosphatases, which are all very active in this compartment.

Finally, endosomes can influence signaling either by sustaining signals originating from the PM or by contributing to signal specificity through their provision of a platform for the assembly of specific signaling complexes that are prevented at the PM (reviewed in Gould and Lippincott-Schwartz 2009 and Scita and Di Fiore 2010).

These observations highlight the need to construct a full-predictive and time-resolved model of EGFR endocytosis and signaling by extending the early model from the first 2 minutes following EGF stimulation to the triggering process of EGFR internalization, sorting and downstream signaling. To this end, we will perform the modeling phases in parallel by:

- experimentally exploring the EGFR biology in terms of endocytosis, as well as the later steps of endosomal sorting and EGF-dependent signaling outputs;
- formalizing into the model the biological observations that result from these experiments, and generating a feedback loop in which modeling and wet lab experiments are interconnected in a back-to-back approach.

1.1 - Converting EAM into a quantitative model

The EAM is able to adequately account for the early stages of the EGFR activation, generating results in good agreement with data and parameters coming from the literature. However, the EAM is intrinsically a semi-quantitative model, built on normalized curves for both EGFR Ub and pY, and it is not time-resolved.

To develop a full quantitative model, we are faced with the problem of deconvoluting the simplifications that we introduced into the EAM. First of all, the fact that EGFR phosphorylation is modeled based on the contribution of only 3 phosphotyrosines (the ones relevant to Cbl/Grb2 recruitment), rather than the contribution of the 9 pY residues present in the system. In addition, the model was designed in such a way that the three tyrosines become

phosphorylated independently of each other and follow the same kinetics of phosphorylation. Although these properties were validated experimentally (in the case of independency, **Fig.27**) or taken from literature (in the case of time kinetics Olsen, Blagoev et al. 2006), we acknowledge that we do not know whether these properties can be extended to the other phosphorylation sites.

We have now extended the EAM to consider all 9 tyrosine residues and performed a preliminary analysis *in silico* of how the extended EAM model performs in terms of absolute phosphorylation within the first 2 min of EGF stimulation (**Fig. 47**). The model showed that the kinetics of phosphorylation of the single tyrosine residues is very rapid and almost all receptors are modified at least in one residue after few seconds of EGFR activation. Overall, in a time window of 20-40 seconds, almost all receptors are phosphorylated in five out of nine sites. Only ~10% of receptors are predicted to be modified on all the nine tyrosine residues after 2 minutes.

In order to understand whether this behavior reflects what happens *in vivo*, we are currently performing a series of experiments, including the analysis of EGFR phosphorylation kinetics at early (< 2 minutes) and late (> 2 minutes) time points after EGFR activation, with a particular focus on the kinetics of single phosphosites in the contest of the WT EGFR. For this analysis, we will use specific phospho-EGFR antibodies (we have available 8 antibodies out of the 9 phosphosites).

Moreover, we are performing a quantitative analysis of total EGFR phosphorylation and ubiquitination, in an attempt to measure the number of receptors that are modified at 2 minutes and eventually at earlier or later time points. This will be done by comparing the amount of EGFR immunoprecipitated with anti-Ub or anti-pY antibodies under stringent lysis conditions (i.e., 1% SDS, see Methods Section Protein Buffers), in order to understand the absolute number of EGFR molecules that undergo ubiquitination and phosphorylation at the

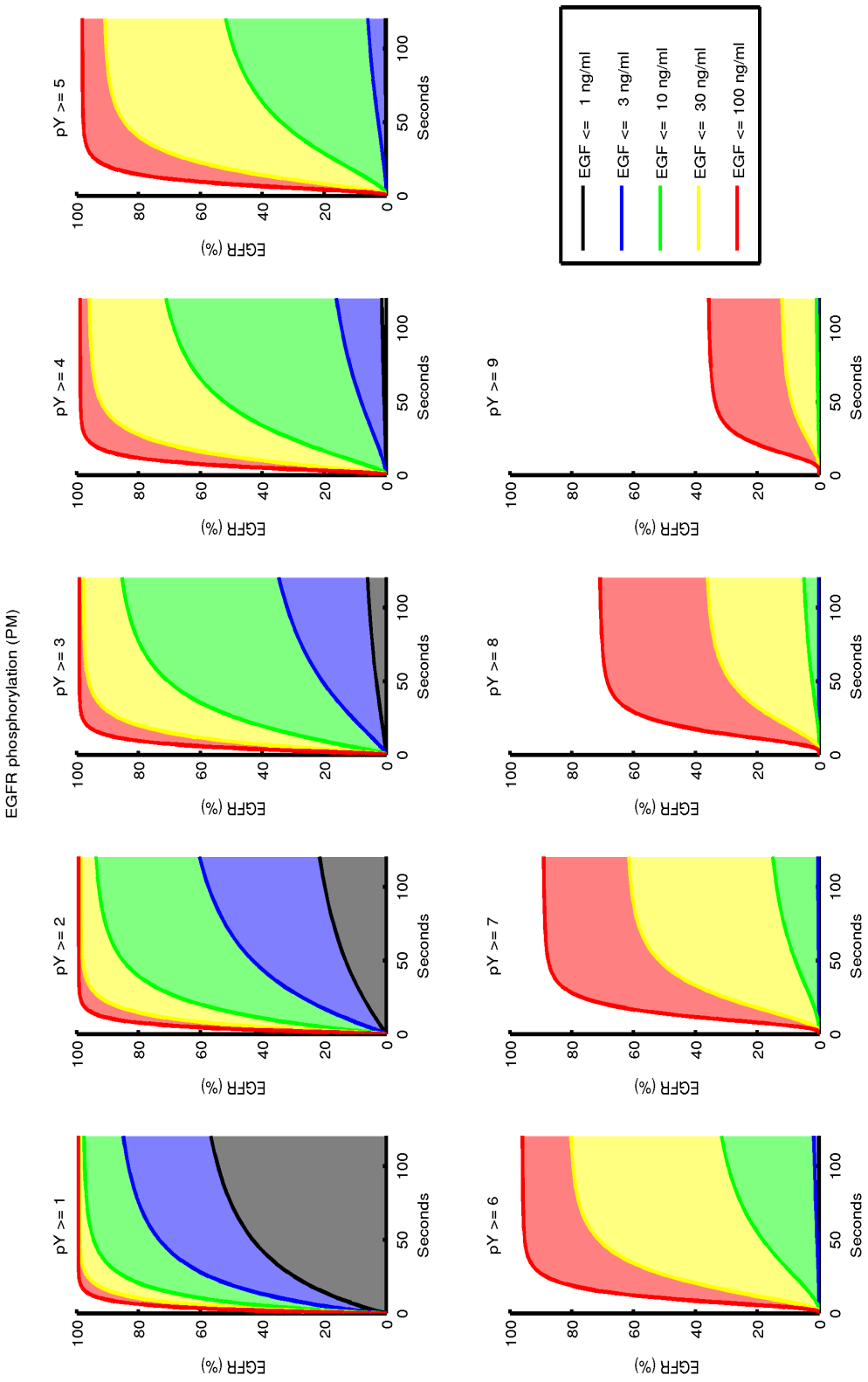


Figure 47: Extended EAM simulation of EGFR phosphorylation following differentially phosphorylated receptors. The EAM was challenged to describe the behavior of EGFR phosphorylation by grouping together receptors with the same number of phosphate groups. The nine panels represent the number of EGFR that are phosphorylated at increasing number of tyrosines, at the indicated times on the same EGFR moiety at different EGF doses. The behavior at different EGF concentrations is indicated with different color lines: black 0.1-1.0 ng/ml EGF, blue 1.0-3.0 ng/ml EGF, green 3.0-10 ng/ml EGF, yellow 10-30 ng/ml EGF, red 30-300 ng/ml EGF. On the x-axis the time after EGF stimulation is indicated and on the y-axis the percentage of EGFR molecules phosphorylated with respect to the total number of EGFRs/cell. [This figure was generated by Marchetti, L. (COSBI) with the developing version of the advanced model]

different EGF doses. Experimentally, this estimation will be performed using the same approach used for the quantitation of pY-Cbl (see Results Section 3.1).

We are also interested in studying the behavior of each single tyrosine in an EGFR backbone where all the other tyrosines are absent, through the generation of single EGFR add-back mutants. In our lab, we previously engineered an EGFR mutant in which all the nine phosphorylatable tyrosines in the cytoplasmic tail were mutagenized into phenylalanine (9Y-, **Fig. 27A**), in order to obtain a phosphorylation null receptor. We are now adding back single tyrosine residues to the 9Y- tyrosine null backbone to generate 9 single add-back mutants that harbor only one out of nine phosphorylatable tyrosines, in order to characterize the phosphorylation kinetics of the single phosphosites and to understand the impact of their phosphorylation on the remaining sites.

These experiments will allow us to understand whether, and, if so, how, we have to modify the EAM, in order to convert it in absolute quantitative terms. The extension of the 3 phosphosite-based EAM to a 9 phosphosite-based EAM, while increasing the computational efforts in terms of the number of ODE equations, represents a pre-requisite to convert the model in absolute quantitative terms. The newly generated model will serve as a starting point from which to extend the modeling of EGFR activation, endocytosis and downstream signaling.

2 - EGFR knockout HeLa cells

Data presented in Results Section 5.1 confirmed that the NIH 3T3 clones represent a useful tool to test EGFR-Ub/pY predictions in a homogeneous background, successfully validating the EAM. However, we have collected evidence indicating that this cell model system might not be suitable for the study of the downstream EGFR response. Indeed, NIH 3T3 cells:

- are mouse cells engineered to express human EGFR;

- express ~5000 endogenous mouse EGFR molecules, which may affect the EGF-dependent signaling response;
- are not characterized in terms of EGFR endocytosis.

We initially attempted to investigate the EGFR entry routes in NIH 3T3 cells by performing ^{125}I-EGF internalization kinetic analysis following KD of key players in EGFR endocytosis. We optimized the KD conditions for clathrin (specifically blocks CME), dynamin 1/2 (should block both CME and NCE), and RTN3 (specifically blocks NCE, Caldieri, Barbier et al. 2015) in NIH 3T3 cells. However, the results we obtained from the internalization experiments were difficult to interpret because the inhibition of one internalization pathway seemed to upregulate the other endocytic routes. We are currently performing additional experiments to better clarify the endocytic mechanisms acting in these cells.

However, since the EAM was built on experimental observations generated in HeLa cells, we are now generating HeLa knockout (KO) cells to be reconstituted with human EGFR. The goal of this alternative approach is to obtain a well-characterized cellular setting in which to perform internalization assays aimed at correlating different EGFR expression levels with the selection of a particular entry route.

We managed to generate a HeLa-derived cell line in which endogenous EGFR has been knocked out using the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)/Cas (CRISP ASsociated protein) technology for genome editing (Ran, Hsu et al. 2013). The CRISPR/Cas system originated as part of a prokaryotic immune system to confer resistance to foreign genetic elements, but in the last two years, the system has been used for gene editing. By delivering the Cas9 protein and an appropriate guide RNA (gRNA) into a cell, the organism's genome can be cut at any desired location. Cas9 is a nuclease that was first discovered as a component of the CRISPR system in *Streptococcus pyogenes* and has been adapted for the use in mammalian cells (Jinek, Chylinski et al. 2012). The WT Cas9

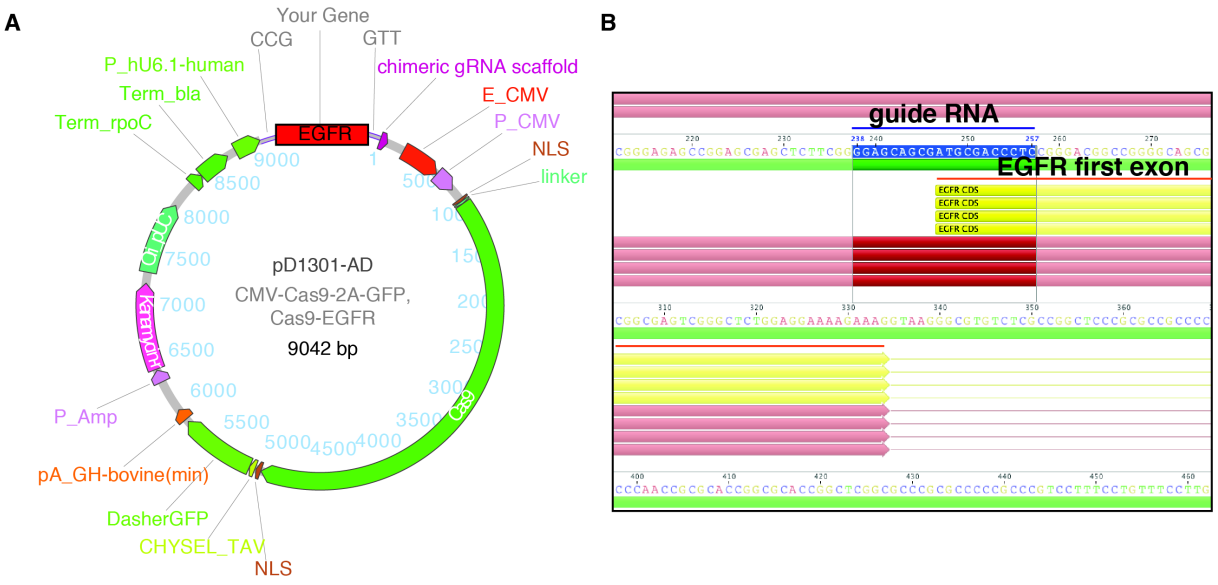


Figure 48: CRISPR/Cas9 technology. A - pD1301-AD plasmid for mammalian transient transfection. The vector carries the Cas9 DNA nuclease and the RNA guide (gRNA), together with a GFP selection element that was used to enrich transfected HeLa cells by FACS sorting. **B -** schematic representation of the mapping of gRNA on the human EGFR coding sequence (CDS). Note that the gRNA is centered exactly on the ATG of the EGFR CDS. (Image generated with Geneious software)

endonuclease is guided to the target site by the gRNA, which anneals to the specific target sequence on the genome, and is able to introduce double-stranded breaks (DSB) with high efficiency. Usually this kind of lesion, in the absence of a DNA template that can be used by the homology-directed repair (HDR) machinery, is likely to activate the non-homologous end joining (NHEJ) repair system (Sander and Joung 2014) that randomly repairs DNA lesions. If overlapping or located in proximity to the ATG of the gene of interest, this lesion can determine a functional KO by various mechanisms, including frameshift mutations, disruption of the promoter or elimination of the initiation codon.

To generate the HeLa EGFR KO cell line, we transfected WT HeLa cells with a plasmid encoding both the gRNA targeting the first exon of EGFR gene and the Cas9 enzyme (**Fig. 48A**), along with GFP for selection. Cas9 is guided to the correct place on the genome by the complementarity of the gRNA with its target site, inducing a DSB. In our case, we guided the gRNA on the first EGFR exon (**Fig. 48B**). This event will leave on the genome a lesion, usually identified as a small insertion or deletion (**Fig. 49A**). After enrichment of the

transfected population by FACS sorting (**Fig. 49B**), HeLa cells were sub-cloned and some candidates were tested for EGFR level both by ^{125}I-EGF saturation binding assay (data not shown) and WB analysis (**Fig. 49C**). The technology revealed to be efficient, allowing the identification of several clones that have lost EGFR protein expression, concomitant to impaired ^{125}I-EGF binding.

Based on the morphology and the growth rate of the clones tested in **Fig. 49C**, we chose to further characterize clones #9 and #13 (red asterisks in **Fig. 49C**). To better understand the type of lesion that was introduced in each clone, we are now performing DNA sequencing at the single allele level, in order to exclude aberrant EGFR truncated products. To this aim, we are setting up, together with the Sequencing Facility at our Institute, a Fragment Analysis protocol coupled to standard Sanger sequencing (Yang, Steentoft et al. 2015); a procedure that will allow the definition - up to the single base level – of the kind of the lesion introduced.

2.1 - Reconstitution of HeLa EGFR KO cells with WT or mutated EGFR

The HeLa EGFR KO cells will be the starting point for the generation of isogenic cell lines expressing different levels of WT EGFR, through the reintroduction of the human EGFR gene cloned into a viral vector, as performed in NIH-3T3 clones (see Results Section 8.1). The resulting cell lines will be investigated for EGFR-Ub/pY levels and for the selection of the internalization route (e.g., CME *vs.* NCE) upon variation of EGFR/EGF levels. The HeLa cell model system represents an ideal starting point for the characterization of the EGFR endocytic route in relation to varying EGFR/EGF levels, because the parental cell line has been extensively studied by our group in terms of endocytosis (Sigismund, Argenzio et al. 2008; Sigismund, Algisi et al. 2013).

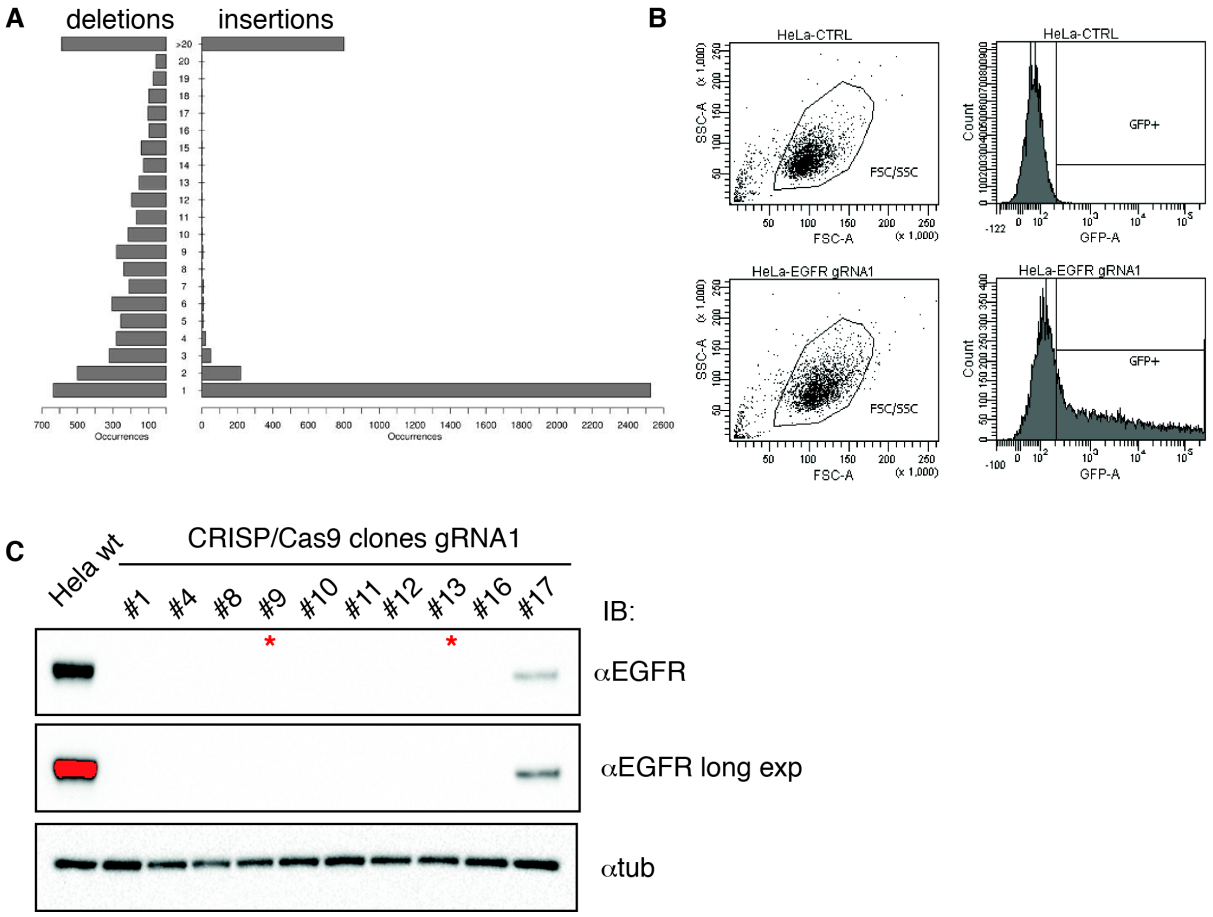


Figure 49: Generation of HeLa EGFR KO cell lines. A - Schematic representation of the frequency of occurrence of insertions/deletions (in/del) at the targeted locus. Note the most frequent mutations that usually occour are very small (1-2 bp) in/dels. **B -** FACS analysis of GFP-positive HeLa cells. Cells were either mock treated (CTRL, top panels) or transfected with pD1301-AD-gRNA1 plasmid (HeLa-EGFR-gRNA1, bottom panels). The threshold for selection of positive cells was set at the threshold determined by WT cells. Positive cells were sorted and replated for clone selection. **C -** Ten clones were subjected to WB analysis with the indicated antibodies in comparison to HeLa WT cells. Tubulin (tub) was used as loading control. Long exp: long exposure. EGFR protein was undetectable (compared to the HeLa WT control, left) in all clones, with the exception of one clone (#17). Asterisks indicates the selected clones that are currently under analysis by sequencing to characterize the genetic lesion introduced, and that will be used for reconstitution experiments with WT or mutated human EGFR.

We also intend to exploit HeLa-reconstituted clones to investigate the possible correlation between variations in EGFR/EGF levels and EGF-dependent biological readouts. We will follow downstream EGFR signaling (e.g. AKT/PI3K, ERK1/2, Ras activation) and biological readouts (e.g. migration and/or proliferation) in the selected clones, expressing different EGFR surface levels, at different ligand concentrations.

Finally, we will generate HeLa EGFR KO cells expressing EGFR add-back mutants to study the kinetics of the single phosphosites (see Results Section 3). The single add-back mutants that have already been generated will be expressed in the newly generated HeLa EGFR KO cells and clones expressing levels of mutant EGFR comparable to levels of endogenous EGFR will be selected. Analysis of the phosphorylation kinetics and the dose-response curves of the single phosphosites will be performed. These add-back mutants will be also useful to study the role of individual phosphorylation sites in the entry of EGFR into a given endocytic pathway or in the coupling with a specific signaling response.

Together, these experiments we will generate a more complete picture of the EGFR system, from the very early activation steps to late events that take place inside the cell after receptor stimulation, and possibly their biological implications. Experiments will drive mathematical modeling by exploring EGFR endocytosis and EGF-dependent signaling, and will provide a way to validate predictions that will eventually be derived from the time-resolved model.

3 - Validation of cancer-relevant model predictions using primary cells

Since the EAM was successful in generating cancer-relevant predictions that were validated in cancer cell lines, we plan to extend the validation of these predictions to primary human cancer cells, using samples from the primary cell culture facility at the European Institute of Oncology (IEO) in Milan. The analysis will include lung and/or breast primary tumor cell cultures stratified on the basis of their EGFR status (WT, mutated or overexpressed) alongside other standard clinico-pathological parameters. We will determine the basal status of EGFR by immune-histochemistry on formalin-fixed, paraffin-embedded (FFPE) specimens.

With the obtained primary cell lines, we plan to test the impact of EGFR/EGF levels on:

- EGFR-Ub and -pY curves by DELFIA;
- differential CME/NCE entry by the ^{125}I-EGF internalization assay;
- EGFR degradation/signaling.

Finally, new experiments will be designed according to new predictions generated by the time-resolved model, with the aim to shed light on the link between deregulation of EGFR biology and human cancers.

Appendix 1

Reaction type or species	Symbol	Optimization Range (Model)	Value (Model)	Optimization Range (Typical units)	Value (Typical units)	Reference
Total concentration of EGFR[a]	R_T	N.O.	0.83 nM	N.O.	250000 molecules	This work
Concentration of Cbl available for EGFR interaction[b]	Cbl_T	N.O.	$1.5x10^{-2}$ nM	N.O.	5000 molecules	This work
Total concentration of Grb2[c]	$Grb2_T$	N.O.	3.3 nM	N.O.	1000000 molecules	This work
Hill coefficient for phopshorylation curve	n_H	0.5-1.5	1.13	0.5-1.5	1.13	This work
Half-maximum phosphorylation of EGFR	H	0.1-10 nM	4.6 nM	0.58-58.8 ng/ml	27 ng/ml	This work
Maximum phosphorylation rate constant	k_{KIN}^{MAX}	0.005-10 s^{-1}	0.0759 s^{-1}	0.005-10 s^{-1}	0.0759 s^{-1}	> 0.005 s^{-1} (Olsen, Blagoev et al. 2006); 0.0733 s^{-1} (Pryor, Low-Nam et al. 2013); > 0.075 s^{-1}/ 2 s^{-1} [depending on model in (Kleiman, Maiwald et al. 2011)]; 0.2 s^{-1} (Monast, Furcht et al. 2012); 1s^{-1} (Kozer, Barua et al. 2013)
Dephosphorylation rate constant	kptp	N.O.	0.016 s^{-1}	N.O.	0.016 s^{-1}	0.016 s^{-1} (Kleiman, Maiwald et al. 2011); <0.1 s^{-1} (Olsen, Blagoev et al. 2006); 0.01 s^{-1} (Kholodenko, Demin et al. 1999); 0.13 s^{-1} (Monast, Furcht et al. 2012); 1s^{-1} (Kozer, Barua et al. 2013)
Kd of of pY1068/pY1086 and Grb2 binding (and of pY1045 and Cbl)	K45 (K68)	0.2-150 nM	0.2 nM	0.01-7 µM	0.01 µM	0.4 µM (Lemmon, Ladbury et al. 1994); 0.7 µM (Chook, Gish et al. 1996); 0.1 µM (Cussac, Frech et al. 1994); 0.35 µM (Hu and Hubbard 2005)
Kd of Cbl and Grb2 binding	Kcg	0.006-30 nM	0.006 nM	0.0003-1.5 µM	0.3 nM	1.48 nM (Sastry, Lin et al. 1995); 400 nM (Chook, Gish et al. 1996); 1.8 µM (Lemmon, Bu et al. 1997); 300–500 µM (Houtman, Yamaguchi et al. 2006); 4 µM and 120 µM [Nter-SH3 and CterSH3of Grb2 (Cussac, Frech et al. 1994)]
Rate constant of dissociation of pY1068/pY1086 and Grb2 pY1045 and Cbl (and of pY1045 and Cbl)	ku_{45} (ku_{68})	0.001-40 s^{-1}	0.001 s^{-1}	0.001-40 s^{-1}	0.01 s^{-1}	0.2 s^{-1} (Chook, Gish et al. 1996); 0.4-4 s^{-1} (Cussac, Frech et al. 1994)
Rate constant of dissociation of Cbl and Grb2	ku_{cg}	0.01-100 s^{-1}	0.3 s^{-1}	0.01-100 s^{-1}	0.3 s^{-1}	Not available
Localization factor	f_{LOC}	10-10^5 nM	$2x10^4$ nM	0.5 µM -5 mM	1 mM	This work
Maximum ubiquitination	Ubmax	0.0005-0.05 nM	0.0143 nM	0.003-0.3 ng/ml	0.083 ng/ml	This work
Maximum phosphorylation (MPMB)	pYmax	0-2.5 nM	2.28 nM	0-15 ng/ml	13 ng/ml	This work

EAM parameters						
Maximum phosphorylation	pYmax	0-2.5 nM	2.35 nM	0-15 ng/ml	13.8 ng/ml	This work
K_d of EGF-EGFR binding (high affinity)	K	0.1-20 nM	3.4 nM	0.1-20 nM	3.4 nM	0.2 nM (Macdonald et al. 2008), 4.3 nM (French et al. 1995); 0.67 nM (Waters et al. 1990); 20 nM **(Kholodenko** et al. **1999)**; 20 nM **(Klein** et al. **2004)**
EGF-EGFR association rate constant (to a preformed EGFR dimer)	kb	0.02-20 $nM^{-1}s^{-1}$	5 $nM^{-1}s^{-1}$	0.02-20 (5) $nM^{-1}s^{-1}$	5 $nM^{-1}s^{-1}$	0.2–0.4 $nM^{-1}s^{-1}$ (first EGF moiety), 2 $nM^{-1}s^{-1}$ (second EGF moiety) (Teramura et al. 2006)
K_d of EGF-EGFR binding (low affinity)	K1	N.O.	50 nM	N.O.	50 nM	2.9 nM (Macdonald et al. 2008); 600 nM **(Klein** et al. **2004)**
EGF-EGFR association rate constant (low affinity)	k1b	N.O.	$3x10^{-3}$ $nM^{-1}s^{-1}$	N.O.	0.003 $nM^{-1}s^{-1}$	0.001 $nM^{-1}s^{-1}$ (French et al. 1995); 0.003 $nM^{-1}s^{-1}$ (Waters et al. 1990); 0.004 $nM^{-1}s^{-1}$ (Teramura et al. 2006); 0.001 $nM^{-1}s^{-1}$ **(Monast** et al. **2012)**; $3x10^{-3}$ $nM^{-1}s^{-1}$ **(Kholodenko** et al. **1999)**
Equilibrium ratio of EGFR closed/open	K_e	N.O.	30	N.O.	30	20-30 **(Klein** et al. **2004)**
Rate of closing of extracellular EGFR domain	k_c	N.O.	10 s^{-1}	N.O.	10 s^{-1}	Not available
As k_c when EGFR has EGF bound	$k1_c$	N.O.	10 s^{-1}	N.O.	10 s^{-1}	Not available
Dimerization rate constant	k_b^{DIM} (D)	1-200 $nM^{-1}s^{-1}$	200 nM^{-1} s^{-1}	0.01-0.2 μm^2 s^{-1}	0.200 μm^2 s^{-1}	0.01-0.02 μm^2 s^{-1} (Kusumi et al. 1993); 0.05 μm^2 s^{-1} (Low-Nam et al. 2011); 0.2 μm^2 s^{-1} (Chung et al. 2010)
Dimer dissociation rate constant	k_u^{DIM}	0.07–20 s^{-1}	1.1s^{-1}	0.07–20 s^{-1}	1.1s^{-1}	2 s^{-1} (Chung et al.2010); 1.24s^{-1} (no EGF bound), 0.738s^{-1} (1 EGF bound), and 0.271s^{-1} (2 EGF bound) (Low-Nam et al. 2011)
Phosphorylation rate constant	k_{KIN}	0.005-10 s^{-1}	0.289 s^{-1}	0.005-10 s^{-1}	0.289 s^{-1}	> 0.005 s^{-1} (Olsen et al. 2006); 0.0733 s^{-1} (Pryor et al. 2013); 0.2 s^{-1} **(Monast** et al. **2012)**; > 0.075 s^{-1}/ 2 s^{-1} [depending on model in (Kleiman et al. 2011)]; 1s^{-1} Kozer et al. 2013)

Appendix 1: Model Parameters. Parameters or molecular species were either measured ("this work") or taken from the literature. We distinguish between experimental measurements (references in black) and estimates produced by previous modeling studies (references in red). Parameters were either fitted in a given interval or not optimized (N.O.).

a. Assuming that HeLa cells contain an average of 2.5-3 x 10^5 EGFRs (measured by saturation binding assay, see Materials and Methods) and that each cell is surrounded by 0.5×10^{-6} ml of medium.

b. We estimated the amount of Cbl available for EGF interaction by measuring the levels of Cbl phosphorylated by the EGFR at the maximal EGF concentration (see Result Section 8.1). This represents the maximal "active Cbl" fraction, i.e. able to recruit E2 ubiquitin-conjugating enzymes and to ubiquitinate the EGFR. Thus, we take this as a proxy for the maximal amount of Cbl actually available for binding to the EGFR (measured in **Fig. 35**).

c. Grb2 levels as measured in **Fig. 35**.

d. In the MPM, we set the active EGFR fraction as an Hill function of EGF and optimized the parameters via a best-fit procedure; see Supplementary Information in (Capuani, Conte et al. 2015).

Acknowledgements

I thank professor Pier Paolo Di Fiore for allowing me to join his group in such an excellent institute. I started my adventure in IFOM 6 years ago, for the University master thesis and I was given the opportunity to stay to follow also my PhD project.

I really thank Sara Sigismund for everyday help and support and for directing me with patience during all this years.

I also thank my internal supervisor, Andrea Ciliberto, and my external supervisor, Ivan Dikic, for their useful suggestions and for the enthusiasm they showed in following this project.

I thank Rosalind Gunby for critically reading this work and for her advices.

This PhD project was supported by IFOM fellowship and a personal grant from the Fondazione Italiana per la Ricerca sul Cancro (IG 12972).

I feel also to say thanks to all lab A, to girls and boys who became not only colleagues but most importantly friends. Some of them are still in the lab, some others decided to leave for new job and personal experiences. I fondly remember Veronica, who welcomed me the very first day in the lab and taught me so many things. My lab-bench companions, Claudia and Roberta, with which I spent so many hours inside and outside the lab! *Vipers-island* will never die in our memory :-)

Time goes by and endo group still remains: Elisa, Giusi, Irene, Giovanni… thank you all!

All lab A members were essential to make lab A what it is... (actually a nursery school, but sometimes we also talk about science!)

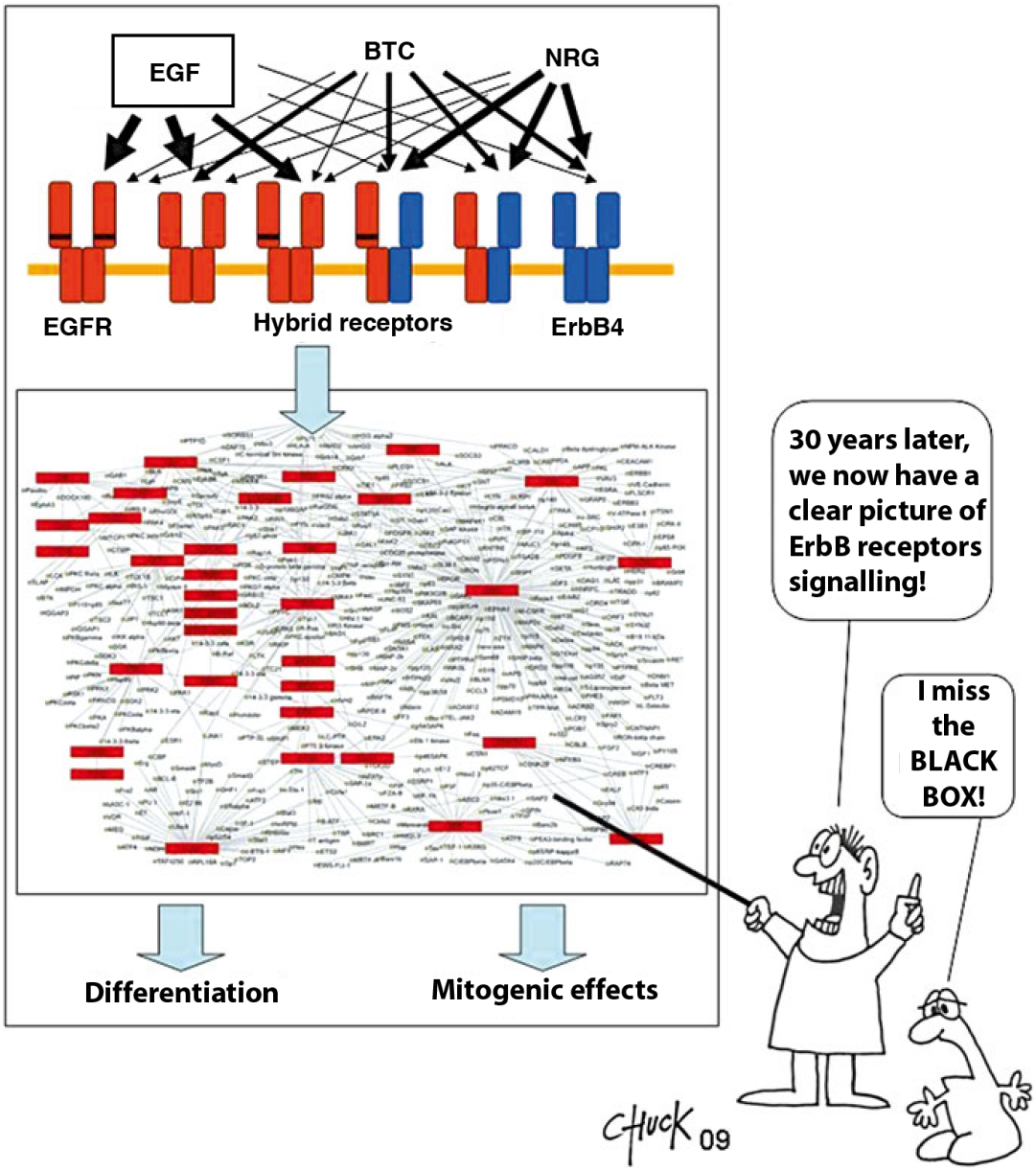
There are some lab-mates that become great friends in real life: Eleonora, Simona, Chiara, Marilena and their boys… What to say? Thanks for Japanese dinners, parties and nights out, trips and weddings!!!

Now let's switch to Italian for those who would never know they've been acknowledged in English... Grazie alla mia famiglia per avermi cresciuto e fatto diventare la persona che sono, nel bene e nel male. Grazie alla mamma per i sacrifici che ha fatto per me e per essere stata una amica oltre che una madre… mi dispiace non potere stare vicine (il che è anche una protezione per entrambe!). Un affettuoso pensiero alla nonna, che ci ha lasciato esattamente un anno fa, una nonna che avrebbe fatto di tutto per me, anche se non le ho mai dimostrato molta riconoscenza; grazie a mio zio Claudio, Sandra e Laura (e famiglia).

Come nella tesi della laurea triennale e della laurea specialistica, un pensiero particolare va ad Aaron, cui ora si aggiungono Casper e Pedro, perché ripeto ormai con una frase di rito: *senza di loro non esisterebbe (la mia) vita sulla terra e io non sarei ciò che sono!*

In questi quattro anni una persona molto importante si è aggiunta nella mia vita, e forse a lui più di tutti vanno i miei ringraziamenti per avermi sopportato, consolato, fatto arrabbiare, fatto irare, nei momenti più bui di questo cammino, quando anche avevo pensato di non potere reggere il peso fino alla fine. Bruno, la mia metà opposta, la mia vita.

L'ultimo pensiero va a te, che avresti voluto esserci e che mi manchi terribilmente sempre di più, ma la vita a volte decide diversamente. Anche se non siamo d'accordo. Grazie anche a te perché so che più di tutti, tu saresti fiero di me…



(Adapted from Pierre De Meyts 2015)

Bibliography

Abella, J. V., Peschard, P., Naujokas, M. A., Lin, T., Saucier, C., Urbe, S. and Park, M. (2005). **Met/Hepatocyte growth factor receptor ubiquitination suppresses transformation and is required for Hrs phosphorylation.** Molecular and cellular biology 25(21): 9632-9645.

Alroy, I. and Yarden, Y. (1997). **The ErbB signaling network in embryogenesis and oncogenesis: signal diversification through combinatorial ligand-receptor interactions.** FEBS Lett 410(1): 83-86.

Amit, I., Wides, R. and Yarden, Y. (2007). **Evolvable signaling networks of receptor tyrosine kinases: relevance of robustness to malignancy and to cancer therapy.** Mol Syst Biol 3: 151.

Anderson, R. G., Brown, M. S. and Goldstein, J. L. (1977). **Role of the coated endocytic vesicle in the uptake of receptor-bound low density lipoprotein in human fibroblasts.** Cell 10(3): 351-364.

Andl, C. D., Mizushima, T., Oyama, K., Bowser, M., Nakagawa, H. and Rustgi, A. K. (2004). **EGFR-induced cell migration is mediated predominantly by the JAK-STAT pathway in primary esophageal keratinocytes.** Am J Physiol Gastrointest Liver Physiol 287(6): G1227-1237.

Ang, K. K., Berkey, B. A., Tu, X., Zhang, H. Z., Katz, R., Hammond, E. H., Fu, K. K. and Milas, L. (2002). **Impact of epidermal growth factor receptor expression on survival and pattern of relapse in patients with advanced head and neck carcinoma.** Cancer Res 62(24): 7350-7356.

Anklesaria, P., Teixido, J., Laiho, M., Pierce, J. H., Greenberger, J. S. and Massague, J. (1990). **Cell-cell adhesion mediated by binding of membrane-anchored transforming growth factor alpha to epidermal growth factor receptors promotes cell proliferation.** Proc Natl Acad Sci U S A 87(9): 3289-3293.

Arteaga, C. L. (2002). **Epidermal growth factor receptor dependence in human tumors: more than just expression?** The oncologist 7 Suppl 4: 31-39.

Arteaga, C. L. and Engelman, J. A. (2014). **ERBB receptors: from oncogene discovery to basic science to mechanism-based cancer therapeutics.** Cancer cell 25(3): 282-303.

Bach, I. and Ostendorff, H. P. (2003). **Orchestrating nuclear functions: ubiquitin sets the rhythm.** Trends Biochem Sci 28(4): 189-195.

Bang, Y. J., Van Cutsem, E., Feyereislova, A., Chung, H. C., Shen, L., Sawaki, A., Lordick, F., Ohtsu, A., Omuro, Y., Satoh, T., Aprile, G., Kulikov, E., Hill, J., Lehle, M., Ruschoff, J., Kang, Y. K. and To, G. A. T. I. (2010). **Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial.** Lancet 376(9742): 687-697.

Bargmann, C. I., Hung, M. C. and Weinberg, R. A. (1986). **Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185.** Cell 45(5): 649-657.

Barker, A. J., Gibson, K. H., Grundy, W., Godfrey, A. A., Barlow, J. J., Healy, M. P., Woodburn, J. R., Ashton, S. E., Curry, B. J., Scarlett, L., Henthorn, L. and Richards, L. (2001). **Studies leading to the identification of ZD1839 (IRESSA): an orally active, selective epidermal growth factor receptor tyrosine kinase inhibitor targeted to the treatment of cancer.** Bioorg Med Chem Lett 11(14): 1911-1914.

Barton, W. A., Tzvetkova-Robev, D., Miranda, E. P., Kolev, M. V., Rajashankar, K. R., Himanen, J. P. and Nikolov, D. B. (2006). **Crystal structures of the Tie2 receptor**

ectodomain and the angiopoietin-2-Tie2 complex. Nature structural & molecular biology 13(6): 524-532.

Batzer, A. G., Rotin, D., Urena, J. M., Skolnik, E. Y. and Schlessinger, J. (1994). **Hierarchy of binding sites for Grb2 and Shc on the epidermal growth factor receptor.** Mol Cell Biol 14(8): 5192-5201.

Birchmeier, C., Birchmeier, W., Gherardi, E. and Vande Woude, G. F. (2003). **Met, metastasis, motility and more.** Nat Rev Mol Cell Biol 4(12): 915-925.

Birtwistle, M. R., Hatakeyama, M., Yumoto, N., Ogunnaike, B. A., Hoek, J. B. and Kholodenko, B. N. (2007). **Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses.** Molecular systems biology 3: 144.

Boeri Erba, E., Matthiesen, R., Bunkenborg, J., Schulze, W. X., Di Stefano, P., Cabodi, S., Tarone, G., Defilippi, P. and Jensen, O. N. (2007). **Quantitation of multisite EGF receptor phosphorylation using mass spectrometry and a novel normalization approach.** J Proteome Res 6(7): 2768-2785.

Boucrot, E., Ferreira, A. P., Almeida-Souza, L., Debard, S., Vallis, Y., Howard, G., Bertot, L., Sauvonnet, N. and McMahon, H. T. (2015). **Endophilin marks and controls a clathrin-independent endocytic pathway.** Nature 517(7535): 460-465.

Brankatschk, B., Wichert, S. P., Johnson, S. D., Schaad, O., Rossner, M. J. and Gruenberg, J. (2012). **Regulation of the EGF transcriptional response by endocytic sorting.** Sci Signal 5(215): ra21.

Bromberg, J. (2002). **Stat proteins and oncogenesis.** J Clin Invest 109(9): 1139-1142.

Brown, P. I., Lam, R., Lakshmanan, J. and Fisher, D. A. (1990). **Transforming growth factor alpha in developing rats.** Am J Physiol 259(2 Pt 1): E256-260.

Bublil, E. M. and Yarden, Y. (2007). **The EGF receptor family: spearheading a merger of signaling and therapeutics.** Current opinion in cell biology 19(2): 124-134.

Burgess, A. W., Cho, H. S., Eigenbrot, C., Ferguson, K. M., Garrett, T. P., Leahy, D. J., Lemmon, M. A., Sliwkowski, M. X., Ward, C. W. and Yokoyama, S. (2003). **An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors.** Mol Cell 12(3): 541-552.

Busser, B., Sancey, L., Josserand, V., Niang, C., Khochbin, S., Favrot, M. C., Coll, J. L. and Hurbin, A. (2010). **Amphiregulin promotes resistance to gefitinib in nonsmall cell lung cancer cells by regulating Ku70 acetylation.** Mol Ther 18(3): 536-543.

Caldieri, G., Barbieri, E., Nappo, G., Raimondi, A., Conte, A., Polo, S., Di Fiore, P. and Sigismund, S. (2015). A Novel Pathway of Internalization depend on Contacts between the Endoplasmic Reticulum and the Plasma Membrane.

Capuani*, F., Conte*, A., Argenzio, E., Marchetti, L., Priami, C., Polo, S., Di Fiore, P. P., Sigismund, S. and Ciliberto, A. (2015). **Quantitative analysis reveals how EGFR activation and downregulation are coupled in normal but not in cancer cells.** Nat Commun 6: 7999.

Carbone, R., Fre, S., Iannolo, G., Belleudi, F., Mancini, P., Pelicci, P. G., Torrisi, M. R. and Di Fiore, P. P. (1997). **eps15 and eps15R are essential components of the endocytic pathway.** Cancer Res 57(24): 5498-5504.

Carpenter, C. L., Auger, K. R., Chanudhuri, M., Yoakim, M., Schaffhausen, B., Shoelson, S. and Cantley, L. C. (1993). **Phosphoinositide 3-kinase is activated by phosphopeptides that bind to the SH2 domains of the 85-kDa subunit.** J Biol Chem 268(13): 9478-9483.

Carpenter, G. and Cohen, S. (1990). **Epidermal growth factor.** The Journal of biological chemistry 265(14): 7709-7712.

Carter, T. A., Wodicka, L. M., Shah, N. P., Velasco, A. M., Fabian, M. A., Treiber, D. K., Milanov, Z. V., Atteridge, C. E., Biggs, W. H., 3rd, Edeen, P. T., Floyd, M., Ford, J. M., Grotzfeld, R. M., Herrgard, S., Insko, D. E., Mehta, S. A., Patel, H. K., Pao, W.,

Sawyers, C. L., Varmus, H., Zarrinkar, P. P. and Lockhart, D. J. (2005). **Inhibition of drug-resistant mutants of ABL, KIT, and EGF receptor kinases.** Proc Natl Acad Sci U S A 102(31): 11011-11016.

Chattopadhyay, A., Vecchi, M., Ji, Q., Mernaugh, R. and Carpenter, G. (1999). **The role of individual SH2 domains in mediating association of phospholipase C-gamma1 with the activated EGF receptor.** J Biol Chem 274(37): 26091-26097.

Chen, W. W., Schoeberl, B., Jasper, P. J., Niepel, M., Nielsen, U. B., Lauffenburger, D. A. and Sorger, P. K. (2009). **Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data.** Molecular systems biology 5: 239.

Cheng, Z. J., Singh, R. D., Sharma, D. K., Holicky, E. L., Hanada, K., Marks, D. L. and Pagano, R. E. (2006). **Distinct mechanisms of clathrin-independent endocytosis have unique sphingolipid requirements.** Molecular biology of the cell 17(7): 3197-3210.

Chook, Y. M., Gish, G. D., Kay, C. M., Pai, E. F. and Pawson, T. (1996). **The Grb2-mSos1 complex binds phosphopeptides with higher affinity than Grb2.** The Journal of biological chemistry 271(48): 30472-30478.

Chung, I., Akita, R., Vandlen, R., Toomre, D., Schlessinger, J. and Mellman, I. (2010). **Spatial control of EGF receptor activation by reversible dimerization on living cells.** Nature 464(7289): 783-787.

Ciehanover, A., Hod, Y. and Hershko, A. (1978). **A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes.** Biochem Biophys Res Commun 81(4): 1100-1105.

Citri, A. and Yarden, Y. (2006). **EGF-ERBB signalling: towards the systems level.** Nature reviews. Molecular cell biology 7(7): 505-516.

Clague, M. J., Liu, H. and Urbe, S. (2012). **Governance of endocytic trafficking and signaling by reversible ubiquitylation.** Developmental cell 23(3): 457-467.

Clayton, A. H., Walker, F., Orchard, S. G., Henderson, C., Fuchs, D., Rothacker, J., Nice, E. C. and Burgess, A. W. (2005). **Ligand-induced dimer-tetramer transition during the activation of the cell surface epidermal growth factor receptor-A multidimensional microscopy analysis.** The Journal of biological chemistry 280(34): 30392-30399.

Comoglio, P. M. and Trusolino, L. (2002). **Invasive growth: from development to metastasis.** J Clin Invest 109(7): 857-862.

Cussac, D., Frech, M. and Chardin, P. (1994). **Binding of the Grb2 SH2 domain to phosphotyrosine motifs does not change the affinity of its SH3 domains for Sos proline-rich motifs.** The EMBO journal 13(17): 4011-4021.

D'Souza-Schorey, C. and Chavrier, P. (2006). **ARF proteins: roles in membrane traffic and beyond.** Nat Rev Mol Cell Biol 7(5): 347-358.

Dikic, I. (2003). **Mechanisms controlling EGF receptor endocytosis and degradation.** Biochem Soc Trans 31(Pt 6): 1178-1181.

Dikic, I. and Schmidt, M. H. (2007). **Malfunctions within the Cbl interactome uncouple receptor tyrosine kinases from destructive transport.** European journal of cell biology 86(9): 505-512.

Discafani, C. M., Carroll, M. L., Floyd, M. B., Jr., Hollander, I. J., Husain, Z., Johnson, B. D., Kitchen, D., May, M. K., Malo, M. S., Minnick, A. A., Jr., Nilakantan, R., Shen, R., Wang, Y. F., Wissner, A. and Greenberger, L. M. (1999). **Irreversible inhibition of epidermal growth factor receptor tyrosine kinase with in vivo activity by N-[4-[(3-bromophenyl)amino]-6-quinazolinyl]-2-butynamide (CL-387,785).** Biochem Pharmacol 57(8): 917-925.

Doody, J. F., Wang, Y., Patel, S. N., Joynes, C., Lee, S. P., Gerlak, J., Rolser, R. L., Li, Y., Steiner, P., Bassi, R., Hicklin, D. J. and Hadari, Y. R. (2007). **Inhibitory activity of cetuximab on epidermal growth factor receptor mutations in non small cell lung cancers.** Mol Cancer Ther 6(10): 2642-2651.

Dou, H., Buetow, L., Hock, A., Sibbet, G. J., Vousden, K. H. and Huang, D. T. (2012). **Structural basis for autoinhibition and phosphorylation-dependent activation of c-Cbl.** Nature structural & molecular biology 19(2): 184-192.

Dubois, E. A. and Cohen, A. F. (2009). **Panitumumab.** Br J Clin Pharmacol 68(4): 482-483.

Edeling, M. A., Smith, C. and Owen, D. (2006). **Life of a clathrin coat: insights from clathrin and AP structures.** Nature reviews. Molecular cell biology 7(1): 32-44.

Ellis, S. and Mellor, H. (2000). **Regulation of endocytic traffic by rho family GTPases.** Trends Cell Biol 10(3): 85-88.

Emde, A., Kostler, W. J., Yarden, Y., Association of, R. and Oncology of the Mediterranean, a. (2012). **Therapeutic strategies and mechanisms of tumorigenesis of HER2-overexpressing breast cancer.** Crit Rev Oncol Hematol 84 Suppl 1: e49-57.

Endres, N. F., Das, R., Smith, A. W., Arkhipov, A., Kovacs, E., Huang, Y., Pelton, J. G., Shan, Y., Shaw, D. E., Wemmer, D. E., Groves, J. T. and Kuriyan, J. (2013). **Conformational coupling across the plasma membrane in activation of the EGF receptor.** Cell 152(3): 543-556.

Enuka, Y., Feldman, E. and Yarden, Y. (2015). **Computational and Modeling Aspects of RTK Networks**. Receptor Tyrosine Kinases: Structure, Functions and Role in Human Disease. Wheeler D. and Y. Yarden, Humana Press.

Farah, R. A., Clinchy, B., Herrera, L. and Vitetta, E. S. (1998). **The development of monoclonal antibodies for the therapy of cancer.** Critical reviews in eukaryotic gene expression 8(3-4): 321-356.

Fivaz, M., Vilbois, F., Thurnheer, S., Pasquali, C., Abrami, L., Bickel, P. E., Parton, R. G. and van der Goot, F. G. (2002). **Differential sorting and fate of endocytosed GPI-anchored proteins.** EMBO J 21(15): 3989-4000.

Frangioni, J. V., Beahm, P. H., Shifrin, V., Jost, C. A. and Neel, B. G. (1992). **The nontransmembrane tyrosine phosphatase PTP-1B localizes to the endoplasmic reticulum via its 35 amino acid C-terminal sequence.** Cell 68(3): 545-560.

Frick, M., Bright, N. A., Riento, K., Bray, A., Merrified, C. and Nichols, B. J. (2007). **Coassembly of flotillins induces formation of membrane microdomains, membrane curvature, and vesicle budding.** Current biology : CB 17(13): 1151-1156.

Fry, D. W., Bridges, A. J., Denny, W. A., Doherty, A., Greis, K. D., Hicks, J. L., Hook, K. E., Keller, P. R., Leopold, W. R., Loo, J. A., McNamara, D. J., Nelson, J. M., Sherwood, V., Smaill, J. B., Trumpp-Kallmeyer, S. and Dobrusin, E. M. (1998). **Specific, irreversible inactivation of the epidermal growth factor receptor and erbB2, by a new class of tyrosine kinase inhibitor.** Proc Natl Acad Sci U S A 95(20): 12022-12027.

Gaestel, M. (2006). **MAPKAP kinases - MKs - two's company, three's a crowd.** Nat Rev Mol Cell Biol 7(2): 120-130.

Gajria, D. and Chandarlapaty, S. (2011). **HER2-amplified breast cancer: mechanisms of trastuzumab resistance and novel targeted therapies.** Expert Rev Anticancer Ther 11(2): 263-275.

Galan, J. M. and Haguenauer-Tsapis, R. (1997). **Ubiquitin lys63 is involved in ubiquitination of a yeast plasma membrane protein.** EMBO J 16(19): 5847-5854.

Gan, H. K., Cvrljevic, A. N. and Johns, T. G. (2013). **The epidermal growth factor receptor variant III (EGFRvIII): where wild things are altered.** FEBS J 280(21): 5350-5370.

Garrett, T. P., McKern, N. M., Lou, M., Elleman, T. C., Adams, T. E., Lovrecz, G. O., Zhu, H. J., Walker, F., Frenkel, M. J., Hoyne, P. A., Jorissen, R. N., Nice, E. C., Burgess, A. W. and Ward, C. W. (2002). **Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha.** Cell 110(6): 763-773.

Gilbertson, R., Hernan, R., Pietsch, T., Pinto, L., Scotting, P., Allibone, R., Ellison, D., Perry, R., Pearson, A. and Lunec, J. (2001). **Novel ERBB4 juxtamembrane splice variants are frequently expressed in childhood medulloblastoma.** Genes Chromosomes Cancer 31(3): 288-294.

Glebov, O. O., Bright, N. A. and Nichols, B. J. (2006). **Flotillin-1 defines a clathrin-independent endocytic pathway in mammalian cells.** Nature cell biology 8(1): 46-54.

Goh, L. K., Huang, F., Kim, W., Gygi, S. and Sorkin, A. (2010). **Multiple mechanisms collectively regulate clathrin-mediated endocytosis of the epidermal growth factor receptor.** J Cell Biol 189(5): 871-883.

Gould, G. W. and Lippincott-Schwartz, J. (2009). **New roles for endosomes: from vesicular carriers to multi-purpose platforms.** Nat Rev Mol Cell Biol 10(4): 287-292.

Grandis, J. R., Chakraborty, A., Zeng, Q., Melhem, M. F. and Tweardy, D. J. (1998). **Downmodulation of TGF-alpha protein expression with antisense oligonucleotides inhibits proliferation of head and neck squamous carcinoma but not normal mucosal epithelial cells.** J Cell Biochem 69(1): 55-62.

Grovdal, L. M., Stang, E., Sorkin, A. and Madshus, I. H. (2004). **Direct interaction of Cbl with pTyr 1045 of the EGF receptor (EGFR) is required to sort the EGFR to lysosomes for degradation.** Experimental cell research 300(2): 388-395.

Gschwind, A., Fischer, O. M. and Ullrich, A. (2004). **The discovery of receptor tyrosine kinases: targets for cancer therapy.** Nat Rev Cancer 4(5): 361-370.

Guo, L., Kozlosky, C. J., Ericsson, L. H., Daniel, T. O., Cerretti, D. P. and Johnson, R. S. (2003). **Studies of ligand-induced site-specific phosphorylation of epidermal growth factor receptor.** J Am Soc Mass Spectrom 14(9): 1022-1031.

Haber, D. A., Bell, D. W., Sordella, R., Kwak, E. L., Godin-Heymann, N., Sharma, S. V., Lynch, T. J. and Settleman, J. (2005). **Molecular targeted therapy of lung cancer: EGFR mutations and response to EGFR inhibitors.** Cold Spring Harb Symp Quant Biol 70: 419-426.

Haglund, K. and Dikic, I. (2012). **The role of ubiquitylation in receptor endocytosis and endosomal sorting.** Journal of cell science 125(Pt 2): 265-275.

Haglund, K., Sigismund, S., Polo, S., Szymkiewicz, I., Di Fiore, P. P. and Dikic, I. (2003). **Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation.** Nature cell biology 5(5): 461-466.

Hallberg, B., Rayter, S. I. and Downward, J. (1994). **Interaction of Ras and Raf in intact mammalian cells upon extracellular stimulation.** J Biol Chem 269(6): 3913-3916.

Hammond, D. E., Urbe, S., Vande Woude, G. F. and Clague, M. J. (2001). **Down-regulation of MET, the receptor for hepatocyte growth factor.** Oncogene 20(22): 2761-2770.

Harris, R. C., Chung, E. and Coffey, R. J. (2003). **EGF receptor ligands.** Exp Cell Res 284(1): 2-13.

Hatakeyama, H., Cheng, H., Wirth, P., Counsell, A., Marcrom, S. R., Wood, C. B., Pohlmann, P. R., Gilbert, J., Murphy, B., Yarbrough, W. G., Wheeler, D. L., Harari, P. M., Guo, Y., Shyr, Y., Slebos, R. J. and Chung, C. H. (2010). **Regulation of heparin-binding EGF-like growth factor by miR-212 and acquired cetuximab-resistance in head and neck squamous cell carcinoma.** PLoS One 5(9): e12702.

Haura, E. B., Zheng, Z., Song, L., Cantor, A. and Bepler, G. (2005). **Activated epidermal growth factor receptor-Stat-3 signaling promotes tumor survival in vivo in non-small cell lung cancer.** Clin Cancer Res 11(23): 8288-8294.

Hicke, L. and Riezman, H. (1996). **Ubiquitination of a yeast plasma membrane receptor signals its ligand-stimulated endocytosis.** Cell 84(2): 277-287.

Hill, A. V. (1913). **The Combinations of Haemoglobin with Oxygen and with Carbon Monoxide. I.** The Biochemical journal 7(5): 471-480.

Hill, C. S. and Treisman, R. (1995). **Transcriptional regulation by extracellular signals: mechanisms and specificity.** Cell 80(2): 199-211.

Himanen, J. P. and Nikolov, D. B. (2003). **Eph signaling: a structural view.** Trends in neurosciences 26(1): 46-51.

Hirsh, V. (2011). **Managing treatment-related adverse events associated with egfr tyrosine kinase inhibitors in advanced non-small-cell lung cancer.** Curr Oncol 18(3): 126-138.

Holbro, T. and Hynes, N. E. (2004). **ErbB receptors: directing key signaling networks throughout life.** Annu Rev Pharmacol Toxicol 44: 195-217.

Hou, X., Johnson, A. C. and Rosner, M. R. (1994). **Identification of an epidermal growth factor receptor transcriptional repressor.** J Biol Chem 269(6): 4307-4312.

Houtman, J. C., Yamaguchi, H., Barda-Saad, M., Braiman, A., Bowden, B., Appella, E., Schuck, P. and Samelson, L. E. (2006). **Oligomerization of signaling complexes by the multipoint binding of GRB2 to both LAT and SOS1.** Nature structural & molecular biology 13(9): 798-805.

Hu, J. and Hubbard, S. R. (2005). **Structural characterization of a novel Cbl phosphotyrosine recognition motif in the APS family of adapter proteins.** The Journal of biological chemistry 280(19): 18943-18949.

Huang, F., Goh, L. K. and Sorkin, A. (2007). **EGF receptor ubiquitination is not necessary for its internalization.** Proceedings of the National Academy of Sciences of the United States of America 104(43): 16904-16909.

Huang, F., Kirkpatrick, D., Jiang, X., Gygi, S. and Sorkin, A. (2006). **Differential regulation of EGF receptor internalization and degradation by multiubiquitination within the kinase domain.** Molecular cell 21(6): 737-748.

Huang, T. T. and D'Andrea, A. D. (2006). **Regulation of DNA repair by ubiquitylation.** Nat Rev Mol Cell Biol 7(5): 323-334.

Huang, T. T., Nijman, S. M., Mirchandani, K. D., Galardy, P. J., Cohn, M. A., Haas, W., Gygi, S. P., Ploegh, H. L., Bernards, R. and D'Andrea, A. D. (2006). **Regulation of monoubiquitinated PCNA by DUB autocleavage.** Nat Cell Biol 8(4): 339-347.

Hubbard, S. R. and Miller, W. T. (2007). **Receptor tyrosine kinases: mechanisms of activation and signaling.** Current opinion in cell biology 19(2): 117-123.

Hubbard, S. R. and Till, J. H. (2000). **Protein tyrosine kinase structure and function.** Annu Rev Biochem 69: 373-398.

Hudson, L. G., Thompson, K. L., Xu, J. and Gill, G. N. (1990). **Identification and characterization of a regulated promoter element in the epidermal growth factor receptor gene.** Proc Natl Acad Sci U S A 87(19): 7536-7540.

Hudson, P. J. (1999). **Recombinant antibody constructs in cancer therapy.** Current opinion in immunology 11(5): 548-557.

Huyer, G., Liu, S., Kelly, J., Moffat, J., Payette, P., Kennedy, B., Tsaprailis, G., Gresser, M. J. and Ramachandran, C. (1997). **Mechanism of inhibition of protein-tyrosine phosphatases by vanadate and pervanadate.** J Biol Chem 272(2): 843-851.

Hynes, N. E. and Lane, H. A. (2005). **ERBB receptors and cancer: the complexity of targeted inhibitors.** Nature reviews. Cancer 5(5): 341-354.

Inui, S., Higashiyama, S., Hashimoto, K., Higashiyama, M., Yoshikawa, K. and Taniguchi, N. (1997). **Possible role of coexpression of CD9 with membrane-anchored heparin-binding EGF-like growth factor and amphiregulin in cultured human keratinocyte growth.** J Cell Physiol 171(3): 291-298.

Jaiswal, B. S., Kljavin, N. M., Stawiski, E. W., Chan, E., Parikh, C., Durinck, S., Chaudhuri, S., Pujara, K., Guillory, J., Edgar, K. A., Janakiraman, V., Scholz, R. P., Bowman, K. K., Lorenzo, M., Li, H., Wu, J., Yuan, W., Peters, B. A., Kan, Z., Stinson, J., Mak, M., Modrusan, Z., Eigenbrot, C., Firestein, R., Stern, H. M., Rajalingam, K., Schaefer, G., Merchant, M. A., Sliwkowski, M. X., de Sauvage, F. J. and Seshagiri, S. (2013). **Oncogenic ERBB3 mutations in human cancers.** Cancer Cell 23(5): 603-617.

Jedlinski, A., Ansell, A., Johansson, A. C. and Roberg, K. (2013). **EGFR status and EGFR ligand expression influence the treatment response of head and neck cancer cell lines.** Journal of oral pathology & medicine : official publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology 42(1): 26-36.

Ji, H., Zhao, X., Yuza, Y., Shimamura, T., Li, D., Protopopov, A., Jung, B. L., McNamara, K., Xia, H., Glatt, K. A., Thomas, R. K., Sasaki, H., Horner, J. W., Eck, M., Mitchell, A., Sun, Y., Al-Hashem, R., Bronson, R. T., Rabindran, S. K., Discafani, C. M., Maher, E., Shapiro, G. I., Meyerson, M. and Wong, K. K. (2006). **Epidermal growth factor receptor variant III mutations in lung tumorigenesis and sensitivity to tyrosine kinase inhibitors.** Proc Natl Acad Sci U S A 103(20): 7817-7822.

Jiang, X., Huang, F., Marusyk, A. and Sorkin, A. (2003). **Grb2 regulates internalization of EGF receptors through clathrin-coated pits.** Mol Biol Cell 14(3): 858-870.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012). **A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.** Science 337(6096): 816-821.

Jorissen, R. N., Walker, F., Pouliot, N., Garrett, T. P., Ward, C. W. and Burgess, A. W. (2003). **Epidermal growth factor receptor: mechanisms of activation and signalling.** Experimental cell research 284(1): 31-53.

Kaleko, M., Rutter, W. J. and Miller, A. D. (1990). **Overexpression of the human insulinlike growth factor I receptor promotes ligand-dependent neoplastic transformation.** Mol Cell Biol 10(2): 464-473.

Katara, G. K., Jaiswal, M. K., Kulshrestha, A., Kolli, B., Gilman-Sachs, A. and Beaman, K. D. (2014). **Tumor-associated vacuolar ATPase subunit promotes tumorigenic characteristics in macrophages.** Oncogene 33(49): 5649-5654.

Kholodenko, B. N., Demin, O. V., Moehren, G. and Hoek, J. B. (1999). **Quantification of short term signaling by the epidermal growth factor receptor.** The Journal of biological chemistry 274(42): 30169-30181.

Kholodenko, B. N., Kiyatkin, A., Bruggeman, F. J., Sontag, E., Westerhoff, H. V. and Hoek, J. B. (2002). **Untangling the wires: a strategy to trace functional interactions in signaling and gene networks.** Proc Natl Acad Sci U S A 99(20): 12841-12846.

Kirchhausen, T., Owen, D. and Harrison, S. C. (2014). **Molecular structure, function, and dynamics of clathrin-mediated membrane traffic.** Cold Spring Harbor perspectives in biology 6(5): a016725.

Kirkham, M., Fujita, A., Chadda, R., Nixon, S. J., Kurzchalia, T. V., Sharma, D. K., Pagano, R. E., Hancock, J. F., Mayor, S. and Parton, R. G. (2005). **Ultrastructural identification of uncoated caveolin-independent early endocytic vehicles.** J Cell Biol 168(3): 465-476.

Kleiman, L. B., Maiwald, T., Conzelmann, H., Lauffenburger, D. A. and Sorger, P. K. (2011). **Rapid phospho-turnover by receptor tyrosine kinases impacts downstream signaling and drug binding.** Molecular cell 43(5): 723-737.

Klein, P., Mattoon, D., Lemmon, M. A. and Schlessinger, J. (2004). **A structure-based model for ligand binding and dimerization of EGF receptors.** Proc Natl Acad Sci U S A 101(4): 929-934.

Kloth, M. T., Catling, A. D. and Silva, C. M. (2002). **Novel activation of STAT5b in response to epidermal growth factor.** J Biol Chem 277(10): 8693-8701.

Kobayashi, S., Ji, H., Yuza, Y., Meyerson, M., Wong, K. K., Tenen, D. G. and Halmos, B. (2005). **An alternative inhibitor overcomes resistance caused by a mutation of the epidermal growth factor receptor.** Cancer Res 65(16): 7096-7101.

Kolling, R. and Hollenberg, C. P. (1994). **The ABC-transporter Ste6 accumulates in the plasma membrane in a ubiquitinated form in endocytosis mutants.** EMBO J 13(14): 3261-3271.

Kon, S., Kobayashi, N. and Satake, M. (2014). **Altered trafficking of mutated growth factor receptors and their associated molecules: implication for human cancers.** Cellular logistics 4: e28461.

Kopetz, S. (2007). **Targeting SRC and epidermal growth factor receptor in colorectal cancer: rationale and progress into the clinic.** Gastrointest Cancer Res 1(4 Suppl 2): S37-41.

Kozer, N., Barua, D., Orchard, S., Nice, E. C., Burgess, A. W., Hlavacek, W. S. and Clayton, A. H. (2013). **Exploring higher-order EGFR oligomerisation and phosphorylation--a combined experimental and theoretical approach.** Molecular bioSystems 9(7): 1849-1863.

Kujubu, D. A., Norman, J. T., Herschman, H. R. and Fine, L. G. (1991). **Primary response gene expression in renal hypertrophy and hyperplasia: evidence for different growth initiation processes.** Am J Physiol 260(6 Pt 2): F823-827.

Lammers, R., Bossenmaier, B., Cool, D. E., Tonks, N. K., Schlessinger, J., Fischer, E. H. and Ullrich, A. (1993). **Differential activities of protein tyrosine phosphatases in intact cells.** J Biol Chem 268(30): 22456-22462.

Le Roy, C. and Wrana, J. L. (2005). **Clathrin- and non-clathrin-mediated endocytic regulation of cell signalling.** Nat Rev Mol Cell Biol 6(2): 112-126.

Lemmon, M. A. (2009). **Ligand-induced ErbB receptor dimerization.** Experimental cell research 315(4): 638-648.

Lemmon, M. A., Bu, Z., Ladbury, J. E., Zhou, M., Pinchasi, D., Lax, I., Engelman, D. M. and Schlessinger, J. (1997). **Two EGF molecules contribute additively to stabilization of the EGFR dimer.** EMBO J 16(2): 281-294.

Lemmon, M. A., Ladbury, J. E., Mandiyan, V., Zhou, M. and Schlessinger, J. (1994). **Independent binding of peptide ligands to the SH2 and SH3 domains of Grb2.** The Journal of biological chemistry 269(50): 31653-31658.

Lemmon, M. A. and Schlessinger, J. (2010). **Cell signaling by receptor tyrosine kinases.** Cell 141(7): 1117-1134.

Lemmon, M. A., Schlessinger, J. and Ferguson, K. M. (2014). **The EGFR family: not so prototypical receptor tyrosine kinases.** Cold Spring Harbor perspectives in biology 6(4): a020768.

Leu, T. H. and Maa, M. C. (2003). **Functional implication of the interaction between EGF receptor and c-Src.** Front Biosci 8: s28-38.

Levkowitz, G., Waterman, H., Ettenberg, S. A., Katz, M., Tsygankov, A. Y., Alroy, I., Lavi, S., Iwai, K., Reiss, Y., Ciechanover, A., Lipkowitz, S. and Yarden, Y. (1999). **Ubiquitin ligase activity and tyrosine phosphorylation underlie suppression of growth factor signaling by c-Cbl/Sli-1.** Mol Cell 4(6): 1029-1040.

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

ubiquitination of the epidermal growth factor receptor. Genes Dev 12(23): 3663-3674.

Li, Y., Macdonald-Obermann, J., Westfall, C., Piwnica-Worms, D. and Pike, L. J. (2012). **Quantitation of the effect of ErbB2 on epidermal growth factor receptor binding and dimerization.** The Journal of biological chemistry 287(37): 31116-31125.

Liebmann, C. (2001). **Regulation of MAP kinase activity by peptide receptor signalling pathway: paradigms of multiplicity.** Cell Signal 13(11): 777-785.

Lipkowitz, S. and Weissman, A. M. (2011). **RINGs of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis.** Nat Rev Cancer 11(9): 629-643.

Loeffler-Ragg, J., Schwentner, I., Sprinzl, G. M. and Zwierzina, H. (2008). **EGFR inhibition as a therapy for head and neck squamous cell carcinoma.** Expert Opin Investig Drugs 17(10): 1517-1531.

Lonardo, F., Di Marco, E., King, C. R., Pierce, J. H., Segatto, O., Aaronson, S. A. and Di Fiore, P. P. (1990). **The normal erbB-2 product is an atypical receptor-like tyrosine kinase with constitutive activity in the absence of ligand.** New Biol 2(11): 992-1003.

Lowenstein, E. J., Daly, R. J., Batzer, A. G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E. Y., Bar-Sagi, D. and Schlessinger, J. (1992). **The SH2 and SH3 domain-containing protein GRB2 links receptor tyrosine kinases to ras signaling.** Cell 70(3): 431-442.

Lundmark, R., Doherty, G. J., Howes, M. T., Cortese, K., Vallis, Y., Parton, R. G. and McMahon, H. T. (2008). **The GTPase-activating protein GRAF1 regulates the CLIC/GEEC endocytic pathway.** Curr Biol 18(22): 1802-1808.

Lynch, T. J., Bell, D. W., Sordella, R., Gurubhagavatula, S., Okimoto, R. A., Brannigan, B. W., Harris, P. L., Haserlat, S. M., Supko, J. G., Haluska, F. G., Louis, D. N., Christiani, D. C., Settleman, J. and Haber, D. A. (2004). **Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib.** N Engl J Med 350(21): 2129-2139.

Macdonald, J. L. and Pike, L. J. (2008). **Heterogeneity in EGF-binding affinities arises from negative cooperativity in an aggregating system.** Proceedings of the National Academy of Sciences of the United States of America 105(1): 112-117.

Mak, H. H., Peschard, P., Lin, T., Naujokas, M. A., Zuo, D. and Park, M. (2007). **Oncogenic activation of the Met receptor tyrosine kinase fusion protein, Tpr-Met, involves exclusion from the endocytic degradative pathway.** Oncogene 26(51): 7213-7221.

Marks, B., Stowell, M. H., Vallis, Y., Mills, I. G., Gibson, A., Hopkins, C. R. and McMahon, H. T. (2001). **GTPase activity of dynamin and resulting conformation change are essential for endocytosis.** Nature 410(6825): 231-235.

Marshall, C. J. (1995). **Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation.** Cell 80(2): 179-185.

Maruyama, I. N. (2014). **Mechanisms of activation of receptor tyrosine kinases: monomers or dimers.** Cells 3(2): 304-330.

Mattoon, D. R., Lamothe, B., Lax, I. and Schlessinger, J. (2004). **The docking protein Gab1 is the primary mediator of EGF-stimulated activation of the PI-3K/Akt cell survival pathway.** BMC Biol 2: 24.

Mayor, S. and Pagano, R. E. (2007). **Pathways of clathrin-independent endocytosis.** Nature reviews. Molecular cell biology 8(8): 603-612.

McClellan, M., Kievit, P., Auersperg, N. and Rodland, K. (1999). **Regulation of proliferation and apoptosis by epidermal growth factor and protein kinase C in human ovarian surface epithelial cells.** Exp Cell Res 246(2): 471-479.

McKay, M. M. and Morrison, D. K. (2007). **Integrating signals from RTKs to ERK/MAPK.** Oncogene 26(22): 3113-3121.

McMahon, H. T. and Boucrot, E. (2011). **Molecular mechanism and physiological functions of clathrin-mediated endocytosis.** Nature reviews. Molecular cell biology 12(8): 517-533.

Mettlen, M., Pucadyil, T., Ramachandran, R. and Schmid, S. L. (2009). **Dissecting dynamin's role in clathrin-mediated endocytosis.** Biochemical Society transactions 37(Pt 5): 1022-1026.

Meyts, P. D. (2015). **Receptor Tyrosine Kinase Signal Transduction and the Molecular Basis of Signalling Specificity**. Receptor Tyrosine Kinases: Structure, Functions and Role in Human Disease. Wheeler D. and Y. Yarden, Humana Press.

Mialon, A., Sankinen, M., Soderstrom, H., Junttila, T. T., Holmstrom, T., Koivusalo, R., Papageorgiou, A. C., Johnson, R. S., Hietanen, S., Elenius, K. and Westermarck, J. (2005). **DNA topoisomerase I is a cofactor for c-Jun in the regulation of epidermal growth factor receptor expression and cancer cell proliferation.** Molecular and cellular biology 25(12): 5040-5051.

Moasser, M. M. (2007). **The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis.** Oncogene 26(45): 6469-6487.

Monast, C. S., Furcht, C. M. and Lazzara, M. J. (2012). **Computational analysis of the regulation of EGFR by protein tyrosine phosphatases.** Biophysical journal 102(9): 2012-2021.

Moriki, T., Maruyama, H. and Maruyama, I. N. (2001). **Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain.** Journal of molecular biology 311(5): 1011-1026.

Mousavi, S. A., Malerod, L., Berg, T. and Kjeken, R. (2004). **Clathrin-dependent endocytosis.** Biochem J 377(Pt 1): 1-16.

Mroczkowski, B. and Reich, M. (1993). **Identification of biologically active epidermal growth factor precursor in human fluids and secretions.** Endocrinology 132(1): 417-425.

Mullin, J. M. (2004). **Epithelial barriers, compartmentation, and cancer.** Sci STKE 2004(216): pe2.

Nelson, A. L., Dhimolea, E. and Reichert, J. M. (2010). **Development trends for human monoclonal antibody therapeutics.** Nat Rev Drug Discov 9(10): 767-774.

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

Nishikawa, R., Ji, X. D., Harmon, R. C., Lazar, C. S., Gill, G. N., Cavenee, W. K. and Huang, H. J. (1994). **A mutant epidermal growth factor receptor common in human glioma confers enhanced tumorigenicity.** Proc Natl Acad Sci U S A 91(16): 7727-7731.

Normanno, N., Bianco, C., Strizzi, L., Mancino, M., Maiello, M. R., De Luca, A., Caponigro, F. and Salomon, D. S. (2005). **The ErbB receptors and their ligands in cancer: an overview.** Curr Drug Targets 6(3): 243-257.

Normanno, N., Campiglio, M., Maiello, M. R., De Luca, A., Mancino, M., Gallo, M., D'Alessio, A. and Menard, S. (2008). **Breast cancer cells with acquired resistance to the EGFR tyrosine kinase inhibitor gefitinib show persistent activation of MAPK signaling.** Breast Cancer Res Treat 112(1): 25-33.

Noselli, S. and Perrimon, N. (2000). **Signal transduction. Are there close encounters between signaling pathways?** Science 290(5489): 68-69.

Oda, K., Matsuoka, Y., Funahashi, A. and Kitano, H. (2005). **A comprehensive pathway map of epidermal growth factor receptor signaling.** Molecular systems biology 1: 2005 0010.

Offterdinger, M., Georget, V., Girod, A. and Bastiaens, P. I. (2004). **Imaging phosphorylation dynamics of the epidermal growth factor receptor.** J Biol Chem 279(35): 36972-36981.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J. H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M. and Yokoyama, S. (2002). **Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains.** Cell 110(6): 775-787.

Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P. and Mann, M. (2006). **Global, in vivo, and site-specific phosphorylation dynamics in signaling networks.** Cell 127(3): 635-648.

Organ, S. L. and Tsao, M. S. (2011). **An overview of the c-MET signaling pathway.** Ther Adv Med Oncol 3(1 Suppl): S7-S19.

Ostman, A. and Bohmer, F. D. (2001). **Regulation of receptor tyrosine kinase signaling by protein tyrosine phosphatases.** Trends Cell Biol 11(6): 258-266.

P, O. C., Rhys-Evans, P. and Eccles, S. (2000). **Expression and regulation of c-ERBB ligands in human head and neck squamous carcinoma cells.** Int J Cancer 88(5): 759-765.

Paez, J. G., Janne, P. A., Lee, J. C., Tracy, S., Greulich, H., Gabriel, S., Herman, P., Kaye, F. J., Lindeman, N., Boggon, T. J., Naoki, K., Sasaki, H., Fujii, Y., Eck, M. J., Sellers, W. R., Johnson, B. E. and Meyerson, M. (2004). **EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy.** Science 304(5676): 1497-1500.

Pao, W., Miller, V., Zakowski, M., Doherty, J., Politi, K., Sarkaria, I., Singh, B., Heelan, R., Rusch, V., Fulton, L., Mardis, E., Kupfer, D., Wilson, R., Kris, M. and Varmus, H. (2004). **EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib.** Proc Natl Acad Sci U S A 101(36): 13306-13311.

Pao, W., Miller, V. A. and Kris, M. G. (2004). **'Targeting' the epidermal growth factor receptor tyrosine kinase with gefitinib (Iressa) in non-small cell lung cancer (NSCLC).** Semin Cancer Biol 14(1): 33-40.

Pao, W., Miller, V. A., Politi, K. A., Riely, G. J., Somwar, R., Zakowski, M. F., Kris, M. G. and Varmus, H. (2005). **Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain.** PLoS Med 2(3): e73.

Patterson, R. L., van Rossum, D. B., Nikolaidis, N., Gill, D. L. and Snyder, S. H. (2005). **Phospholipase C-gamma: diverse roles in receptor-mediated calcium signaling.** Trends Biochem Sci 30(12): 688-697.

Payne, C. K., Jones, S. A., Chen, C. and Zhuang, X. (2007). **Internalization and trafficking of cell surface proteoglycans and proteoglycan-binding ligands.** Traffic 8(4): 389-401.

Pearce, M. J., Mintseris, J., Ferreyra, J., Gygi, S. P. and Darwin, K. H. (2008). **Ubiquitin-like protein involved in the proteasome pathway of Mycobacterium tuberculosis.** Science 322(5904): 1104-1107.

Pelkmans, L. and Zerial, M. (2005). **Kinase-regulated quantal assemblies and kiss-and-run recycling of caveolae.** Nature 436(7047): 128-133.

Penengo, L., Mapelli, M., Murachelli, A. G., Confalonieri, S., Magri, L., Musacchio, A., Di Fiore, P. P., Polo, S. and Schneider, T. R. (2006). **Crystal structure of the ubiquitin binding domains of rabex-5 reveals two modes of interaction with ubiquitin.** Cell 124(6): 1183-1195.

Peruzzi, B. and Bottaro, D. P. (2006). **Targeting the c-Met signaling pathway in cancer.** Clin Cancer Res 12(12): 3657-3660.

Peschard, P., Fournier, T. M., Lamorte, L., Naujokas, M. A., Band, H., Langdon, W. Y. and Park, M. (2001). **Mutation of the c-Cbl TKB domain binding site on the Met receptor tyrosine kinase converts it into a transforming protein.** Mol Cell 8(5): 995-1004.

Peschard, P., Kozlov, G., Lin, T., Mirza, I. A., Berghuis, A. M., Lipkowitz, S., Park, M. and Gehring, K. (2007). **Structural basis for ubiquitin-mediated dimerization and activation of the ubiquitin protein ligase Cbl-b.** Molecular cell 27(3): 474-485.

Pike, L. J. (2009). **The challenge of lipid rafts.** Journal of lipid research 50 Suppl: S323-328.

Platta, H. W. and Stenmark, H. (2011). **Endocytosis and signaling.** Current opinion in cell biology 23(4): 393-403.

Prenzel, N., Zwick, E., Leserer, M. and Ullrich, A. (2000). **Tyrosine kinase signalling in breast cancer. Epidermal growth factor receptor: convergence point for signal integration and diversification.** Breast Cancer Res 2(3): 184-190.

Prickett, T. D., Agrawal, N. S., Wei, X., Yates, K. E., Lin, J. C., Wunderlich, J. R., Cronin, J. C., Cruz, P., Rosenberg, S. A. and Samuels, Y. (2009). **Analysis of the tyrosine kinome in melanoma reveals recurrent mutations in ERBB4.** Nat Genet 41(10): 1127-1132.

Pryor, M. M., Low-Nam, S. T., Halasz, A. M., Lidke, D. S., Wilson, B. S. and Edwards, J. S. (2013). **Dynamic transition states of ErbB1 phosphorylation predicted by spatial stochastic modeling.** Biophysical journal 105(6): 1533-1543.

Raab, G., Kover, K., Paria, B. C., Dey, S. K., Ezzell, R. M. and Klagsbrun, M. (1996). **Mouse preimplantation blastocysts adhere to cells expressing the transmembrane form of heparin-binding EGF-like growth factor.** Development 122(2): 637-645.

Raiborg, C. and Stenmark, H. (2009). **The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins.** Nature 458(7237): 445-452.

Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. and Zhang, F. (2013). **Genome engineering using the CRISPR-Cas9 system.** Nat Protoc 8(11): 2281-2308.

Reddy, C. C., Wells, A. and Lauffenburger, D. A. (1998). **Comparative mitogenic potencies of EGF and TGF alpha and their dependence on receptor-limitation versus ligand-limitation.** Med Biol Eng Comput 36(4): 499-507.

Renard, H. F., Simunovic, M., Lemiere, J., Boucrot, E., Garcia-Castillo, M. D., Arumugam, S., Chambon, V., Lamaze, C., Wunder, C., Kenworthy, A. K., Schmidt, A. A., McMahon, H. T., Sykes, C., Bassereau, P. and Johannes, L. (2015). **Endophilin-A2 functions in membrane scission in clathrin-independent endocytosis.** Nature 517(7535): 493-496.

Resat, H., Ewald, J. A., Dixon, D. A. and Wiley, H. S. (2003). **An integrated model of epidermal growth factor receptor trafficking and signal transduction.** Biophysical journal 85(2): 730-743.

Revillion, F., Lhotellier, V., Hornez, L., Bonneterre, J. and Peyrat, J. P. (2008). **ErbB/HER ligands in human breast cancer, and relationships with their receptors, the bio-pathological features and prognosis.** Ann Oncol 19(1): 73-80.

Reynolds, A. R., Tischer, C., Verveer, P. J., Rocks, O. and Bastiaens, P. I. (2003). **EGFR activation coupled to inhibition of tyrosine phosphatases causes lateral signal propagation.** Nature cell biology 5(5): 447-453.

Riese, D. J., 2nd and Stern, D. F. (1998). **Specificity within the EGF family/ErbB receptor family signaling network.** Bioessays 20(1): 41-48.

Rimawi, M. F., Shetty, P. B., Weiss, H. L., Schiff, R., Osborne, C. K., Chamness, G. C. and Elledge, R. M. (2010). **Epidermal growth factor receptor expression in breast cancer association with biologic phenotype and clinical outcomes.** Cancer 116(5): 1234-1242.

Robertson, S. C., Tynan, J. and Donoghue, D. J. (2000). **RTK mutations and human syndromes: when good receptors turn bad.** Trends in genetics : TIG 16(8): 368.

Roda-Navarro, P. and Bastiaens, P. I. (2014). **Dynamic recruitment of protein tyrosine phosphatase PTPD1 to EGF stimulation sites potentiates EGFR activation.** PloS one 9(7): e103203.

Roepstorff, K., Grandal, M. V., Henriksen, L., Knudsen, S. L., Lerdrup, M., Grovdal, L., Willumsen, B. M. and van Deurs, B. (2009). **Differential effects of EGFR ligands on endocytic sorting of the receptor.** Traffic 10(8): 1115-1127.

Roth, T. F. and Porter, K. R. (1964). **Yolk Protein Uptake in the Oocyte of the Mosquito Aedes Aegypti. L.** The Journal of cell biology 20: 313-332.

Rothberg, K. G., Heuser, J. E., Donzell, W. C., Ying, Y. S., Glenney, J. R. and Anderson, R. G. (1992). **Caveolin, a protein component of caveolae membrane coats.** Cell 68(4): 673-682.

Rusch, V., Klimstra, D., Venkatraman, E., Pisters, P. W., Langenfeld, J. and Dmitrovsky, E. (1997). **Overexpression of the epidermal growth factor receptor and its ligand transforming growth factor alpha is frequent in resectable non-small cell lung cancer but does not predict tumor progression.** Clin Cancer Res 3(4): 515-522.

Ryan, P. E., Davies, G. C., Nau, M. M. and Lipkowitz, S. (2006). **Regulating the regulator: negative regulation of Cbl ubiquitin ligases.** Trends in biochemical sciences 31(2): 79-88.

Sabharanjak, S., Sharma, P., Parton, R. G. and Mayor, S. (2002). **GPI-anchored proteins are delivered to recycling endosomes via a distinct cdc42-regulated, clathrin-independent pinocytic pathway.** Dev Cell 2(4): 411-423.

Saeki, T., Salomon, D. S., Johnson, G. R., Gullick, W. J., Mandai, K., Yamagami, K., Moriwaki, S., Tanada, M., Takashima, S. and Tahara, E. (1995). **Association of epidermal growth factor-related peptides and type I receptor tyrosine kinase receptors with prognosis of human colorectal carcinomas.** Jpn J Clin Oncol 25(6): 240-249.

Sahin, U., Weskamp, G., Kelly, K., Zhou, H. M., Higashiyama, S., Peschon, J., Hartmann, D., Saftig, P. and Blobel, C. P. (2004). **Distinct roles for ADAM10 and ADAM17 in ectodomain shedding of six EGFR ligands.** The Journal of cell biology 164(5): 769-779.

Salomon, D. S., Normanno, N., Ciardiello, F., Brandt, R., Shoyab, M. and Todaro, G. J. (1995). **The role of amphiregulin in breast cancer.** Breast Cancer Res Treat 33(2): 103-114.

Sander, J. D. and Joung, J. K. (2014). **CRISPR-Cas systems for editing, regulating and targeting genomes.** Nature biotechnology 32(4): 347-355.

Sastry, L., Lin, W., Wong, W. T., Di Fiore, P. P., Scoppa, C. A. and King, C. R. (1995). **Quantitative analysis of Grb2-Sos1 interaction: the N-terminal SH3 domain of Grb2 mediates affinity.** Oncogene 11(6): 1107-1112.

Scala, S., Saeki, T., Lynch, A., Salomon, D., Merino, M. J. and Bates, S. E. (1995). **Coexpression of TGF alpha, epidermal growth factor receptor, and P-glycoprotein in normal and benign diseased breast tissues.** Diagn Mol Pathol 4(2): 136-142.

Schechter, A. L., Stern, D. F., Vaidyanathan, L., Decker, S. J., Drebin, J. A., Greene, M. I. and Weinberg, R. A. (1984). **The neu oncogene: an erb-B-related gene encoding a 185,000-Mr tumour antigen.** Nature 312(5994): 513-516.

Schlesinger, D. H., Goldstein, G. and Niall, H. D. (1975). **The complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells.** Biochemistry 14(10): 2214-2218.

Schlessinger, J. (2000). **Cell signaling by receptor tyrosine kinases.** Cell 103(2): 211-225.

Schlessinger, J. and Lemmon, M. A. (2003). **SH2 and PTB domains in tyrosine kinase signaling.** Sci STKE 2003(191): RE12.

Schmid, E. M. and McMahon, H. T. (2007). **Integrating molecular and network biology to decode endocytosis.** Nature 448(7156): 883-888.

Schmidt, M. H. and Dikic, I. (2005). **The Cbl interactome and its functions.** Nature reviews. Molecular cell biology 6(12): 907-918.

Schneider, M. R. and Wolf, E. (2009). **The epidermal growth factor receptor ligands at a glance.** Journal of cellular physiology 218(3): 460-466.

Schneider, M. R. and Yarden, Y. (2014). **Structure and function of epigen, the last EGFR ligand.** Semin Cell Dev Biol 28: 57-61.

Schonwasser, D. C., Marais, R. M., Marshall, C. J. and Parker, P. J. (1998). **Activation of the mitogen-activated protein kinase/extracellular signal-regulated kinase pathway by conventional, novel, and atypical protein kinase C isotypes.** Mol Cell Biol 18(2): 790-798.

Schulze, A., Lehmann, K., Jefferies, H. B., McMahon, M. and Downward, J. (2001). **Analysis of the transcriptional program induced by Raf in epithelial cells.** Genes Dev 15(8): 981-994.

Schulze, W. X., Deng, L. and Mann, M. (2005). **Phosphotyrosine interactome of the ErbB-receptor kinase family.** Molecular systems biology 1: 2005 0008.

Scita, G. and Di Fiore, P. P. (2010). **The endocytic matrix.** Nature 463(7280): 464-473.

Selvaggi, G., Novello, S., Torri, V., Leonardo, E., De Giuli, P., Borasio, P., Mossetti, C., Ardissone, F., Lausi, P. and Scagliotti, G. V. (2004). **Epidermal growth factor receptor overexpression correlates with a poor prognosis in completely resected non-small-cell lung cancer.** Ann Oncol 15(1): 28-32.

Sengupta, P., Bosis, E., Nachliel, E., Gutman, M., Smith, S. O., Mihalyne, G., Zaitseva, I. and McLaughlin, S. (2009). **EGFR juxtamembrane domain, membranes, and calmodulin: kinetics of their interaction.** Biophys J 96(12): 4887-4895.

Shankaran, H., Zhang, Y., Chrisler, W. B., Ewald, J. A., Wiley, H. S. and Resat, H. (2012). **Integrated experimental and model-based analysis reveals the spatial aspects of EGFR activation dynamics.** Molecular bioSystems 8(11): 2868-2882.

Sharma, S. V., Bell, D. W., Settleman, J. and Haber, D. A. (2007). **Epidermal growth factor receptor mutations in lung cancer.** Nature reviews. Cancer 7(3): 169-181.

Shoelson, S. E. (1997). **SH2 and PTB domain interactions in tyrosine kinase signal transduction.** Curr Opin Chem Biol 1(2): 227-234.

Shtiegman, K., Kochupurakkal, B. S., Zwang, Y., Pines, G., Starr, A., Vexler, A., Citri, A., Katz, M., Lavi, S., Ben-Basat, Y., Benjamin, S., Corso, S., Gan, J., Yosef, R. B., Giordano, S. and Yarden, Y. (2007). **Defective ubiquitinylation of EGFR mutants of lung cancer confers prolonged signaling.** Oncogene 26(49): 6968-6978.

Siegelin, M. D. and Borczuk, A. C. (2014). **Epidermal growth factor receptor mutations in lung adenocarcinoma.** Laboratory investigation; a journal of technical methods and pathology 94(2): 129-137.

Sigismund, S., Algisi, V., Nappo, G., Conte, A., Pascolutti, R., Cuomo, A., Bonaldi, T., Argenzio, E., Verhoef, L. G., Maspero, E., Bianchi, F., Capuani, F., Ciliberto, A., Polo, S. and Di Fiore, P. P. (2013). **Threshold-controlled ubiquitination of the EGFR directs receptor fate.** The EMBO journal 32(15): 2140-2157.

Sigismund, S., Argenzio, E., Tosoni, D., Cavallaro, E., Polo, S. and Di Fiore, P. P. (2008). **Clathrin-mediated internalization is essential for sustained EGFR signaling but dispensable for degradation.** Developmental cell 15(2): 209-219.

Sigismund, S., Confalonieri, S., Ciliberto, A., Polo, S., Scita, G. and Di Fiore, P. P. (2012). **Endocytosis and signaling: cell logistics shape the eukaryotic cell plan.** Physiol Rev 92(1): 273-366.

Sigismund, S., Woelk, T., Puri, C., Maspero, E., Tacchetti, C., Transidico, P., Di Fiore, P. P. and Polo, S. (2005). **Clathrin-independent endocytosis of ubiquitinated cargos.** Proceedings of the National Academy of Sciences of the United States of America 102(8): 2760-2765.

Singh, A. B. and Harris, R. C. (2005). **Autocrine, paracrine and juxtacrine signaling by EGFR ligands.** Cellular signalling 17(10): 1183-1193.

Singh, J., Dobrusin, E. M., Fry, D. W., Haske, T., Whitty, A. and McNamara, D. J. (1997). **Structure-based design of a potent, selective, and irreversible inhibitor of the catalytic domain of the erbB receptor subfamily of protein tyrosine kinases.** J Med Chem 40(7): 1130-1135.

Slamon, D. and Pegram, M. (2001). **Rationale for trastuzumab (Herceptin) in adjuvant breast cancer trials.** Semin Oncol 28(1 Suppl 3): 13-19.

Slamon, D. J., Clark, G. M., Wong, S. G., Levin, W. J., Ullrich, A. and McGuire, W. L. (1987). **Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene.** Science 235(4785): 177-182.

Smaill, J. B., Rewcastle, G. W., Loo, J. A., Greis, K. D., Chan, O. H., Reyner, E. L., Lipka, E., Showalter, H. D., Vincent, P. W., Elliott, W. L. and Denny, W. A. (2000). **Tyrosine kinase inhibitors. 17. Irreversible inhibitors of the epidermal growth factor receptor: 4-(Phenylamino)quinazoline- and 4-(Phenylamino)pyrido.** J Med Chem 43(16): 3199.

Smith, J. J., Derynck, R. and Korc, M. (1987). **Production of transforming growth factor alpha in human pancreatic cancer cells: evidence for a superagonist autocrine cycle.** Proc Natl Acad Sci U S A 84(21): 7567-7570.

Sorkin, A., Di Fiore, P. P. and Carpenter, G. (1993). **The carboxyl terminus of epidermal growth factor receptor/erbB-2 chimerae is internalization impaired.** Oncogene 8(11): 3021-3028.

Sorkin, A., Mazzotti, M., Sorkina, T., Scotto, L. and Beguinot, L. (1996). **Epidermal growth factor receptor interaction with clathrin adaptors is mediated by the Tyr974-containing internalization motif.** J Biol Chem 271(23): 13377-13384.

Sorkin, A. and von Zastrow, M. (2009). **Endocytosis and signalling: intertwining molecular networks.** Nature reviews. Molecular cell biology 10(9): 609-622.

Sorkin, A., Waters, C., Overholser, K. A. and Carpenter, G. (1991). **Multiple autophosphorylation site mutations of the epidermal growth factor receptor. Analysis of kinase activity and endocytosis.** J Biol Chem 266(13): 8355-8362.

Spano, J. P., Lagorce, C., Atlan, D., Milano, G., Domont, J., Benamouzig, R., Attar, A., Benichou, J., Martin, A., Morere, J. F., Raphael, M., Penault-Llorca, F., Breau, J. L., Fagard, R., Khayat, D. and Wind, P. (2005). **Impact of EGFR expression on colorectal cancer patient prognosis and survival.** Annals of oncology : official journal of the European Society for Medical Oncology / ESMO 16(1): 102-108.

Spence, J., Sadis, S., Haas, A. L. and Finley, D. (1995). **A ubiquitin mutant with specific defects in DNA repair and multiubiquitination.** Mol Cell Biol 15(3): 1265-1273.

Stang, E., Blystad, F. D., Kazazic, M., Bertelsen, V., Brodahl, T., Raiborg, C., Stenmark, H. and Madshus, I. H. (2004). **Cbl-dependent ubiquitination is required for progression of EGF receptors into clathrin-coated pits.** Mol Biol Cell 15(8): 3591-3604.

Sugawa, N., Ekstrand, A. J., James, C. D. and Collins, V. P. (1990). **Identical splicing of aberrant epidermal growth factor receptor transcripts from amplified rearranged genes in human glioblastomas.** Proceedings of the National Academy of Sciences of the United States of America 87(21): 8602-8606.

Summy, J. M. and Gallick, G. E. (2006). **Treatment for advanced tumors: SRC reclaims center stage.** Clin Cancer Res 12(5): 1398-1401.

Tajeddine, N. and Gailly, P. (2012). **TRPC1 protein channel is major regulator of epidermal growth factor receptor signaling.** J Biol Chem 287(20): 16146-16157.

Tao, R. H. and Maruyama, I. N. (2008). **All EGF(ErbB) receptors have preformed homo- and heterodimeric structures in living cells.** Journal of cell science 121(Pt 19): 3207-3217.

Tatulian, S. A. (2015). **Structural Dynamics of Insulin Receptor and Transmembrane Signaling.** Biochemistry 54(36): 5523-5532.

Thogersen, V. B., Sorensen, B. S., Poulsen, S. S., Orntoft, T. F., Wolf, H. and Nexo, E. (2001). **A subclass of HER1 ligands are prognostic markers for survival in bladder cancer patients.** Cancer Res 61(16): 6227-6233.

Tomas, A., Futter, C. E. and Eden, E. R. (2014). **EGF receptor trafficking: consequences for signaling and cancer.** Trends in cell biology 24(1): 26-34.

Tonks, N. K. (2006). **Protein tyrosine phosphatases: from genes, to function, to disease.** Nat Rev Mol Cell Biol 7(11): 833-846.

Tonks, N. K., Diltz, C. D. and Fischer, E. H. (1988). **Characterization of the major protein-tyrosine-phosphatases of human placenta.** J Biol Chem 263(14): 6731-6737.

Tonks, N. K., Diltz, C. D. and Fischer, E. H. (1988). **Purification of the major protein-tyrosine-phosphatases of human placenta.** J Biol Chem 263(14): 6722-6730.

Torgersen, M. L., Skretting, G., van Deurs, B. and Sandvig, K. (2001). **Internalization of cholera toxin by different endocytic mechanisms.** J Cell Sci 114(Pt 20): 3737-3747.

Traub, L. M. (2009). **Tickets to ride: selecting cargo for clathrin-regulated internalization.** Nat Rev Mol Cell Biol 10(9): 583-596.

Trusolino, L. and Comoglio, P. M. (2002). **Scatter-factor and semaphorin receptors: cell signalling for invasive growth.** Nat Rev Cancer 2(4): 289-300.

Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J. and et al. (1984). **Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells.** Nature 309(5967): 418-425.

Ullrich, A. and Schlessinger, J. (1990). **Signal transduction by receptors with tyrosine kinase activity.** Cell 61(2): 203-212.

Umebayashi, K., Stenmark, H. and Yoshimori, T. (2008). **Ubc4/5 and c-Cbl continue to ubiquitinate EGF receptor after internalization to facilitate polyubiquitination and degradation.** Mol Biol Cell 19(8): 3454-3462.

Velu, T. J. (1990). **Structure, function and transforming potential of the epidermal growth factor receptor.** Mol Cell Endocrinol 70(3): 205-216.

Villasenor, R., Nonaka, H., Del Conte-Zerial, P., Kalaidzidis, Y. and Zerial, M. (2015). **Regulation of EGFR signal transduction by analogue-to-digital conversion in endosomes.** eLife 4.

Volinsky, N. and Kholodenko, B. N. (2013). **Complexity of receptor tyrosine kinase signal processing.** Cold Spring Harbor perspectives in biology 5(8): a009043.

Wakeling, A. E., Guy, S. P., Woodburn, J. R., Ashton, S. E., Curry, B. J., Barker, A. J. and Gibson, K. H. (2002). **ZD1839 (Iressa): an orally active inhibitor of epidermal growth factor signaling with potential for cancer therapy.** Cancer Res 62(20): 5749-5754.

Waterman, H., Katz, M., Rubin, C., Shtiegman, K., Lavi, S., Elson, A., Jovin, T. and Yarden, Y. (2002). **A mutant EGF-receptor defective in ubiquitylation and endocytosis unveils a role for Grb2 in negative signaling.** EMBO J 21(3): 303-313.

Waterman, H., Levkowitz, G., Alroy, I. and Yarden, Y. (1999). **The RING finger of c-Cbl mediates desensitization of the epidermal growth factor receptor.** J Biol Chem 274(32): 22151-22154.

Wikstrand, C. J., Hale, L. P., Batra, S. K., Hill, M. L., Humphrey, P. A., Kurpad, S. N., McLendon, R. E., Moscatello, D., Pegram, C. N., Reist, C. J. and et al. (1995). **Monoclonal antibodies against EGFRvIII are tumor specific and react with breast and lung carcinomas and malignant gliomas.** Cancer Res 55(14): 3140-3148.

Wiley, H. S., Herbst, J. J., Walsh, B. J., Lauffenburger, D. A., Rosenfeld, M. G. and Gill, G. N. (1991). **The role of tyrosine kinase activity in endocytosis, compartmentation, and down-regulation of the epidermal growth factor receptor.** J Biol Chem 266(17): 11083-11094.

Wilson, T. R., Lee, D. Y., Berry, L., Shames, D. S. and Settleman, J. (2011). **Neuregulin-1-mediated autocrine signaling underlies sensitivity to HER2 kinase inhibitors in a subset of human cancers.** Cancer Cell 20(2): 158-172.

Woelk, T., Sigismund, S., Penengo, L. and Polo, S. (2007). **The ubiquitination code: a signalling problem.** Cell division 2: 11.

Wollert, T., Yang, D., Ren, X., Lee, H. H., Im, Y. J. and Hurley, J. H. (2009). **The ESCRT machinery at a glance.** J Cell Sci 122(Pt 13): 2163-2166.

Xian, C. J. and Zhou, X. F. (1999). **Roles of transforming growth factor-alpha and related molecules in the nervous system.** Mol Neurobiol 20(2-3): 157-183.

Yamamoto, T., Saito, M., Kumazawa, K., Doi, A., Matsui, A., Takebe, S., Amari, T., Oyama, M. and Semba, K. (2011). **ErbB2/HER2: Its Contribution to Basic Cancer Biology and the Development of Molecular Targeted Therapy**. Breast Cancer - Carcinogenesis, Cell Growth and Signalling Pathways.

Yang, Z., Steentoft, C., Hauge, C., Hansen, L., Thomsen, A. L., Niola, F., Vester-Christensen, M. B., Frodin, M., Clausen, H., Wandall, H. H. and Bennett, E. P. (2015). **Fast and sensitive detection of indels induced by precise gene targeting.** Nucleic Acids Res 43(9): e59.

Yang, Z. Y., Liu, L., Mao, C., Wu, X. Y., Huang, Y. F., Hu, X. F. and Tang, J. L. (2014). **Chemotherapy with cetuximab versus chemotherapy alone for chemotherapy-naive advanced non-small cell lung cancer.** Cochrane Database Syst Rev 11: CD009948.

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

Yarden, Y. and Pines, G. (2012). **The ERBB network: at last, cancer therapy meets systems biology.** Nat Rev Cancer 12(8): 553-563.

Yarden, Y. and Sliwkowski, M. X. (2001). **Untangling the ErbB signalling network.** Nat Rev Mol Cell Biol 2(2): 127-137.

Yeatman, T. J. (2004). **A renaissance for SRC.** Nat Rev Cancer 4(6): 470-480.

Yokouchi, M., Kondo, T., Houghton, A., Bartkiewicz, M., Horne, W. C., Zhang, H., Yoshimura, A. and Baron, R. (1999). **Ligand-induced ubiquitination of the epidermal growth factor receptor involves the interaction of the c-Cbl RING finger and UbcH7.** The Journal of biological chemistry 274(44): 31707-31712.

Yonesaka, K., Zejnullahu, K., Okamoto, I., Satoh, T., Cappuzzo, F., Souglakos, J., Ercan, D., Rogers, A., Roncalli, M., Takeda, M., Fujisaka, Y., Philips, J., Shimizu, T., Maenishi, O., Cho, Y., Sun, J., Destro, A., Taira, K., Takeda, K., Okabe, T., Swanson, J., Itoh, H., Takada, M., Lifshits, E., Okuno, K., Engelman, J. A., Shivdasani, R. A., Nishio, K., Fukuoka, M., Varella-Garcia, M., Nakagawa, K. and Janne, P. A. (2011). **Activation of ERBB2 signaling causes resistance to the EGFR-directed therapeutic antibody cetuximab.** Sci Transl Med 3(99): 99ra86.

Yudushkin, I. A., Schleifenbaum, A., Kinkhabwala, A., Neel, B. G., Schultz, C. and Bastiaens, P. I. (2007). **Live-cell imaging of enzyme-substrate interaction reveals spatial regulation of PTP1B.** Science 315(5808): 115-119.

Yun, C. H., Boggon, T. J., Li, Y., Woo, M. S., Greulich, H., Meyerson, M. and Eck, M. J. (2007). **Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity.** Cancer Cell 11(3): 217-227.

Zandi, R., Larsen, A. B., Andersen, P., Stockhausen, M. T. and Poulsen, H. S. (2007). **Mechanisms for oncogenic activation of the epidermal growth factor receptor.** Cellular signalling 19(10): 2013-2023.

Zhang, X., Gureasko, J., Shen, K., Cole, P. A. and Kuriyan, J. (2006). **An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor.** Cell 125(6): 1137-1149.

Zhang, Z., Stiegler, A. L., Boggon, T. J., Kobayashi, S. and Halmos, B. (2010). **EGFR-mutated lung cancer: a paradigm of molecular oncology.** Oncotarget 1(7): 497-514.

Zimmermann, M., Zouhair, A., Azria, D. and Ozsahin, M. (2006). **The epidermal growth factor receptor (EGFR) in head and neck cancer: its role and treatment implications.** Radiat Oncol 1: 11.

Zwick, E., Bange, J. and Ullrich, A. (2001). **Receptor tyrosine kinase signalling as a target for cancer intervention strategies.** Endocr Relat Cancer 8(3): 161-173.