Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

M. Sc. (Engg.)- Mr. Gajendra Suryavanshi

Born in: Parbhani, Maharashtra, India

Oral-examination:. . . . . .

Mathematical model to uncover the role of receptor ubiquitination in dose dependent EGFR trafficking

Referees: Prof. Dr. Roland Eils

Prof. Dr. Lars Kaderali

Dedicated to

My Parents

&

Sisters

Acknowledgement

I wish to express my deepest gratitude to my advisors, ***Prof. Dr. Roland Eils*** *and* ***Prof. Dr. Lars Kaderali*** *for their invaluable guidance, constant encouragement, priceless suggestions, support and sharing their rich experience during the entire course. I am obliged to them for the confidence they have shown in me and for the patience they have exercised during entire course of work.*

I also wish to express my gratitude to ***Prof. Dr. Ursula Kummer*** *for her invaluable guidance, constant encouragement and indispensable support. I am thankful to* ***Prof. Dr. Angel Alonso*** *for his suggestions and sharing his rich experience during the entire course. I am very thankful to Dr. Narsis Kiani for thoughtful discussions and suggestions.*

I am truly thankful to ***HBIGS*** *for financial support and encouragement during my work. I would also like to thank the* ***Graduate Academy*** *of Heidelberg University for providing me the thesis completion grant. Without these my work would not have come to the present state*

My warmest regards to all ***BioQuant*** *faculty members for their direct and indirect help during my work. I am also thankful to* ***BioQuant*** *IT and office staff.*

I am thankful to all my lab mates and my friends for extending their help in all possible ways. Special thanks to Dr. Diana Claußnitzer for helping me with German translation.

Finally, I am thankful to my sisters Varsha and Megha for their love and long continuing support extended to me. I am grateful to my parents for bringing me to this stage of life. I owe them a lot.

Suryavanshi Gajendra W.

Abstract

Precise regulation of epidermal growth factor (EGF) receptor (EGFR) activated signaling pathways is essential in cell fate decisions. Recent experiments observe change in EGFR internalization route from clathrin mediated (CME) to clathrin independent endocytosis (CIE) with rising EGF concentration, which alters the receptor localization and modulates associated signaling. The regulatory mechanism controlling the shift in endocytosis remains unknown. In this study, we present a novel mathematical model that describes the dose dependent regulation of EGFR trafficking through receptor ubiquitination. We assume that the receptor-ubiquitin binding reaction follows a sigmoidal behavior as function of EGF dose, which is responsible for a switch-like activation of CIE. Using the model we illustrate the change in the EGFR localization as function of EGF dose and route of endocytosis. The model is further utilized to explore the effect of defective ubiquitination on the EGFR trafficking. At high EGF concentrations, model results quantitatively capture experimentally observed changes in receptor localization caused by selective inhibition of CME or CIE. These results elucidate the ubiquitin guided sorting of EGFR during internalization. In agreement with experiments for low EGF dose, our model predicts that CIE remains largely inactive causing prolonged EGFR transport and decreased ligand degradation and strengthens our assumption of ultrasensitive receptor-ubiquitin binding. Our model accurately captures the experimentally observed deregulation in EGFR trafficking resulting from mutation induced defective ubiquitination and demonstrates the importance of receptor ubiquitination. The predictions obtained clearly indicate that our model successfully captures the underlying dynamics of ubiquitin regulated EGFR sorting and trafficking and provides valuable insights into the experimental observations. The model may thus provide a framework to study the dose-dependent attenuation of EGFR activated signaling pathways.

Zusammenfassung

Die präzise Regulierung der Signalwege, die vom epidermalen Wachstumsfaktor (EGF)-Rezeptor (EGFR) angeregt werden, ist von wesentlicher Bedeutung für Entscheidungen über das Zellschicksal. In jüngsten Experimenten wurde beobachtet, dass die Internalisierung von EGFR mit steigender EGF-Konzentration von der Clathrin-vermittelten (CME) zu Clathrin-unabhängiger Endozytose (CIE) wechselt, was mit Änderungen der Rezeptorlokalisation einhergeht und die damit verbundenen Signale moduliert. Der Regulationsmechanismus, welcher diese Verschiebung der Endozytose steuert, ist bisher nicht bekannt. In dieser Studie präsentieren wir ein neues mathematisches Modell, das die dosisabhängige Regulierung des EGFR-Transports durch Rezeptor-Ubiquitinierung beschreibt. Wir nehmen an, dass die Rezeptor-Ubiquitin-Bindungsreaktion einer sigmoidalen Funktion der EGF-Dosis folgt, welche als Schalter für die Aktivierung der CIE wirkt. Mithilfe dieses Modells zeigen wir die Änderung in der EGFR-Lokalisation als Funktion der EGF-Dosis und der Art der Endozytose. Weiterhin benutzen wir das Modell, um die Wirkung gestörter Ubiquitinierung auf den EGFR-Transport zu untersuchen. Bei hohen EGF-Konzentrationen, zeigen die Modellergebnisse quantitative Übereinstimmung mit experimentell beobachteten Veränderungen in Rezeptorlokalisation, die durch selektive Hemmung von CME oder CIE verursacht wird, und unterstreichen damit die Ubiquitin-gesteuerte Sortierung des EGFR. In Übereinstimmung mit den Experimenten für niedrige EGF-Dosis, sagt unser Modell voraus, dass CIE weitgehend inaktiv bleibt und damit verlängerten EGFR-Transport und verringerten Ligandenabbau verursacht. Dies unterstützt unsere Annahme der ultrasensitiven Rezeptor-Ubiquitin-Bindung. Unser Modell erfasst die experimentell beobachtete Deregulierung des EGFR-Transports durch mutationsinduzierte, defekte Ubiquitinierung präzise, und zeigt damit die Bedeutung der Rezeptor-Ubiquitinierung. Die gewonnenen Vorhersagen zeigen deutlich, dass unser Modell in der Lage ist, die zugrunde liegende Dynamik der Ubiquitin-regulierten Sortierung und Transport des

EGFR zu erfassen, und liefern eine wertvolle Interpretation der experimentellen Beobachtungen. Das Modell kann somit einen Rahmen für die Untersuchung der dosisabhängigen Dämpfung der EGFR-aktivierten Signalwege bieten.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Cells interact with extracellular stimulus through the membrane bound receptors such as receptor tyrosine kinases (RTKs). These receptors upon binding with the ligand turn active and initiate cascade of signaling pathways. Through the signaling pathways information is transferred and encoded to generate stimuli specific responses. Post binding receptors are internalized and transported across the intracellular compartments. From the inside of a cell, receptors activate a different set of cellular signaling leading to different outcome. These features of the RTK activation and transport are fundamental to cellular response and adaptation. A class of RTKs belonging to the family of ErbB plays pivotal role in cell cycle. ErbB family consists of four receptors namely ErbB1, ErbB2, ErbB3, and ErbB4. Upon binding with specific ligand, receptor sends an input signal which is processed downstream to produce distinct outputs ranging from cell division and migration to adhesion, differentiation and apoptosis (Figure: 1.1). Hence, function of ErbB activated signaling is extremely important in human cancers and other cell cycle related diseases.

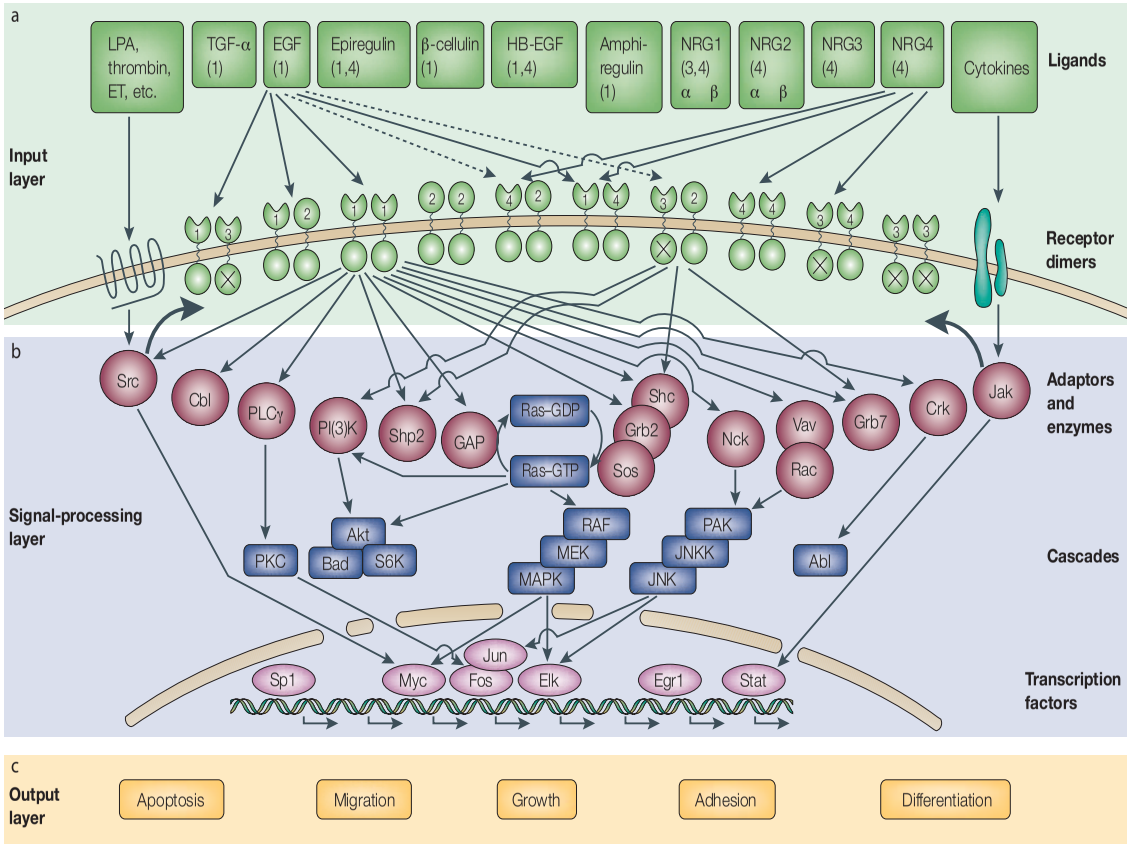


Figure 1.1: Signaling pathways activated by different receptors of ErbB family. Numbers represent the corresponding ErbB receptor (Yarden and Sliwkowski, 2001).

The epidermal growth factor receptor (EGFR) orchestrates a cascade of intracellular signaling crucial in cell growth, survival, proliferation and cell differentiation (Oda *et al*., 2005). It is responsible for the activation of signaling pathways at cell surface as well as inside of the cell (Sorkin and von Zastrow, 2009). At the cell surface, EGFR works as a sensor of the extracellular stimuli and across the intracellular compartments as a carrier of stimulating ligand. Various growth factors bind to the EGFR and form a receptor-ligand complex triggering a series of signaling events at the cell membrane. These complexes then internalize through different modes of endocytosis (Mayor and Pagano, 2007; Orth *et al*., 2006; Sorkin and Goh, 2008). Inside a cell, specific signaling pathways are activated and deactivated depending on the localization of receptors (Murphy *et al*., 2009). Therefore, accurate intracellular positioning of the receptors is essential in modulating cellular signaling. Any error in localization of EGFRs and/or ligands leads to the dysfunction of cellular signaling and may bring a cell on the verge of cancer (Yarden and Sliwkowski, 2001). Cells regulate EGF induced signaling in a dose dependent fashion through systematic regulation of the EGFR trafficking. Yet, the mechanism governing trafficking and localization of EGFR remains unknown, leaving a gap in the understanding of dose-dependent regulation in EGF induced signaling.

Post EGF binding, EGFR undergoes further modification through binding of single or multiple ubiquitin molecules (Dikic *et al*., 2009). Free ubiquitin is attached to its substrates through an E1-E2-E3 multi-enzyme cascade (Dikic *et al*., 2009; Weissman, 2001; Ye and Rape, 2009). In case of EGFR, Cbl ligase (an E3 enzyme) attaches itself to the EGF-EGFR complex and guides the binding of one or several ubiquitin molecules. This attachment of ubiquitin can affect localization, activity, and interaction partners of the ubiquitinated receptors (Woelk *et al*., 2007). Various experiments show that ubiquitination plays an important role in en-

dosomal sorting as it labels ubiquitinated receptors for lysosomal degradation thus reducing recycling of the receptors (Haglund *et al*., 2003; Weissman, 2001). Hence, ubiquitination works as a negative regulator of signaling (Dikic, 2003; Duan *et al*., 2003). Studies show receptor modification by ubiquitin is a signal for initiation of endocytosis. Experiments also show plausible internalization of receptors without ubiquitin binding (Huang *et al*., 2007). Another experimental study suggests a link between external EGF concentration and the extent of EGFR ubiquitination (Sigismund *et al*., 2005). In the study, data show negligible EGFR ubiquitination for low EGF dose, however, amount of EGFR ubiquitination is significant at high EGF concentration. This behavior is seen as a sigmoidal response in receptor-ubiquitin binding as a function of EGF dose (Acconcia *et al*., 2009).

Internalization of receptor from cell surface is a key step that regulates entire EGFR trafficking. Through endocytic routes, ligand bound EGFR travels from the cell membrane to cytoplasm. In clathrin-mediated endocytosis (CME), ligand receptor binding initiates formation of clathrin-coated pits that collect ligand-receptor complexes from the cell membrane. The coated pits detach from the membrane as clathrin-coated vesicles and transport the ligand-receptor complexes to early endosomes. However, many studies provide evidence that in parallel with CME, receptors can also internalize through clathrin independent mode of endocytosis (Mayor and Pagano, 2007; Orth *et al*., 2006). The clathrin independent route of endocytosis exploits cholesterol-rich membrane domains (i.e., rafts and caveolin) (Aguilar and Wendland, 2005; Mayor and Pagano, 2007). Considering the limitations in understanding the molecular nature of clathrin independent endocytosis (CIE), we define the process as a mode of endocytosis alternative to the CME, which is uninfluenced by the ablation of clathrin. In addition, drugs that interact with membrane cholesterol reducing membrane flexibility such as Filipin can impair CI internalization.

Sigismund *et al*. (2008) show that EGFR uses CME or both CME and CIE routes depending on the extracellular concentration of EGF. At low concentration of EGF, Sigismund *et al*. observe only CME to be active whereas at high concentration of EGF both CME and CIE modes are active. Interestingly, the same study shows that receptor and ligand undergo different fate based on their mode of endocytosis. Experiments show receptors entering through CME are more favored for recycling and less for degradation; conversely, those entering through CIE are more prone to degradation and less to recycling. These findings provide evidence that EGFR trafficking is regulated in response to a changing external ligand concentration.

From the experimental observations, it can be interpreted that a majority of ubiquitin-free EGFRs use CME while majority of ubiquitin-bound EGFRs enter through CIE mode. Hence, receptor ubiquitination functions as a regulatory step controlling the EGFR internalization routes in EGF dose-dependent manner. We infer this as a new hypothesis for an EGF dose-dependent regulation of EGFR trafficking. Wherein, at low EGF concentration due to insignificant ubiquitination of EGFR, CIE route remains largely inactive and entry through CME is dominant. This leads to low degradation and high recycling EGFR providing a sustained signaling. However, at high EGF level, elevated ubiquitination of EGFR initiates clathrin independent (CI) internalization. This significantly reduces EGFR count due to an overall increase in degradation of EGFR concluding in the termination of trafficking and associated signaling.

Several existing computational studies help to unravel different aspects of the dynamics of ligand induced EGFR trafficking and consequent signaling. These studies levy emphasis on modeling the ligand-receptor interaction and internalization (Starbuck and Lauffenburger, 1992; Waters *et al*., 1990; Wiley and Cunningham, 1981). In order to study the EGFR intracellular trafficking, early experimental

studies used high EGF concentrations due to technical limitations in data acquisition. Hence, models developed to explain these experimental data are more suitable for the condition of high EGF. Additional experimental data gathered from intracellular compartments and the extension of initial models provide a quantitative description of the intracellular EGFR trafficking (French *et al*., 1994, 1995). Further advances in the experimental techniques provided details about the EGFR associated signaling. Recent models developed using the principles of chemical kinetics are able to explain the cellular signaling that arises post ligand-receptor binding (Kholodenko *et al*., 1999; Resat *et al*., 2003). These theoretical studies provide valuable insights into the nature of signaling and its complexity.

Most of the published models assume a single mode of endocytosis in describing the receptor internalization. This restricts the possibility of studying the influence of different modes of endocytosis on the EGFR trafficking and signaling. In these studies, ubiquitin-bound or ubiquitin-free receptor populations are indistinct. Lund *et al*. (1990) propose that ligand-bound receptors internalize through clathrin coated pits whereas empty receptors use non-coated pits (smooth pits). The model by Lund *et al*., excludes the clathrin independent internalization of ligand-bound EGFR and receptor ubiquitination. Thus, the role of CIE and ubiquitination remains unclear. A recently published computational study models a signaling network activated simultaneously by members of ErbB receptor family (Chen *et al*., 2009). Their study encompasses key components of the signaling that become active in immediate response to the receptor phosphorylation. Chen *et al*. use a set of 499 ordinary differential equations to study time evolution of 471 species of the huge network. Though their model provides details about the dynamic behavior of the signaling cascades with respect to time and EGF concentration, it excludes the components required for ubiquitination and CIE mode. Thus, effect of recep-

tor ubiquitination and CIE mode on the receptor trafficking remains unexplored. Another modeling study proposes that at high EGF dose, saturation of CME route activates the CIE generating an ultrasensitive surface sorting of EGFR (Schmidt-Glenewinkel *et al*., 2008). The assumptions made by the authors appear inconsistent with recent experimental observations this leaves their proposed mechanism unconfirmed. Although various experimental studies and existing mathematical models provide a great deal of information about EGFR trafficking, the regulation of EGFR transport by means of ubiquitination and clathrin independent internalization, a cause and effect dynamics, however remains poorly understood. Insights into EGFR trafficking and its regulation are of foremost importance in order to analyze the dose-dependent attenuation of signaling pathways.

In this report, we present a mathematical model to study the role of ubiquitination and CIE mode in EGFR trafficking. To elucidate the effect of ubiquitination, our model considers two sub-populations of ligand-receptor complexes: ubiquitin-free and ubiquitin-bound complexes. The model also accounts for existence of CME and CIE as two different modes of endocytosis. We fit the model with experimental data to validate it and identify model parameters. We study the impact of CME/CIE inhibition and influence of EGF dosage on the EGFR trafficking. Further, we utilize the model to predict the effect of defective receptor ubiquitination on the dynamics of EGFR transport. The information provided by our model is extremely useful in understanding the modulation of cellular signaling with varying ligand concentration vis-á -vis cellular adaptation in a changing environment.

Chapter 2

Model development

2.1 Model formulation

Scope of our model is to study the influence of receptor ubiquitination, in response to the changing environment, on the localization of EGFR at the cell surface and inside endosome. Using the model, we aim to explore the implications of defects in regulation of EGFR trafficking introduced by selectively disrupting internalization routes. To construct a basic framework for the model we use information from various modeling and experimental studies. We write model equations assuming mass action kinetics for reactions and for the transport and internalization are first order processes. Figure 2.1 depicts the proposed EGF stimuli dependent receptor ubiquitination, EGFR internalization, and intracellular transport. Figure 2.2 presents the flow diagram of the EGFR trafficking events. At the cell surface, a reversible binding between EGF (L_m) and its receptor (R_s) forms a ligand-receptor complex (C_s). In line with the existing models, we assume activation of receptor to be spontaneous and ligand-receptor complex to be active on ligand binding. Unlike other studies in our model, for ligand-receptor complex we propose two modes of endo-

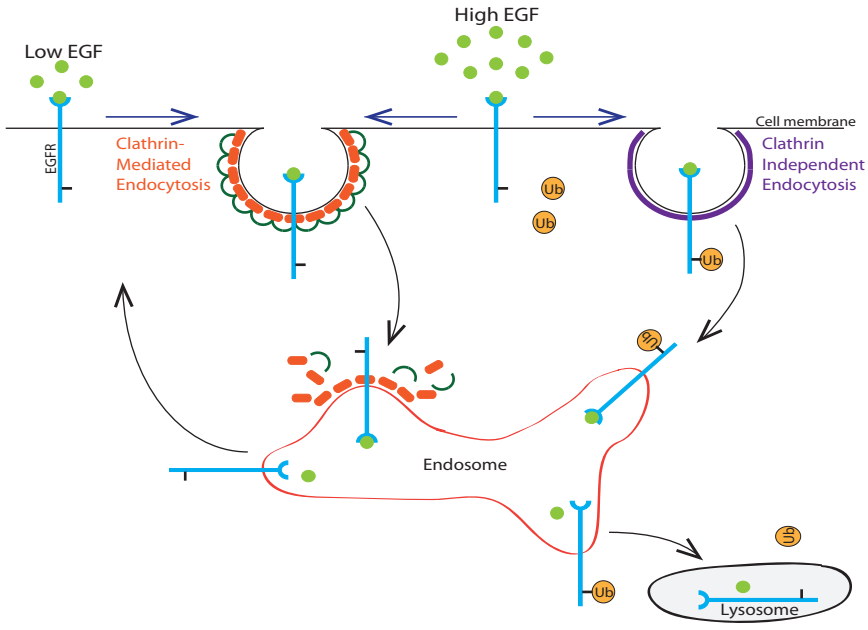


Figure 2.1: Pictorial representation of proposed EGFR surface sorting and trafficking.

cytosis: clathrin dependent (CME) and clathrin independent (CIE). We assume that ubiquitin-free complexes enter through CME route. Considering the recent experimental observations (Sigismund *et al*., 2005), we propose that extent of receptor ubiquitination is dependent on the amplitude of EGF stimuli. Literature suggests that receptor ubiquitination displays sigmoidal behavior in response to the level EGF dose (Acconcia *et al*., 2009). However, the exact mechanism responsible for a switch like response remains unclear. To overcome this limitation we assume rate of receptor ubiquitination to be a function of EGF dose. We relate ubiquitination rate constant and level of EGF stimuli using Hill kinetics to reproduce the proposed ultrasensitivity in extent of ubiquitin-complex binding. An approximate value for the Hill coefficient (n_H) can be obtained from the EGF concentration values used in experiments (as shown in the following section). In the model, we consider that formation of ubiquitin bound complex (C_u) activates the CIE route and these ubiq-

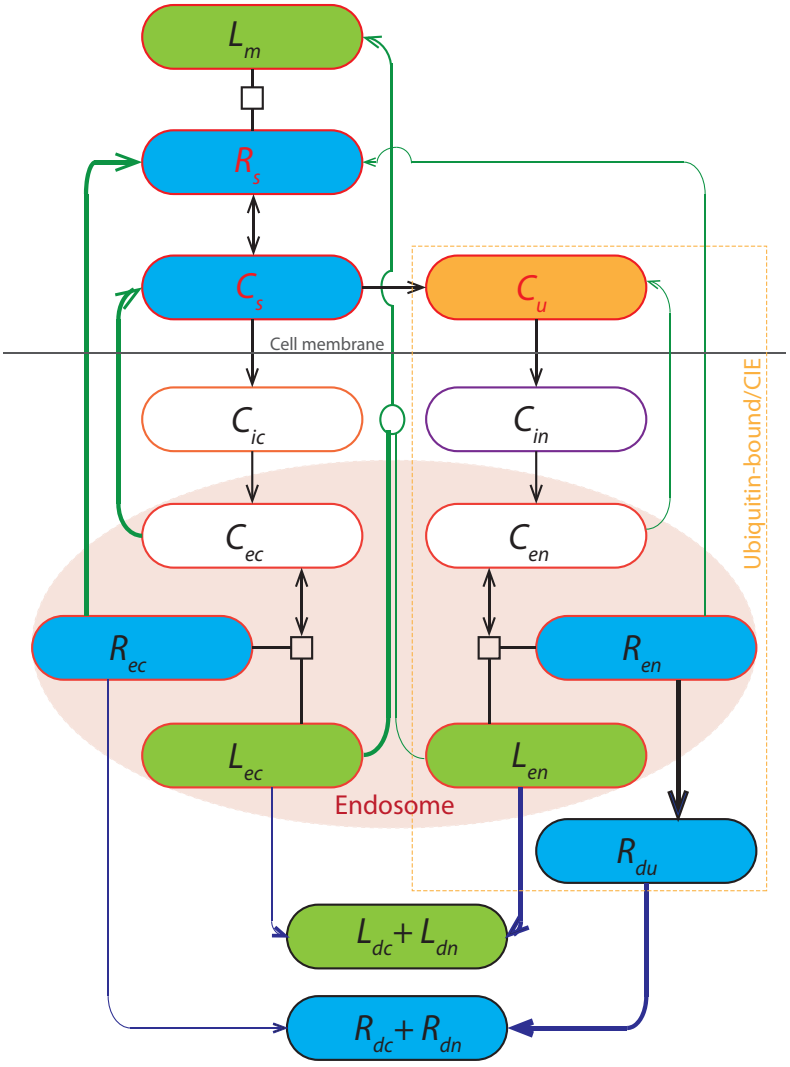


Figure 2.2: Flowchart depicting the proposed reactions and transport events during the EGFR trafficking. Letters in the boxes are variable names as used in the model equations (Table 2.1). Fraction of a population that undergoes recycling (green line) and degradation (blue line) corresponds to thickness of the line.

uitinated receptors-complexes use the CIE mode for internalization.

Using the respective route of endocytosis, both ubiquitin-free complexes (C_{ic}) and ubiquitin-bound complexes (C_{in}) reach the endosomal compartment. Inside the endosome ligand-receptor complex (C_{ec} and C_{en}) is recycled or dissociates giving ligand free receptor and free ligand. The ligand free receptors and free ligand are sorted for recycling and degradation (R_{dc}, R_{dn}, L_{dc}, & L_{dn}). It is evident from various experimental studies that ubiquitin-bound receptors are targeted for lysosomal degradation (Duan *et al*., 2003; Peschard and Park, 2003). Hence, ubiquitin-bound complexes internalizing through CIE mode are assumed to be more prone for degradation and only a small portion of these complexes (C_{en}), receptors (R_{en}), and their ligands (L_{en}) is recycled. Before the degradation of these receptors, ubiquitin molecules dissociate and recycle back to the cell membrane. Experimental evidence suggest that ubiquitin defective receptors are less likely to degrade and show an enhanced recycling (Peschard and Park, 2003). Thus, we suspect that a major faction of ubiquitin-free complexes (C_{ec}), receptors (R_{ec}) and their ligands (L_{ec}) to be recycled and reminder is degraded.

2.1.1 Model equations

We write model equations assuming mass action kinetics for reactions and for the transport and internalization are first order processes. We assume amount ubiquitin molecules to be present in abundance compared to ligand-receptors and the amount of ubiquitin remains constant due to recycling. Thus, we consider receptor ubiquitination to be first order reaction. We propose that value of rate constant for ubiquitin binding reaction to be a function of extracellular ligand concentration.

Model equations representing reactions at the cell surface:

$$V_x N_{av} \frac{dL_m}{dt} = -k_b L_m R_s + k_{ub} C_s + k_{lc} L_{ec} + k_{ln} L_{en} \tag{2.1}$$

$$\frac{dR_s}{dt} = -k_b L_m R_s + k_{ub} C_s + k_{rc} R_{ec} + k_{rn} R_{en} \tag{2.2}$$

$$\frac{dC_s}{dt} = k_b L_m R_s - k_{ub} C_s - k_{ec} C_s - k_u \left(\frac{L_m^{n_H}}{1 + L_m^{n_H}} \right) C_s + k_{rc} C_{ec} \tag{2.3}$$

$$\frac{dC_u}{dt} = k_u \left(\frac{L_m^{n_H}}{1 + L_m^{n_H}} \right) C_s - k_{en} C_u + k_{rn} C_{en} \tag{2.4}$$

Intracellular compartment:

$$\frac{dC_{ic}}{dt} = k_{ec} C_s - k_{ee} C_{ic} \tag{2.5}$$

$$\frac{dC_{in}}{dt} = k_{en} C_u - k_{ee} C_{in} \tag{2.6}$$

Endosomal compartment:

$$\frac{dC_{ec}}{dt} = k_{ee} C_{ic} - k_{eb} C_{ec} + \frac{k_{eub} R_{ec} L_{ec}}{N_{av} V_e} - k_{rc} C_{ec} \tag{2.7}$$

$$\frac{dC_{en}}{dt} = k_{ee} C_{in} - k_{eb} C_{en} + \frac{k_{eub} R_{en} L_{en}}{N_{av} V_e} - k_{rn} C_{en} \tag{2.8}$$

$$\frac{dR_{ec}}{dt} = k_{eb} C_{ec} - \frac{k_{eub} R_{ec} L_{ec}}{N_{av} V_e} - k_{rc} R_{ec} - k_{rdc} R_{ec} \tag{2.9}$$

$$\frac{dR_{en}}{dt} = k_{eb} C_{en} - \frac{k_{eub} R_{en} L_{en}}{N_{av} V_e} - k_{rn} R_{en} - k_{rdn} R_{en} \tag{2.10}$$

$$\frac{dL_{ec}}{dt} = k_{eb} C_{ec} - \frac{k_{eub} R_{ec} L_{ec}}{N_{av} V_e} - k_{lc} L_{ec} - k_{dlc} L_{ec} \tag{2.11}$$

$$\frac{dL_{en}}{dt} = k_{eb} C_{en} - \frac{k_{eub} R_{en} L_{en}}{N_{av} V_e} - k_{ln} L_{en} - k_{dln} L_{en} \tag{2.12}$$

Outside the endosomal compartment:

$$\frac{dR_{du}}{dt} = k_{drn} R_{en} - k_{du} R_{du} \tag{2.13}$$

$$\frac{dR_{dc}}{dt} = k_{drc} R_{ec} \tag{2.14}$$

$$\frac{dR_{dn}}{dt} = k_{du} R_{du} \tag{2.15}$$

$$\frac{dL_{dc}}{dt} = k_{dlc}L_{ec} \tag{2.16}$$

$$\frac{dL_{dn}}{dt} = k_{dln}L_{en} \tag{2.17}$$

Model consists of 17 variables and 17 parameters. Table 2.1 provides the list of all the variables and parameters that are used in the model. Values for the ligand-receptor binding/unbinding rate constants at cell surface and inside the endosome are taken from literature (French *et al*., 1995). The values of endosomal volume (V_e) used for the model simulation is 10^{-14} liter/cell (Tzafriri and Edelman, 2007) and value of extracellular volume (V_x) is 10^{-10} liter/cell considering the cell population range between 10^6 and 10^8 cells/ml. We used standard value of Avogadro's number (N_{av})= 6.0221415×10^{23}. The set of coupled ODEs is solved in MATLAB. A code is written to solve the model equations to mimic the experimental protocols. Model equations are solved with continuous stimulation of 100 ng/ml EGF concentration to calculate the surface and total EGFR concentration. To calculate EGF degradation and recycling the equations are solved with 20 ng/ml EGF for the pulse phase of 6 min and 15 min, respectively. In order to replicate the chase phase, we solve the equations with the external EGF concentration and the ligand-bound receptor population set to zero as initial condition.

2.1.2 Value of Hill coefficient

The Hill equation is commonly used to describe the reaction velocity though its use can be extended study the nature of stimuli-response curve. Sigmoidal behavior of a stimuli-response curve can be quantified using Hill coefficient. Higher the value of the coefficient steeper is the response curve. As common practice coefficient is estimated by fitting the equation against the stimuli-response data. The coefficient

Variable	Symbol	Rate constants	Symbol
Ligand in extracellular medium	L_m	Ligand receptor binding and unbinding rates surface	k_b,k_{ub}
Receptors at cell surface	R_s	Ubiquitin binding constant	k_u
Surface ligand-receptor complex	C_s	CME internalization rate	k_{ec}
Surface ubiquitin-bound complex	C_u	CIE internalization rate	k_{en}
Complexes internalizing through CME route	C_{ic}	Rate of transport from surface to endosome	k_{ee}
Complexes internalizing through CIE route	C_{in}	Ligand receptor binding and unbinding rates in endosome	k_{eb}, k_{eub}
CME internalized complex in endosome	C_{ec}	Recycling rate of C_{ec} and R_{ec}	k_{rc}
CIE internalized complex in endosome	C_{en}	Recycling rate of C_{en} and R_{en}	k_{rn}
Receptor dissociated from C_{ec}	R_{ec}	Recycling rate of L_{ec}	k_{lc}
Ligand dissociated from C_{ec}	L_{ec}	Recycling rate of L_{en}	k_{ln}
Receptor dissociated from C_{en}	R_{en}	Degradation rate of R_{ec}	k_{drc}
Ligand dissociated from C_{en}	L_{en}	Deubiquitination rate of R_{en}	k_{drn}
Deubiquitinated receptor	R_{du}	Degradation rate of L_{ec}	k_{dlc}
Degraded receptor CME, CIE	R_{dc}, R_{dn}	Degradation rate of L_{en}	k_{dln}
Degraded ligand CME, CIE	L_{dc}, L_{dn}	Degradation rate of R_{du}	k_{du}

Table 2.1: List of model variables and parameters.

can be established when data are provided for 10% and 90% response. Goldbeter and Koshland (1981) derived a correlation between the Hill coefficient and stimuli concentrations required for 10% and 90% response. The proposed correlations to calculate Hill coefficient (n_H) is shown below (see also (Huang and Ferrell, 1996))

$$n_H = \frac{\log(81)}{\log(S_{0.9}) - \log(S_{0.1})} \tag{2.18}$$

where, $S_{0.1}$ and $S_{0.9}$ are the stimuli concentrations responsible for generation 10% and 90% response, respectively.

When the stimuli concentrations are unavailable, the output response can be used instead. The above correlation then can also be written in terms of the substrate consumption. In our case, we consider the percentage of total receptor undergoing ubiquitination to be the marker of output response. The new correlation can therefore be written as follows:

$$n_H = \frac{\log(81)}{\log(0.9 \times R_T) - \log(0.1 \times R_T)} = 2 \tag{2.19}$$

where, R_T is the total EGFR count.

An alternative way to establish the coefficient is as follows.

In the experiments, Sigismund *et al*. (2008) use two concentrations of EGF ligand 1.5 ng/ml (low EGF) and 20 ng/ml (high EGF). Their previous study show for low EGF dose receptor ubiquitination is minimal and significant ubiquitination for high EGF (Sigismund *et al*., 2005). In our model to estimate the value of Hill coefficient we assume 10% ubiquitin binding response for low EGF and 90% for high EGF. On substituting the concentration values in the equation 2.18

$$n_H = \frac{\log(81)}{\log(20) - \log(1.5)} = 1.7 \cong 2 \tag{2.20}$$

Hence, in our model we set the Hill coefficient equal to 2.

2.2 Parameter estimation

Parameter estimation is a process of finding the values of the unknown parameters of governing model equations such that the model results match with given experimental measurements with minimum error. Being an inverse problem, parameter estimation does not provide a general solution (Aster *et al*., 2005). Yet, estimation of model parameter values is a crucial step in the model development as model parameters govern the dynamics of the system. Validity of model predictions is strongly connected to the correctness of parameter values. Our model consists of a set of ordinary differential equations related by parameters that are unknown and need to be either measured or estimated. For the optimization algorithm, we implement the standard approach of minimizing a cost function. The function quantifies the value of the difference between the model predicted values and the experimental observed values. Given a model and a set of experimental data, the objective of model calibration is to find the best realistic guess for unknown parameters. In our nonlinear optimization algorithm, only those parameter sets are considered valid solutions that satisfy the experimentally observed biological conditions imposed as constraints, denoted by ($\Gamma(P)$). The best-fit parameters is computed by solving the following nonlinear optimization problem

$$Min \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{o}(y_{ij}(k) - y_{ij}^{p}(k))^2 \tag{2.21}$$

subject to $\Gamma(P)$

where, $y_{ij}(k)$ is experimental observation and $y_{ij}^{p}(k)$ is the model predicted value at k^{th} time point. The index i run over the experimental condition and j over species for which measurements are available. P is set of parameters. In this work, we have two experimental conditions low EGF and high EGF. For parameter estimation we are using only the control set of data for high EGF (combined data for 100 ng/ml

and 20 ng/ml of EGF) hence $i = 1$.

We propose that in the experiments of Sigismund *et al*. the sorting of the receptors is more based on ubiquitination status of the receptor than the endocytosis route used. We consider that in the experimental study receptors population internalized through the CIE route is ubiquitin-bound and receptor population that used CME route is ubiquitin-free. Thus, the experimental data regarding the CIE will correspond to the combined intracellular population of ubiquitin-bound complexes (C_{in} and C_{en}) and dissociation products of these complexes (R_{en}, R_{du} and L_{en}). Similarly, for the CME measurements correlate to collective intracellular population of ubiquitin-free complexes (C_{ic} and C_{ec}) and receptors (R_{ec}) and ligands (L_{ec}) dissociated from these complexes. The analogies allow us to assign experimentally measured values to the corresponding model variables. With this information, we fit the model against the experimental data to obtain the rate constants such as recycling and degradation rates for the ubiquitin-bound and ubiquitin-free complexes, receptors, and associated ligands.

From the internalization assays, Sigismund *et al*. calculated the percentage of EGFR that internalized through CME and CIE. Authors also calculated the percentage of CME/CIE internalized EGF that undergoes degradation and recycling. Using this information a set of conditions have been prepared and converted to point wise constrains, shown in the table 2.2. Model equations are solved in accordance with the experimental protocol. To calculate the total and surface EGFR, equations are solved with initial concentration of EGF (L_m) set to 100 ng/ml and surface EGFR number (R_s) to 4.5×10^5/cell. The initial value for all other variables is set to zero. For the EGF degradation and recycling model equations are simulated for $t = 6$ min and $t = 15$ min, respectively, with initial EGF concentration set to 20 ng/ml. After the this pulse phase, model variable are reinitialized for chase

	CM internalized (ubiquitin-free)	**CI internalized (ubiquitin-bound)**
EGF internalized (after 6 min of stimulation)	≥ 60	≤ 40
EGF Degradation (at the end of chase phase)	≤ 30	≥ 80
EGF recycled (at the end of chase phase)	-	≤ 20

Table 2.2: Constrains $\Gamma(P)$ placed during the model calibration.

phase with EGF concentration and surface ligand bound receptor population set to zero. The amount of internalized EGF is measured by summing up all intracellular ligand. From the total internalized EGF, percentages CME internalized and CIE internalized EGF is calculated. Similarly, amount of EGF internalized during pulse phase is used to calculate the percentage EGF recycling and degradation. In addition, to speed up the parameter search we place constrains, such as percentage of ubiquitin-free EGF recycled should be higher than the ubiquitin-bound EGF vise-versa for the EGF degradation. We place a heavy penalty on the individual parameter set that fails to maintain all the constraints.

Our system contains 13 unknown parameters. To obtain values of these parameters, we use training data comprising of 18 measurements of 4 species (Surface EGFR, Total EGFR, EGF degraded, and EGF recycled) with 3 replicates for control conditions. We used the mean values of 3 replicates to find best set of parameters. To avoid weighting problem and considering the nature of least square function which is very sensitive to outliers, during optimization we omit data point at time point 60 for EGF degradation. Local optimization algorithms like gradient-based

methods perform well when implemented to find a local minimum. In our case, degree of freedom and landscape of the objective function preclude use of local optimization algorithms. Hence, we opted for a global optimization method namely genetic algorithm (GA) (Appendix A provides a brief overview of the genetic algorithm). It is more probable that the optimal solution found using GA will lie in the proximity of a global minimum. We provide a parameter search space by setting largest possible interval around all the parameters. The interval is decided based on the values of the similar parameters used in the existing models (French *et al*., 1994). Depending on the performance of optimization algorithm in searching the minimum, the intervals are retuned to get the best possible value for parameter sets. Each optimization cycle runs up to 500 generations with population size of 100 and for every generation new population is produced using adaptive mutation function (MATLAB code is provided in the Appendix A).

2.3 Sensitivity analysis

Parametric sensitivity analysis provides information about the influence of change in the parameter value on the output of the system. Being a local property, sensitivity can be used to study the change dynamics of a system with variations in the parameter values. Hence, we use the sensitivity analysis to explore the difference in the dynamics of EGFR trafficking across different fits. We use the normalized sensitivity coefficient defined as follows:

$$s_{ij}(t) = \frac{\partial \ln(y_i(t))}{\partial \ln(p_j)} \tag{2.22}$$

where, $s_{ij}(t)$ is the sensitivity of i^{th} observable variable y_i (such as total EGFR percentage) with respect to change in the change in the j^{th} parameter in p_j. We

convert this sensitivity index into a time independent values by integrating it over observation time.

$$S_{ij} = \frac{1}{T}\int_0^T |s_{ij}(t)|dt \tag{2.23}$$

where, T is the final time point of observation and we use absolute values of sensitivity to avoid negative and positive cancellation.

We use software package OpenBio for sensitivity analysis using a code written in Jacobian, a scripting environment in OpenBio, to calculate average sensitivity. To quantify the variability in EGFR dynamics between different fits we calculate the correlation coefficient (R) for parameter sensitivities among the fits.

2.4 Identifiability analysis

Identifiability analysis deals with the problem of uniqueness of the parameter set obtained from model calibration exercise (Cobelli and DiStefano, 1980; Bellman and Åström, 1970). If a model is not uniquely identifiable, several or infinite parameter sets can generate identical results for model fitting. Hence, parameter identifiability analysis plays a critical role in deciding the accuracy of the model predictions. In general, structural or practical unidentifiability precludes the optimization process from providing a unique solution. Structural (priori) identifiability as name suggests depends on the structure of the model, available measured output and it provides an estimate of information that can be obtained from the experiments. On the other hand practical identifiability depends not only on model structure but also on experimental conditions in combination with quantity and quality of the measured data (Bellman and Åström, 1970). Structural unidentifiability can be resolved either by reducing redundant parameters of the model or by increasing the numbers of measured model variables.

We performed identifiability analysis using the software toolbox SensSB developed for MATLAB (Rodriguez-Fernandez and Banga, 2010). We use the toolbox to calculate the correlation matrix that provides the information about interdependence between the model parameters (details in the Appendix B). The experiments cannot always provide data for all the involved species. In most cases, indirect measurements such as concentration of related species or cumulative measurements for a group species are practically possible. Thus, very often it is difficult to distinguished between structural and practical identifiability with the available experimental data. The problem of limited data can be addressed by providing measurements for each variable (specie) of the model. The SensSB toolbox allows us to generate simulated data for each model variable (specie). The data are then used to determine the structural identifiability of the model. In many cases, even if a model is structurally identifiable, parameter optimization may fail to determine definitive parameter values due to practical unidentifiability. Information gathered through SensSB toolbox will be helpful in finding out whether the limitations in parameter estimation are emerging from the data or from model structure.

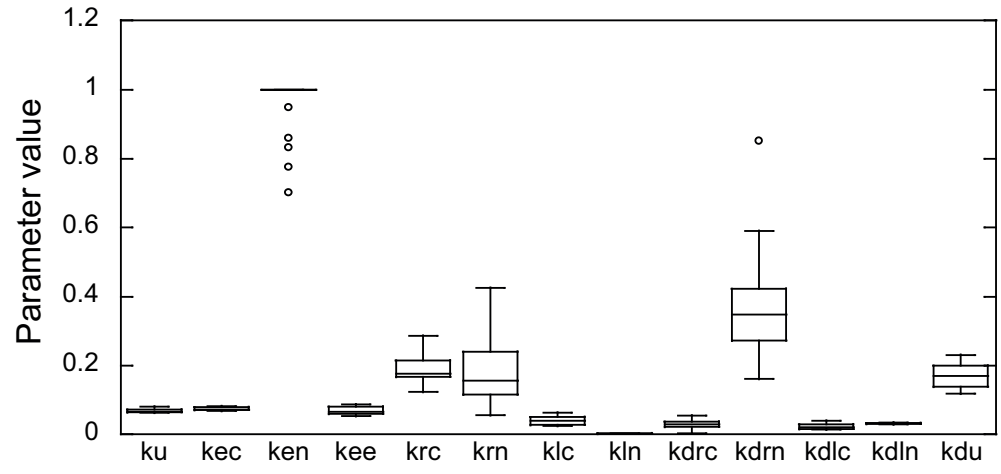
Chapter 3

Model results

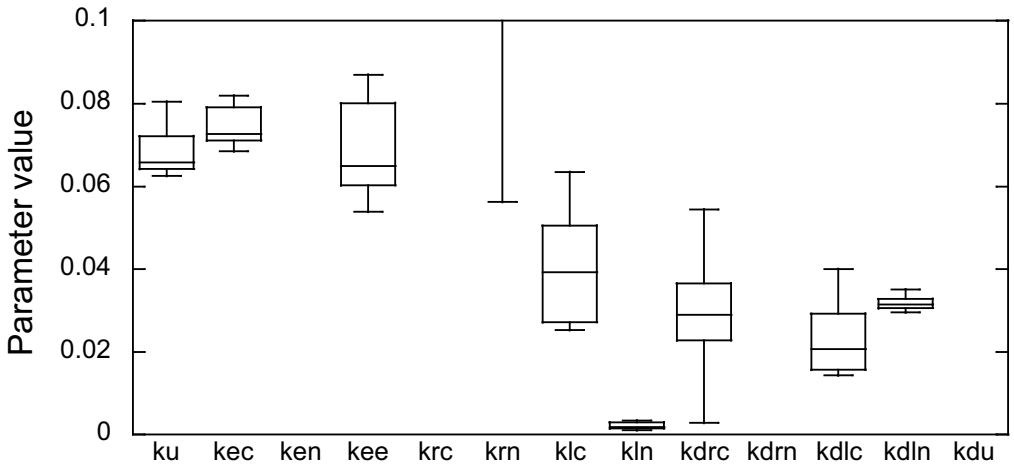
Our primary goal is to study the regulatory influence of ubiquitination on the endocytosis and transport of EGFR. In order to achieve a realistic estimate of the parameters, we fit our model against the control measurements from (Sigismund *et al*., 2008), where normal HeLa cells are exposed to high EGF dose.

3.1 Parameter estimates

Using the genetic algorithm we obtain different parameter sets that fit the experimental data equally well. Coefficient of variation for parameter value shows some parameters such as k_{ec}, k_u and k_{ln} to remain approximately constant across multiple fits, we term them well constrained. While some parameters such as k_{rn}, k_{drc} and k_{drn} to vary, termed poorly constrained. The observations are visible in the box plot (Figure 3.1(a), 3.1(b)) . The characteristics of some parameters being well constrained while some being poorly constrained, define a "sloppy" model (Brown and Sethna, 2003) and it is almost a universal property of nonlinear multi-parameter system biology models (Gutenkunst *et al*., 2007). The observed sloppiness hinders



(a) Box plot range 0 to 1



(b) Box plot range 0 to 0.1

Figure 3.1: Box plot showing parameter values and their variation.

genetic algorithm from finding the exact global minimum of cost function. The existence of different sets of parameters for which the model behavior is consistent with the data, indicates a wide flat valley in objective function domain. To check for the possibility that our algorithm is getting stuck-up in the flat valley of objective function, areas outside of valley have been explored by initiating the algorithm in

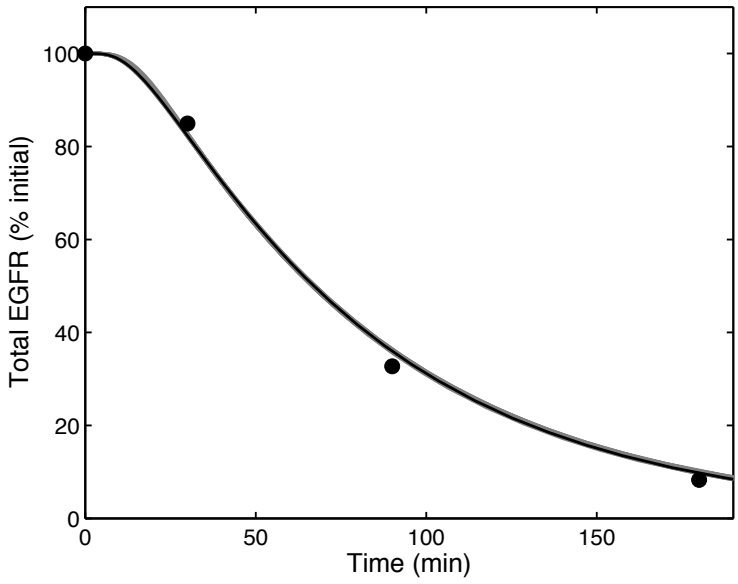
different parts of objective function. The results confirm that the best fit parameters are inside of flat valley. In a collective fit, the parameter set ensemble samples from all consistent sets of parameters. Sloppy models are very insensitive to parameter combinations that lie along sloppy directions. The parameter sets ensemble can extend very far in those directions. Since, discrepancy between different fits with residue value less than experimental error is practically not significant (Chen *et al*., 2009), we consider all fits with a root mean square deviation (RMSD) of approximately 10% or less from real measurement to be equally good. We collect 21 such parameter sets for further analysis and parameter set with lowest RMSD is considered as best-fit and table 3.1 shows the best-fit parameter values.

3.2 Model fitting

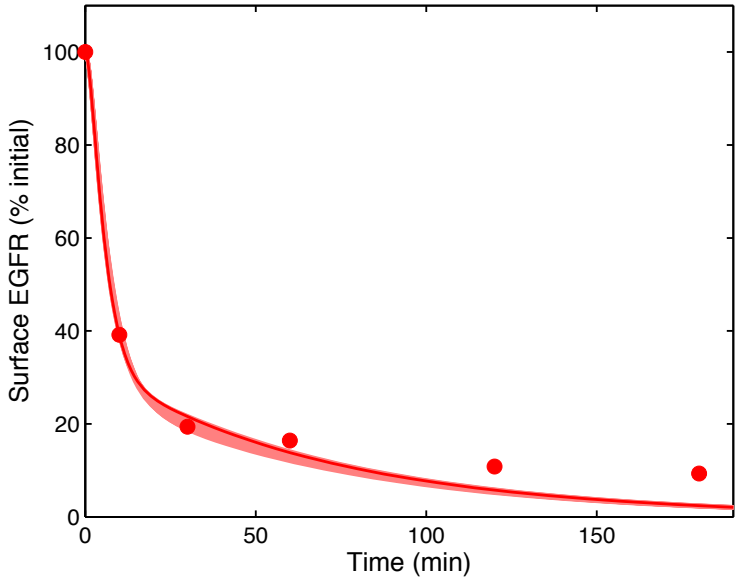
Results of the model simulations with 21 parameter sets are shown in the figures 3.2 and 3.3. Figure 3.2(a) shows simulation results and experimental data of total EGFR. In the experiments, cells were exposed to high EGF does (100 ng/ml) and total EGFR level was monitored using the Immunoblotting. The measurements shows drop in the total EGFR count with time post stimulation. The observed drop is due to the slow and steady degradation of receptors in the lysosomes. The loss of receptors with time slowly ceases trafficking and subsequently terminates the signaling. For the surface EGFR, cells were exposed to high EGF does (100 ng/ml) and surface receptor count was measured using ^{125}I-EGF. The data show that receptors internalize after the EGF binding causing ligand-bound receptors to disappear from the cell membrane. The down-regulation of EGFR lowers the surface count and also reduces the signaling originating from cell membrane. Figure 3.2(b) shows the effect of EGF stimulation on the surface EGFR.

Variable	Best fit values
k_b	0.063 nM^{-1} min^{-1} $*$
k_{ub}	0.16 min^{-1} $*$
k_u	0.069546 min^{-1}
k_{ec}	0.076521 min^{-1}
k_{en}	1 min^{-1}
k_{ee}	0.076204 min^{-1}
k_{eb}	0.0085 nM^{-1} min^{-1} $*$
k_{eub}	0.66 min^{-1} $*$
k_{rc}	0.19147 min^{-1}
k_{rn}	0.11565 min^{-1}
k_{lc}	0.026875 min^{-1}
k_{ln}	0.0030627 min^{-1}
k_{drc}	0.041646 min^{-1}
k_{drn}	0.16102 min^{-1}
k_{dlc}	0.015714 min^{-1}
k_{dln}	0.030311 min^{-1}
k_{du}	0.14123 min^{-1}

Table 3.1: List of parameter with their best-fit values used for the model simulations. $*$ value taken from literature (French *et al*., 1995).

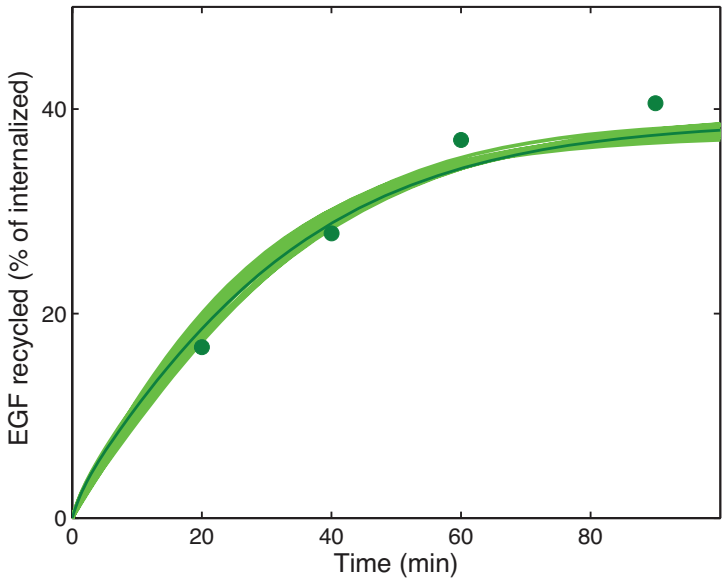


(a) Change in total EGFR in percentage

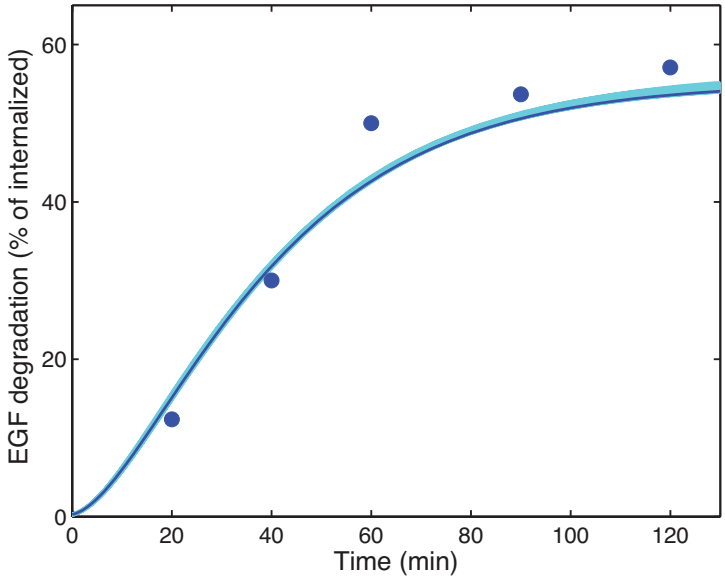


(b) Change in surface EGFR in percentage

Figure 3.2: Model simulation results for 21 equally good fits showing change in EGFR count with time post high EGF stimulation (Dark line represents best-fit results). Symbols represent experimental data (Sigismund *et al.*, 2008).



(a) Recycling of EGF



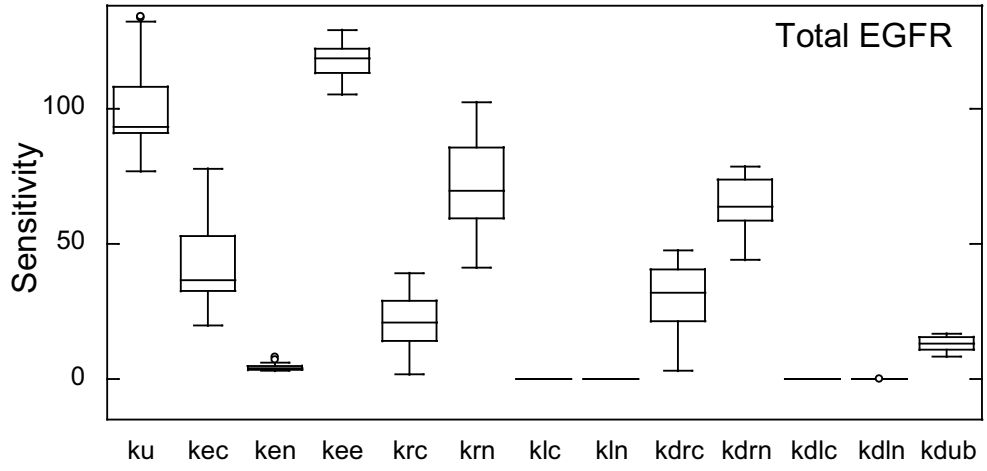
(b) Degradation of EGF

Figure 3.3: Model simulation results for 21 best fits showing EGF recycling and degradation with time post high EGF stimuli. Symbols represent experimental data (Sigismund *et al.*, 2008).

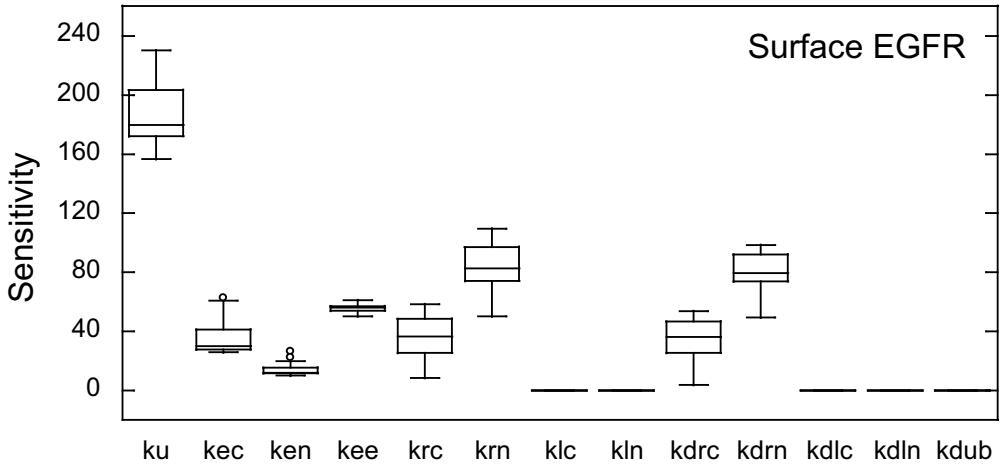
To estimate recycled and degraded EGF pulse-chase experiments were carried out. For the degradation measurements in the pulse phase HeLa cells were exposed to high ^{125}I-EGF (20 ng/ml) for 6 minute and during the chase phase cell were kept in the medium with or without excess EGF ($\sim$ 200 ng/ml) (Sigismund *et al*., 2008) data were collected for percentage of internalized ^{125}I-EGF that degraded with time. For the recycling, pulse phase was of 15 minute and during the chase phase data were collected for percentage of internalized ^{125}I-EGF that was present in the medium. Ligand degradation profile shows an increase in the percentage of degraded EGF (Figure 3.3(b)) with time and similarly there is an increase in the percentage of recycled EGF in the medium (Figure 3.3(a)).

3.3 Variation in parametric sensitivity

The sensitivity analysis of the model revels important information about the behavior of the system. The sensitivity coefficients for four species have been computed for the 21 equally good fits. The results of sensitivity analysis have been shown in figure 3.4 and 3.5. We observe variation in the sensitivity values across different fits. For some parameter the variation in the sensitivity is high and some parameters have no major contribution to the sensitivity. These variations may arise due to the change in the dynamics of the system. To investigate and measure the change in the dynamics we compared the sensitivity values from one fit with another. Comparison between the results of different acceptable fits can be used as a ruler to measure change in dynamic behavior of system due to poorly constrained parameters. The results of correlation between parametric sensitivity coefficients in different fit are shown for 5 representative fits in figure 3.6 and 3.7. The average



(a) Sensitivity box plot for total EGFR

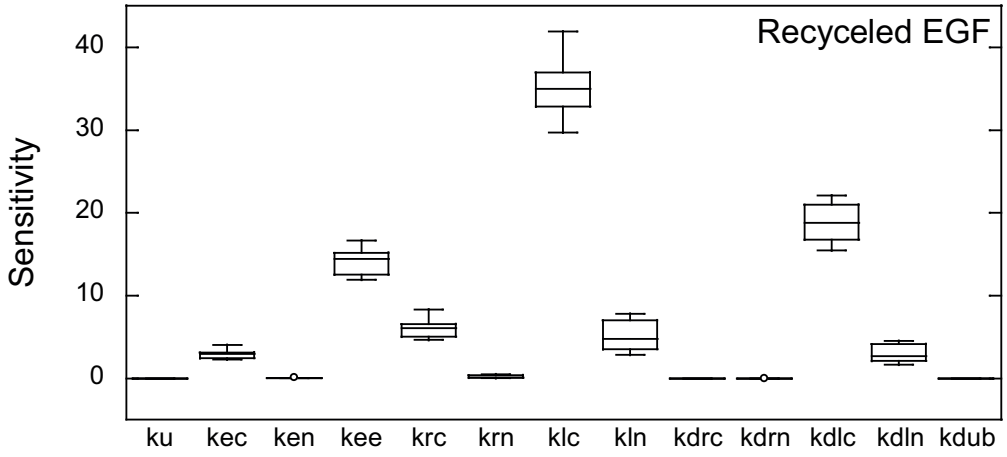


(b) Sensitivity box plot for surface EGFR

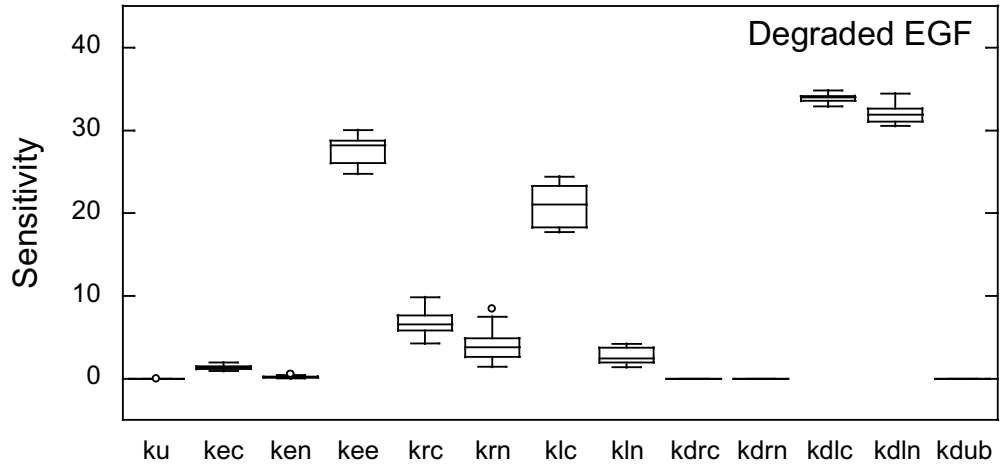
Figure 3.4: Box plot showing sensitivity variation for total and surface EGFR.

correlation, $R \approx 0.98$ confirms consistent and almost unique behavior of the system.

Although, because of sloppiness, the optimization method is not able to find the global minimum of cost function but result of sensitivity analysis for collective fits confirm that our model shows a constant dynamic behavior for all 21 parameter



(a) Sensitivity box plot for recycled EGF



(b) Sensitivity box plot for EGF degradation

Figure 3.5: Box plot showing sensitivity variation for recycled and degraded EGF.

sets. Thus, we consider that a flat valley of the global minimum lies in the parameter search space.

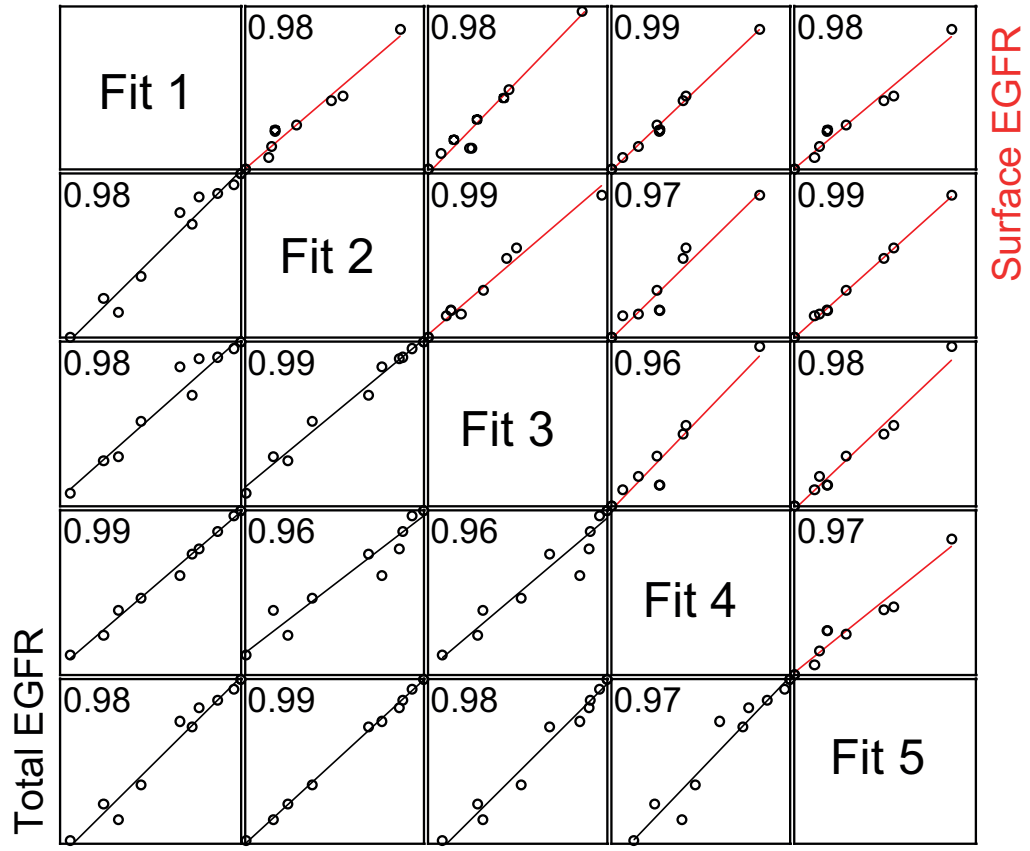


Figure 3.6: Correlation between sensitivity values of 5 representative best fits. Above diagonal plots show correlation of surface EGFR between fits and below diagonal plots show correlation of total EGFR between fits. Values in the plot represents correlation coefficient (R) between each two fits.

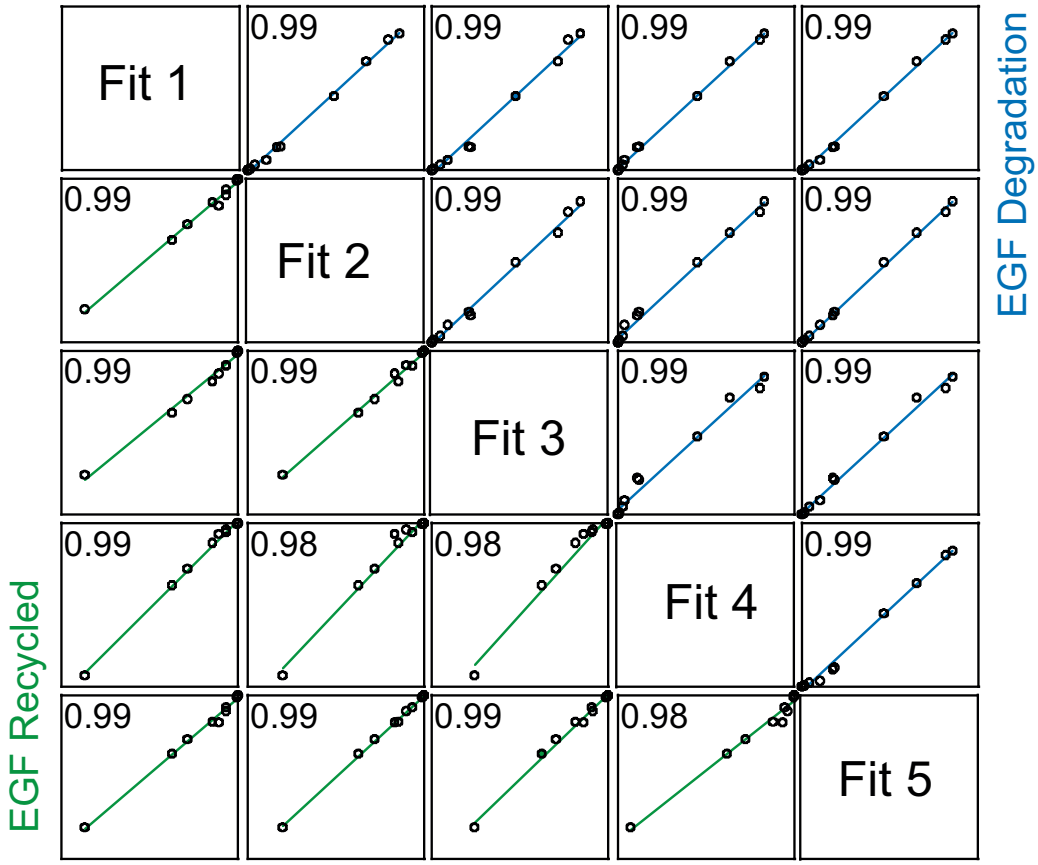


Figure 3.7: Correlation between sensitivity values of 5 representative best fits used in Figure 3.6. Correlation plots for EGF degraded (above diagonal) and EGF recycled (below diagonal). Values in the plot represents correlation coefficient (R) between each two fits.

3.4 Mean values for parameters

Simulations for the 21 fits and the sensitivity correlations indicate presence of a flat valley in the objective function. This hinders the optimization process from finding the exact solution. Considering that the 21 fits is a small sample size to accurately compute mean and standard deviation, we decided to run the optimization algorithm (GA) for large number of time in order to collect big enough sample size that will allow us to compute mean and confidence intervals. With this information we can conclusively propose the most identifiable and least identifiable parameters. In order to achieve this we performed the GA for about 4000 times and with $\approx 25\%$ success rate we gathered 1022 parameter sets with final RMSD score of less than 10%. The distribution over the obtained parameters remains unknown. Hence, we calculate the confidence interval on the mean parameter values (Table 3.2) using Chebyshev's inequality, which states that at least $100(1 - \frac{1}{k^2})\%$ of data lie around the mean μ in the interval $\pm k\sigma$ where σ is standard deviation. Table shows the parameter mean value and 95% confidence interval (values of k is 4.47). Here after we use the mean parameter values for the further model simulations and analysis. Using the mean parameter values we simulated the model to study the wellness of fit. We observe that with the mean values model is able to capture the data well. Figure 3.8 and 3.9 show the model simulation results with mean parameter values.

3.5 Sensitivity analysis with mean values

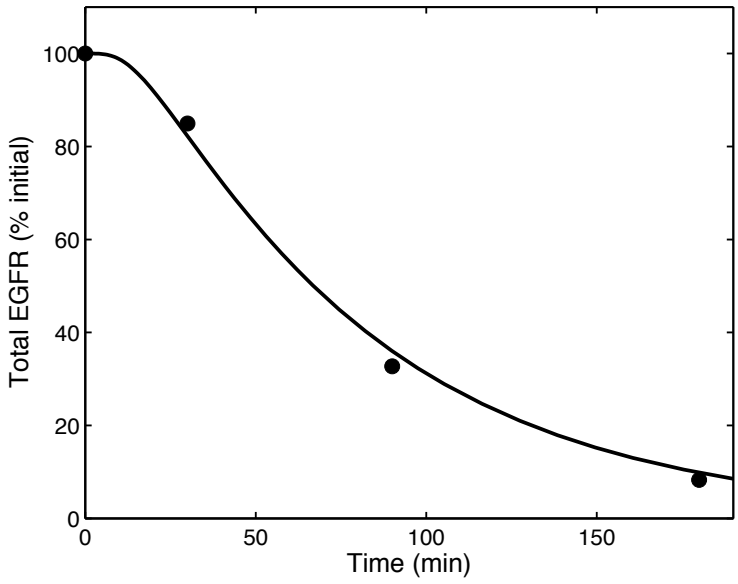
Using the mean parameter values, we conducted sensitivity analysis to observe the contribution of each parameter to the dynamics of the system. We also study the relation between the sensitivity of the parameter and its identifiability, which is related to confidence interval. Large confidence interval and low sensitivity of the

Variable	Mean μ	Std. Deviation σ	Confidence interval $\pm k \times \sigma$
k_u	0.0729 min^{-1}	0.0069	0.0308
k_{ec}	0.0795 min^{-1}	0.0061	0.0271
k_{en}	0.9463 min^{-1}	0.1054	0.4713
k_{ee}	0.0789 min^{-1}	0.0132	0.0589
k_{rc}	0.2256 min^{-1}	0.0713	0.3186
k_{rn}	0.3038 min^{-1}	0.1585	0.7087
k_{lc}	0.0307 min^{-1}	0.0073	0.0328
k_{ln}	0.0028 min^{-1}	0.0009	0.0042
k_{drc}	0.0301 min^{-1}	0.0183	0.0819
k_{drn}	0.5133 min^{-1}	0.2114	0.9448
k_{dlc}	0.0184 min^{-1}	0.0044	0.0197
k_{dln}	0.0332 min^{-1}	0.0022	0.0098
k_{du}	0.1513 min^{-1}	0.0279	0.1245

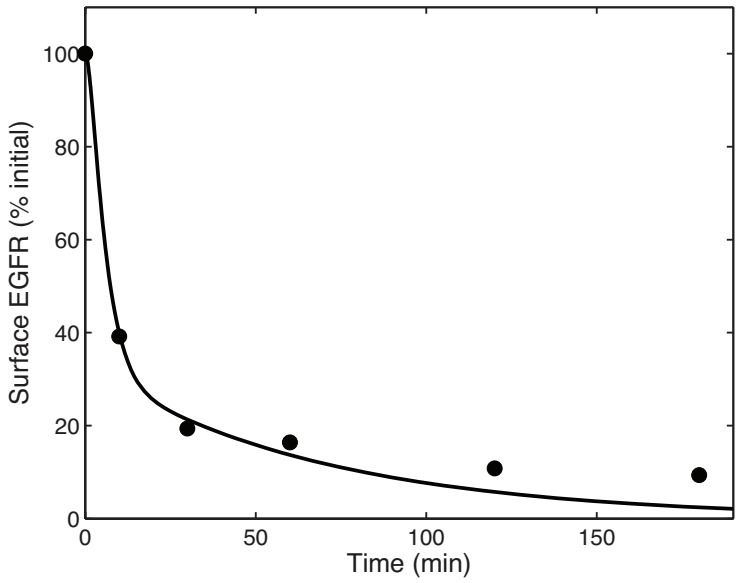
Table 3.2: List of fitted parameters with mean values and confidence interval.

parameter may lead to unidentifiability. In our model we observe that many parameters have huge confidence intervals and low over all sensitivity (Figure 3.10 and Table 3.2).

Sensitivity analysis demonstrates the importance of the parameters in different phases of the dynamics and with respect to species. We find ubiquitination rate (k_u) to be the most sensitive parameter in the system. This finding clearly reflects the importance of receptor-ubiquitin binding in the regulation trafficking. The parameter k_{en} shows low sensitive in all the stages of dynamics thus we speculate that GA is unable to estimate the value for k_{en} (CIE rate) and reaches boundaries of the defined optimization range, even extending the upper bound on the parameter did not provide a unique solution. Thus, we consider the parameter to be unidenti-

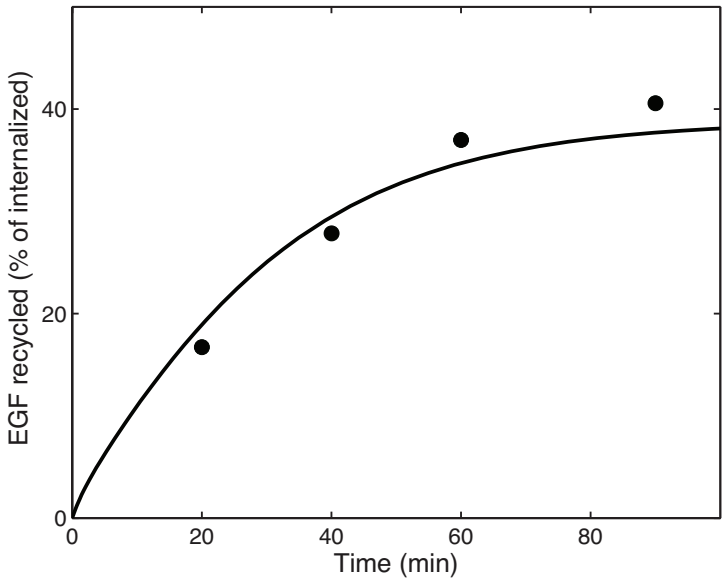


(a) Change in total EGFR in percentage

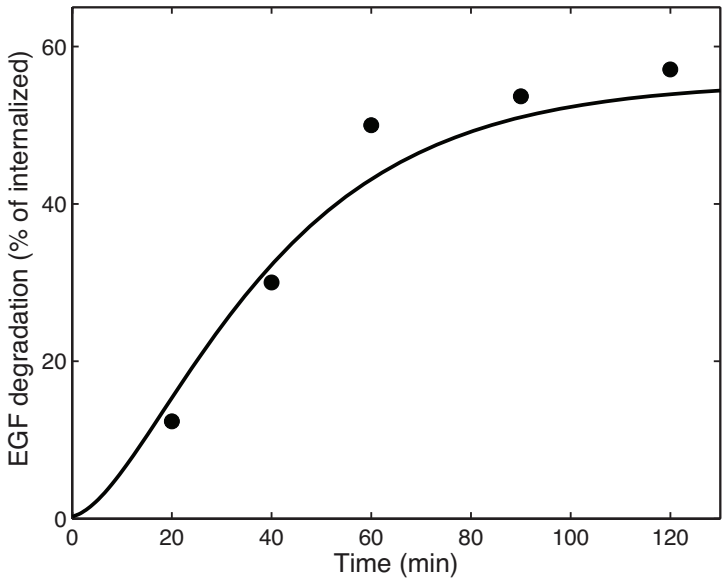


(b) Change in surface EGFR in percentage

Figure 3.8: Model simulation results for mean parameter values showing change in EGFR count with time post high EGF stimulation. Symbols represent experimental data (Sigismund *et al*., 2008).



(a) Recycling of EGF



(b) Degradation of EGF

Figure 3.9: Model simulation results for mean parameter values showing EGF recycling and degradation with time post high EGF stimuli. Symbols represent experimental data (Sigismund *et al*., 2008).

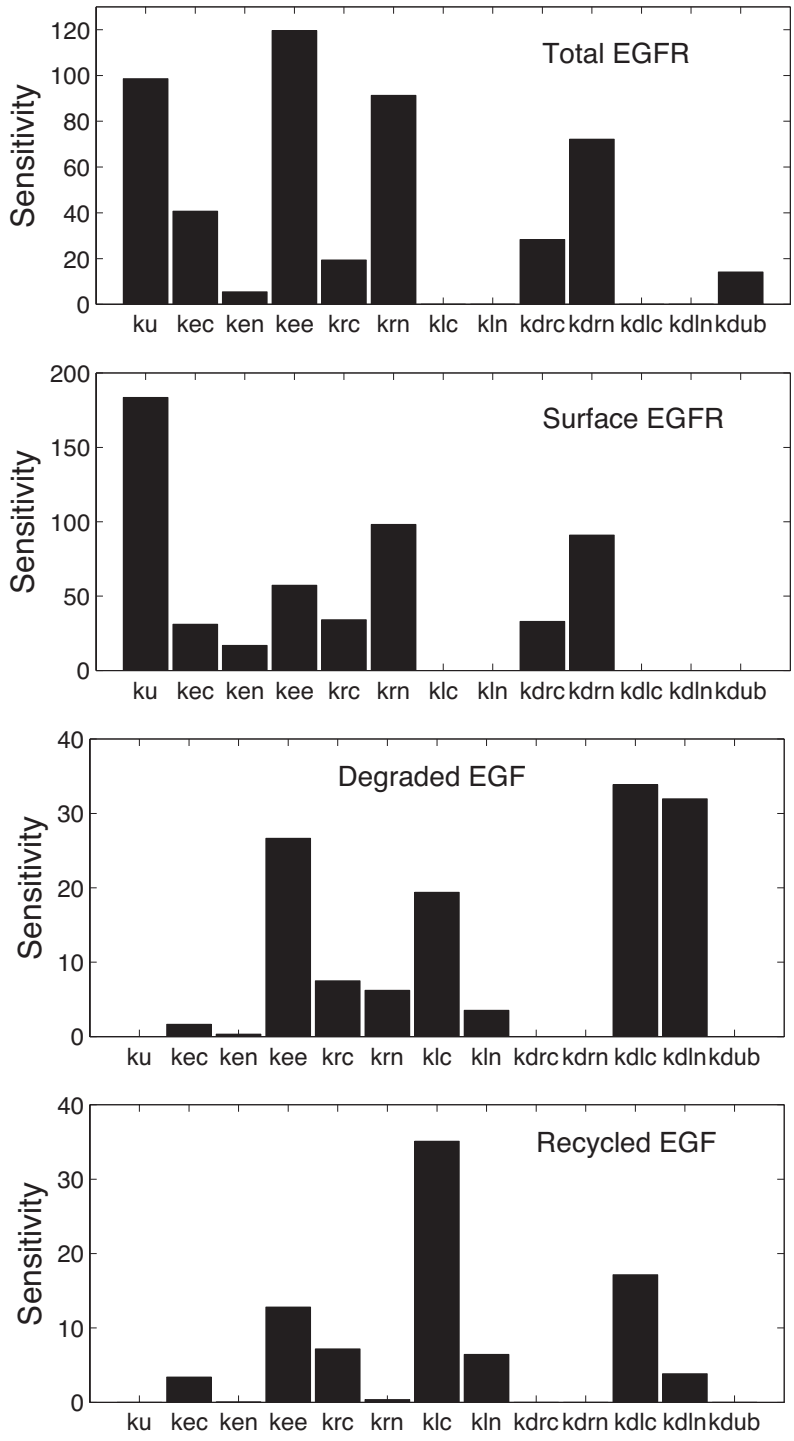


Figure 3.10: Results of sensitivity analysis with mean parameter values.

fiable and define the parameter search space in the range 0.5 to 1. System clearly shows high sensitivity to other trafficking events controlled by k_{ec} and k_{ee} during the initial phase of the dynamics. During the continuous EGF stimuli the parameters controlling the processes at cell surface are more sensitive. The parameters dictating the fate of ligand are insensitive to the total and surface EGFR dynamics. During the chase phase analysis revels the importance of each parameter in EGF recycling and degradation, k_{dlc} and k_{dln} appears to be very sensitive with respect to degradation and k_{lc} with EGF recycling. Parameters controlling the intracellular receptor dynamics show less influence on the EGF recycling and degradation. Due to the absence of EGF stimuli in the chase phase surface parameter do not show any influence on the down stream dynamics. The parameter variability and their sensitivities clearly indicate that not all model parameter can be identified. For most of the systems biology models unidentifiability arises due to the use cumulative measurements for calibration. We argue that the in our study, unidentifiability is more due to the limited data and less due to the structure of the model. In the next section we aim to investigate cause behind the unidentifiability of the model parameters.

3.6 Model identifiability

To further investigated whether the unidentifiability is arising from the model structure or form practical limitations of experimental measurements. We used a MATLAB toolbox 'SensSB' which provides a platform to investigate the structural identifiability. For this purpose the toolbox generate simulated data for all the model variables with zero noise. The simulated data represent an ideal experimental setup where measurements are available for all the species of the system (Figure 3.11). With the simulated data, we can speculate that all the non correlated rate constants

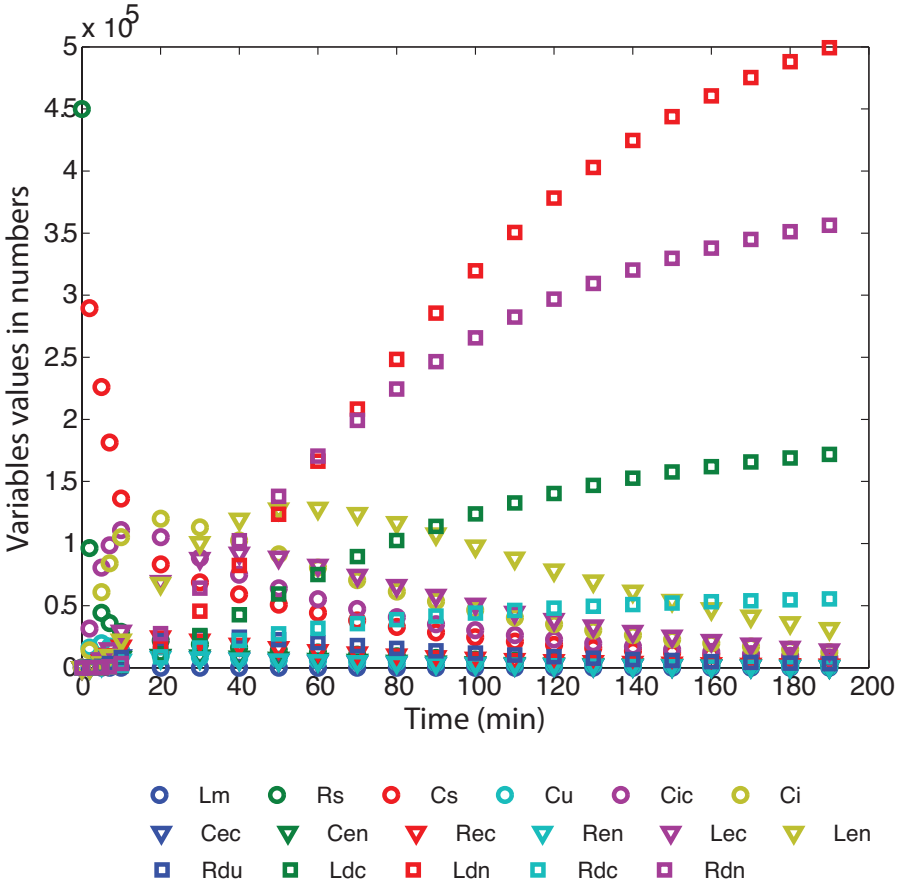


Figure 3.11: Simulated data generated for all the model variables using SensSB. Mean parameter values were used to generate the data.

in the model can be estimated. If there exist a high correlation between parameters then their estimation will be hindered due to the interdependence (Rodriguez-Fernandez *et al*., 2006; Miao *et al*., 2011). The correlation matrix calculated using the toolbox gives information about the interrelationships between the parameters and thereby about the structural identifiability. In the matrix a correlation index value 0.95 or higher signifies high degree of correlation between the two parameters (see Appendix B). For our model, we investigated this interdependence in order to identify the correlated model parameters that are not identifiable due to the structure of the model. Figure 3.12 shows color plot of the parameter correlation matrix

with each squares representing the correlation value between the two parameters. Correlation matrix calculated by SensSB toolbox shows that no pair of parameters

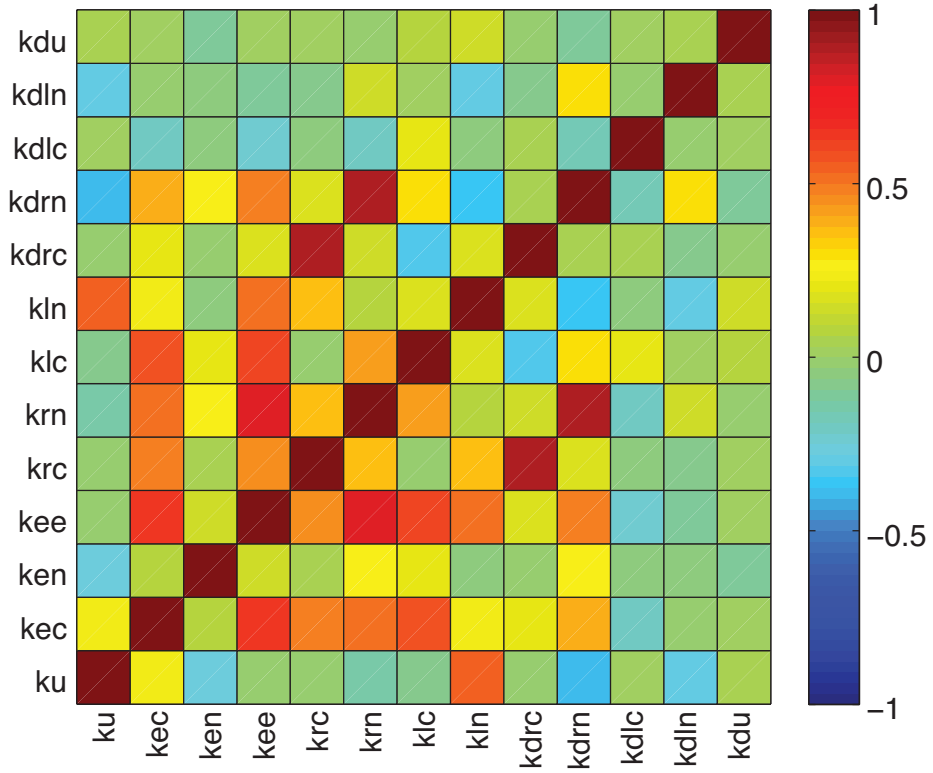


Figure 3.12: Parameter correlation matrix represented as color plot depicting correlations between the parameters.

is highly correlated (no parameter pair with correlation higher than 0.95). These results suggest the model parameters to be identifiable provided that the sufficient data are available. During the analysis we provided data for all the model variables, which is an ideal situation. Under these conditions we find no high interdependence between parameters hence we speculate that model is structurally identifiable. However, the correlation matrix also provides information about the negative and positive correlations between the parameters yet lower than the threshold of 0.95. These low correlations provide information about the interrelationships be-

tween the parameters.

The calculations done so far are based on the local sensitivity analysis carried out in the close range of mean parameter values. Hence, the matrix obtained is also termed as local correlation matrix obtained for local identifiability analysis. A pseudo-global identifiability analysis can be performed using SensSB to investigate the interrelationship between the parameter over a wider range of parameter values. Figure 3.13 shows a color plot of the pseudo-global correlation matrix. The

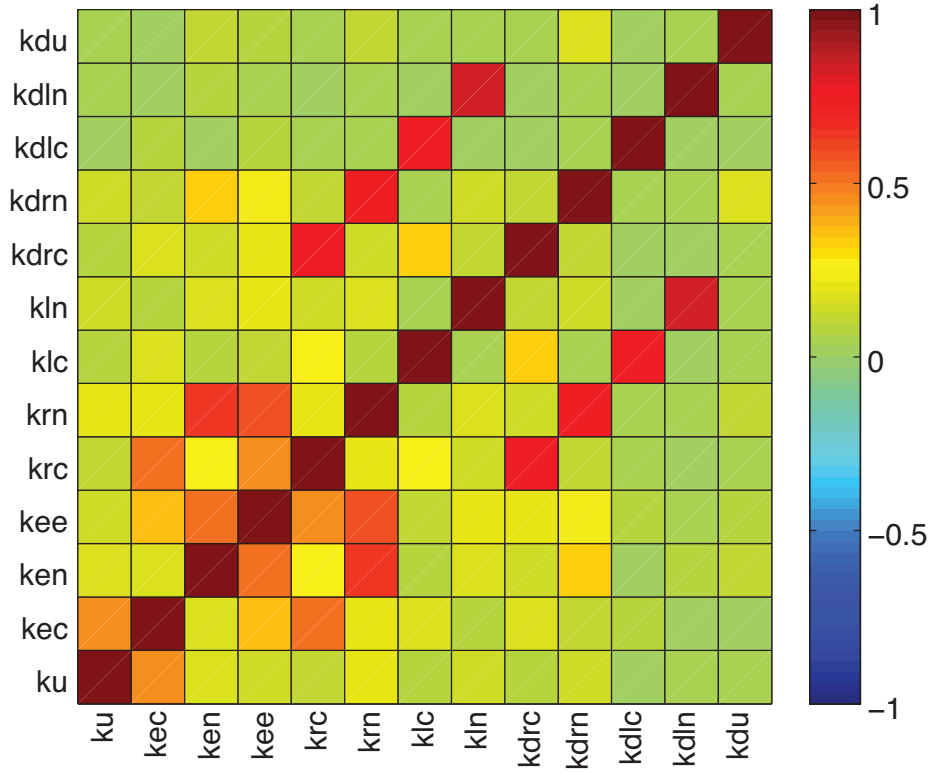


Figure 3.13: Pseudo-global correlation matrix represented as color plot depicting correlations between the parameters.

results of the pseudo-global identifiability analysis reveals that there are no pairs of parameters too highly correlated (no correlations greater than 0.95). The results of local and pseudo-global analysis confirm low interdependence between the

parameters and indicate that model parameters can be identifiable. The findings of identifiability analysis further strengthen our argument that the model is structurally identifiable. Therefore, we propose that the parameter estimation is more limited by the measured data rather than structure of the model.

As mentioned previously the parameter unidentifiability in systems biology models also known as sloppiness is universal and mostly arise due to the technological limitation in the data acquisition (Gutenkunst *et al*., 2007). Considering the limitations in obtaining unique values of the model parameters, it is argued that validation and utility of model should be judged based on the accuracy with which the model can make predictions (Wang *et al*., 2009; Chen *et al*., 2009; Brown and Sethna, 2003; Gutenkunst *et al*., 2007). In the next chapter (Model predictions) we test our model for its capabilities in providing qualitatively as well as quantitatively accurate predictions.

Chapter 4

Model predictions

4.1 Effect of endocytosis inhibitors

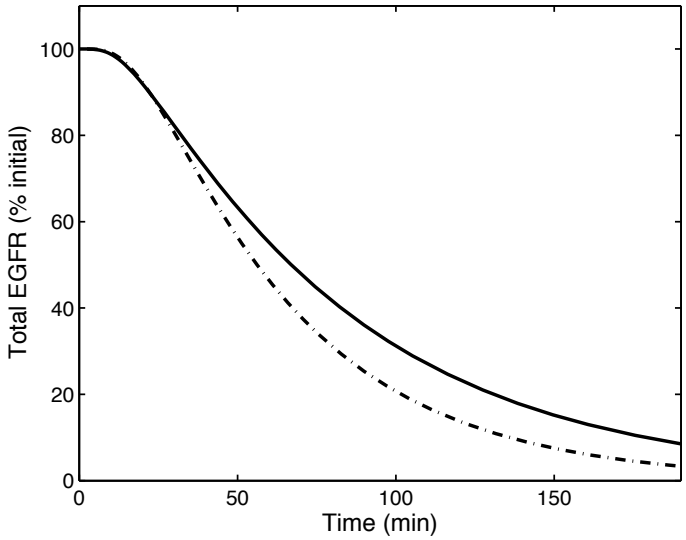
In the model we consider CME and CIE as two independent routes of internalization and rate receptor flow through these routes is governed by the values of rate constants k_{ec} and k_{en}, respectively. Thus, structure of our model permits us to distinguish between two subpopulations of internalized EGFR: CME internalized (ubiquitin-free) and CIE internalized (ubiquitin-bound). It is an advantage over the existing models, which allows us to control each endocytic route and monitor each subpopulation. Model also allows us to specify for each subpopulation a different set of values for rate constants such as internalization rate, recycling and degradation rates for both ligand and receptor. These features of the model let us illustrate the change in EGFR trafficking pattern for obstructed endocytosis route. In order to switch off CME or CIE route in the model, we set the corresponding internalization rate to zero. Simulation results for the closed CME and closed CIE show an altered EGFR localization (Figures: 4.1, 4.2, 4.3, and 4.4).

With CME mode turned off, ligand-receptor complexes undergo ubiquitina-

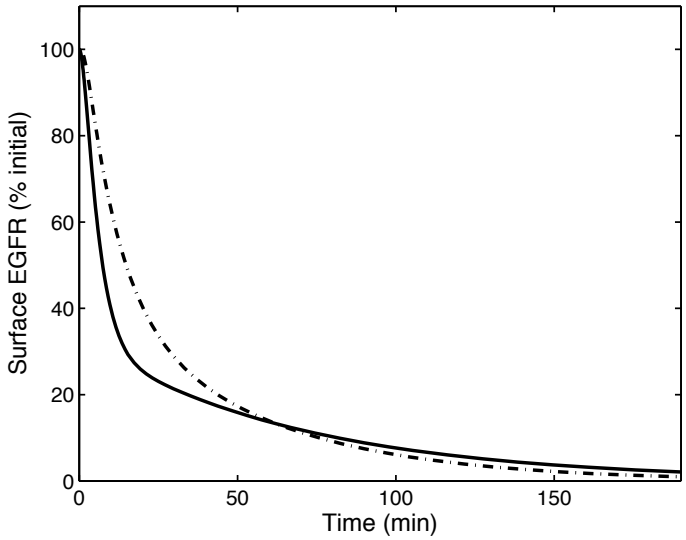
tion activating the CI internalization. As all the internalizing complexes are tagged with ubiquitin, inside the endosome these complexes are targeted to the lysosome at a higher rate compared to their recycling. For the surface EGFR, model results show a slightly delayed drop in the level of EGFR due to a delayed internalization arising from the intermediate ubiquitination reaction (Figure 4.1(b)). In control state, post ligand binding ubiquitin-free receptors can as well internalize through the CME. However, the initial lag in receptor internalization has no significant influence on the later phase of the dynamics. Towards the end, the surface EGFR level for CME inhibition decreases to zero due to the targeting of receptors for lysosomal degradation. In line with the argument that ubiquitin binding efficiently targets the receptors for lysosomal degradation, model results show a steep drop in total EGFR (Figure 4.1(a)).

On the contrary, when CI internalization rate is set to zero, only the ubiquitin-free complexes manage to internalize through CME route and ubiquitin-bound complexes appear to remain at the cell surface. Model simulations show an initial drop in the surface EGFR level due to the clathrin mediated (CM) internalization of ubiquitin-free complexes. The high recycling of these complexes and their dissociated receptors is responsible for the observed reappearance of surface EGFR (Figure 4.2(b)). The total EGFR count drops at a much slower rate than the control because of high recycling and low degradation (Figure 4.2(a)). We compare these model predictions against experimental data. In the experiments, CME route is blocked by clathrin knockdown (clathrin-KD) while a cholesterol-binding drug Filipin is used to hinder the CIE route. We find qualitative agreement between the experimental data and our model predictions. This suggests our model can accurately capture the altered surface and total EGFR localization in presence of endocytosis inhibitors.

Several experimental studies show that the sorting of EGFR in the endo-

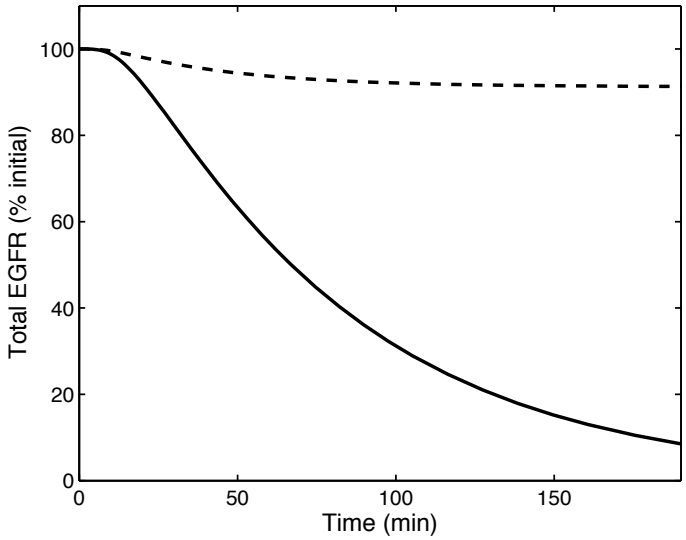


(a) Change in percentage of total EGFR

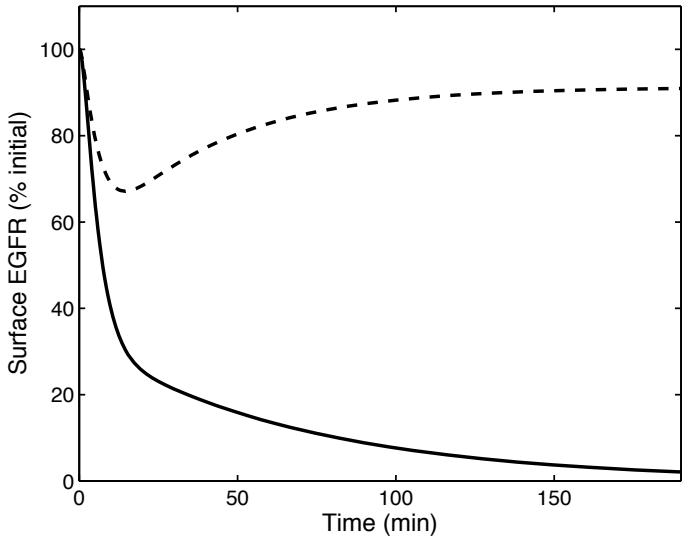


(b) Change in percentage of surface EGFR

Figure 4.1: Model simulation results for mean parameter values showing change in EGFR count with time, post high EGF stimulation. Continuous line shows simulation results for control condition and ( ) represents simulation with k_{ec} set to zero.



(a) Change in percentage of total EGFR

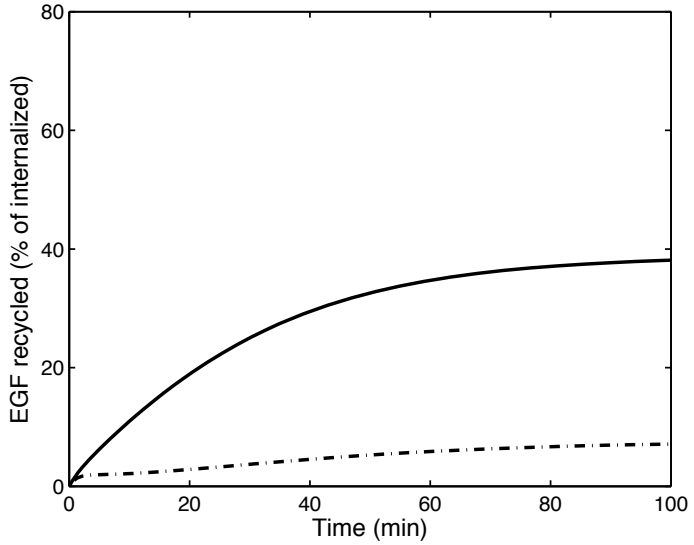


(b) Change in percentage of surface EGFR

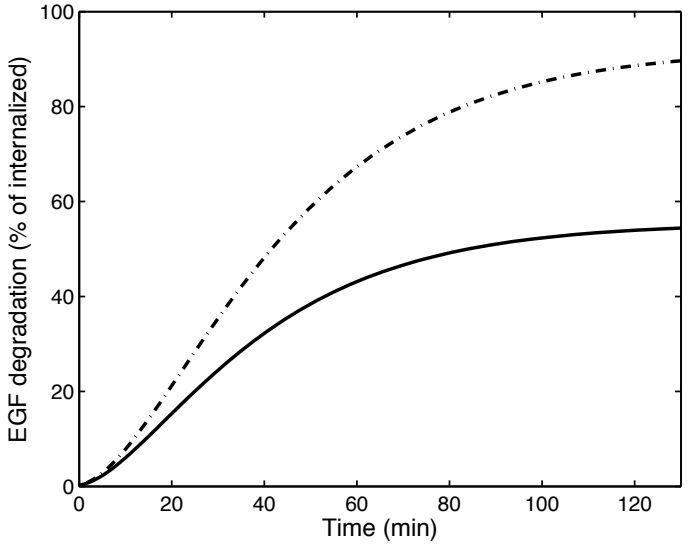
Figure 4.2: Model simulation results for mean parameter values showing change in EGFR count with time, post high EGF stimulation. Continuous line shows simulation results for control condition and (--) represents simulation with k_{en} set to zero.

some depends on the ubiquitination status of the receptor (Haglund *et al*., 2003; Weissman, 2001). Hence, we assume that ubiquitin dependent sorting mechanism is present in the experiments of Sigismund *et al*. and that ubiquitin binding is responsible for targeting the ligand and receptors for lysosomal degradation. Hence, to obtain the proposed sorting, we placed constrains in the optimization algorithm (Table 2.2). This allows the algorithm to search for parameters that not only provide the proposed sorting but also to capture the control measurement data. Remarkably, our model results with endocytosis blocking are in agreement with the inhibition experiments. When the CME route is turned-off, model results show high degradation and low recycling of the internalized EGF in accordance with the proposed sorting (Figure 4.3(b) and 4.3(a)). Upon comparing the model results with clathrin-KD experiments, we find an overestimation of recycling and underestimation of degradation. For the case of CIE inhibition, model results for EGF recycling and degradation appear to show a near quantitative match with Filipin data (Figure 4.4(a) and 4.4(b)). These results show that with the assumption of ubiquitin based sorting, our model can qualitatively explain the change in EGFR localization induced due to the selective inhibition of endocytosis.

Quantitative comparison with inhibition data: Use of clathrin-KD or Filipin fails to provide a complete inhibition of the respective endocytosis. We speculate that the CM internalization with clathrin-KD must be lower than the control. Considering this aspect of the knockdown, we fit the model against the clathrin-KD data with CME rate to be the only free parameter and keep values of rest of the parameter unchanged (Figure 4.5 and 4.6). We use a gradient-based non-linear fitting algorithm, available in MATLAB (lsqcurvefit) to estimate the CME rate. As speculated, the best-fit estimate for CME rate is considerably lower than the con-

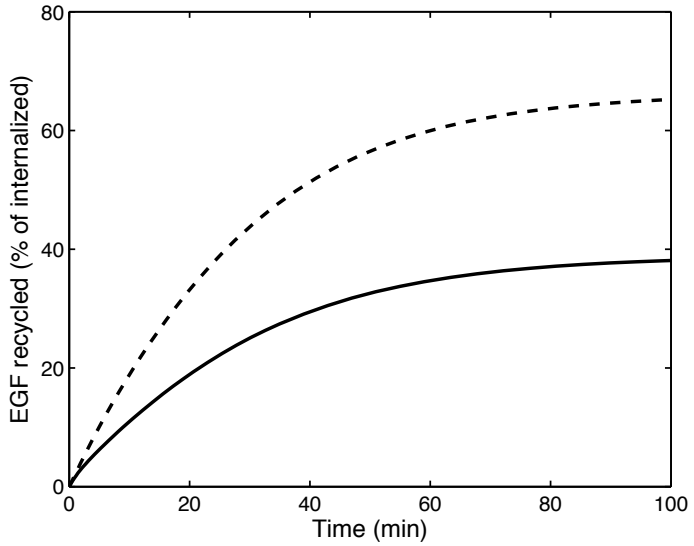


(a) Recycling of EGF

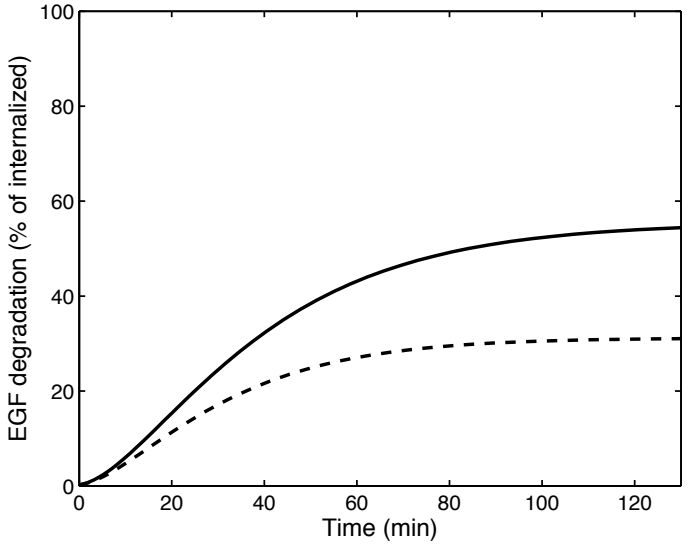


(b) Degradation of EGF

Figure 4.3: Model simulation results for mean parameter values showing EGF recycling and degradation with time, post high EGF stimuli. Continuous line shows simulation results for control condition and (- - -) represents simulation with k_{ec} set to zero.

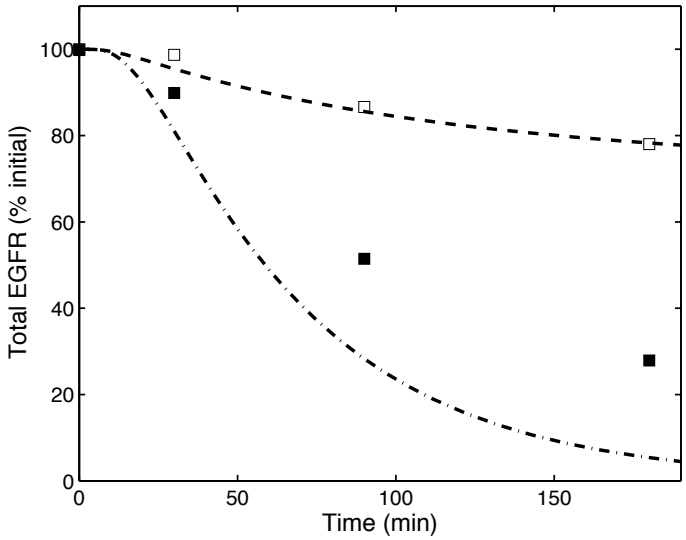


(a) Recycling of EGF

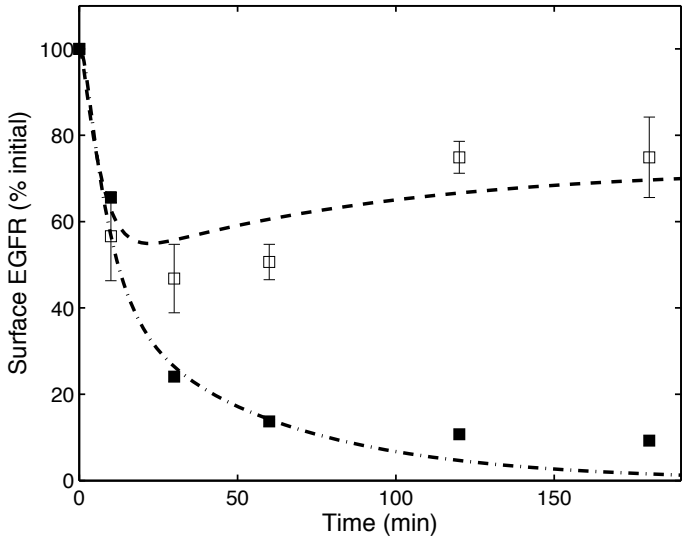


(b) Degradation of EGF

Figure 4.4: Model simulation results for mean parameter values showing EGF recycling and degradation with time, post high EGF stimuli. Continuous line shows simulation results for control condition and (--) represents simulation with k_{en} set to zero.

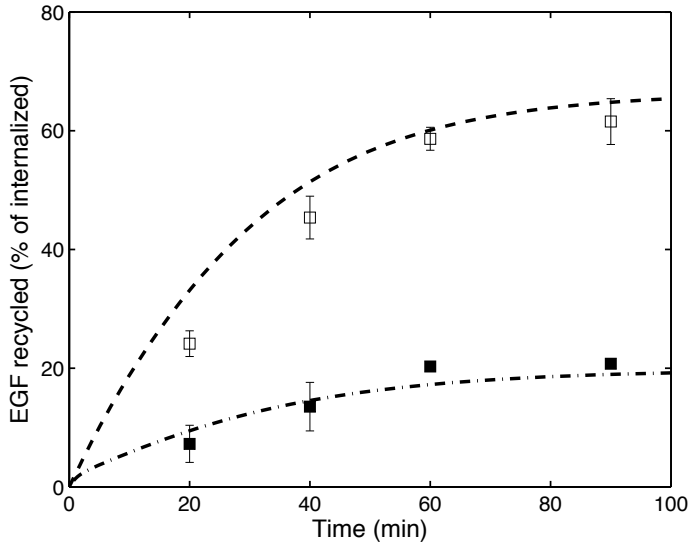


(a) Change in percentage of total EGFR

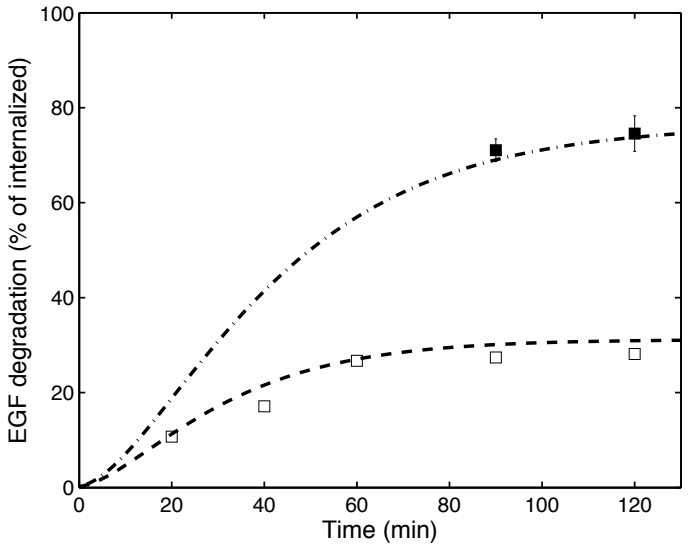


(b) Change in percentage of surface EGFR

Figure 4.5: Model simulation results showing effect of clathrin KD (- - -) and Filipin (– –) on the change in EGFR count with time, post high EGF stimulation. Symbols represent data for clathrin-KD (■) and Filipin (□) (Sigismund *et al*., 2008).



(a) Recycling of EGF



(b) Degradation of EGF

Figure 4.6: Model simulation results showing effect of clathrin KD (- - -) and Filipin (– –) on EGF recycling and degradation with time, post high EGF stimuli. Symbols represent data for clathrin-KD (■) and Filipin (□) (Sigismund *et al*., 2008).

trol value. With the use of estimated CME rate, our model captures dynamics of surface EGFR, degradation and recycling of EGF (Figure 4.5(b), 4.6(b) and 4.6(a)). Compared to the clathrin-KD data, our model over-estimates the loss in total EGFR (Figure 4.5(a)). The reason behind this increase in half-life of total EGFR remains unclear.

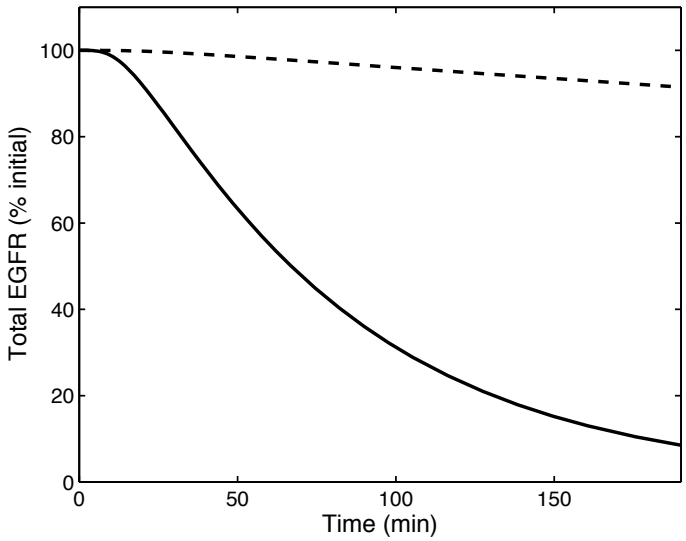
Furthermore, Sigismund *et al*. use a cholesterol-blocking drug called Filipin in order to obstruct the CI internalization. As it has been argued that the change in cholesterol level can influence the EGFR activity (Ringerike *et al*., 2002), we suspect that the presence of Filipin may have an influence on the ubiquitination process. Considering this effect, we set ubiquitination rate and CIE rate as free parameters while fitting the model against the Filipin data and keep the values of other parameters unchanged. We use the same fitting algorithm from MATLAB as above. Results of this fitting provide estimates for both the free parameters that are lower than the control, with the value of CIE rate value being equal to zero. With two refitted parameters, our model successfully captures the data (Figure 4.5 and 4.6).

The model results obtained with and without fitting show that for high EGF dose our model can accurately capture the dynamics of EGFR trafficking. Further, our model is able to reproduce the effect of endocytosis inhibition on surface EGFR, total EGFR and on EGF degradation and recycling. These results further support the assumption that ubiquitination is responsible for surface as well as endosomal sorting of the receptors.

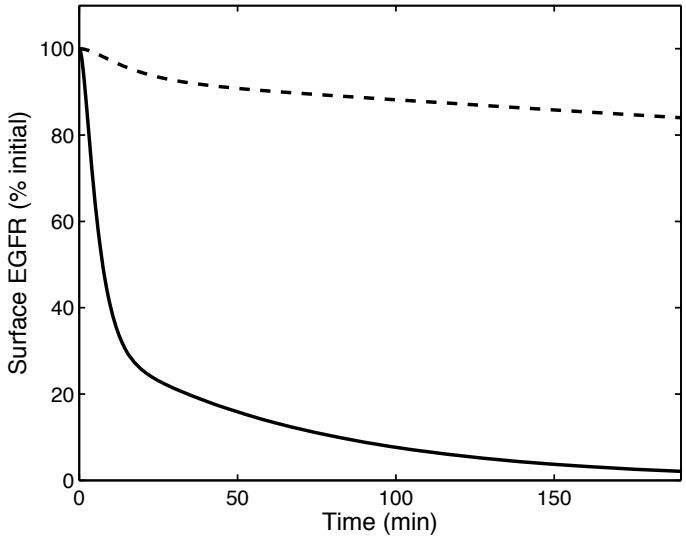
4.2 EGFR transport at low EGF dose

Thus far, we use the model to study the EGFR dynamics in high EGF condition. Due to the use of Hill kinetics, the extent of receptor ubiquitination follows a sigmoidal behavior as a function of EGF concentration. This causes negligible receptor ubiquitination for low EGF level whereas at high EGF level, significant amount of receptors undergo ubiquitination. The negligible amount of ubiquitin binding generates insignificant amount of CI internalization. Hence, at low EGF stimuli, most of the ligand-receptor complexes are ubiquitin-free and internalization is predominately through CME route. Here we use our model to study change in the dynamics of EGFR transport with respect to EGF stimuli. We simulate the model with extracellular ligand concentration (L_m) set to a physiologically low value (1.5 ng/ml) (Sigismund *et al*., 2008, 2005) and use the parameter values obtained through model calibration with high EGF control measurements.

Simulation results when compared to high EGF, show considerably slower drop in the surface EGFR level (Figure 4.7(b)). The total EGFR population drops at a slow pace, showing a prolonged EGFR trafficking (Figure 4.7(a)). Inside the endosome, majority of the internalized complexes, dissociated receptors and ligands are ubiquitin-free. According to proposed ubiquitin dependent endosomal sorting, these complexes, receptors and ligands are more liable for recycling than for degradation. Results for EGF recycling and degradation show a similar trend. Contrary to the high EGF stimuli, model simulations for low stimuli show a decrease in degradation and an increase in recycling of EGF (Figure 4.8(b) and 4.8(a)). Remarkably, results of the model simulation are in quantitative agreement with the low EGF data (Sigismund *et al*., 2008) (Figure 4.8(b)). These results show that our model can explain the dynamics of EGFR for low EGF stimuli. It also implies that at low EGF

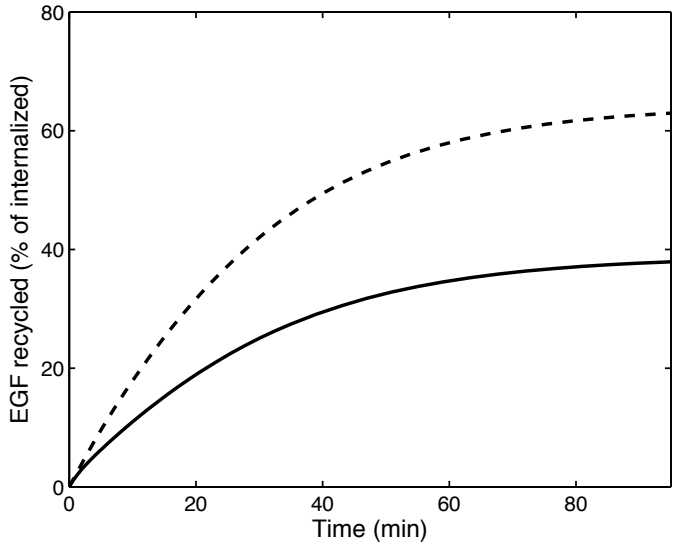


(a) Change in percentage of total EGFR

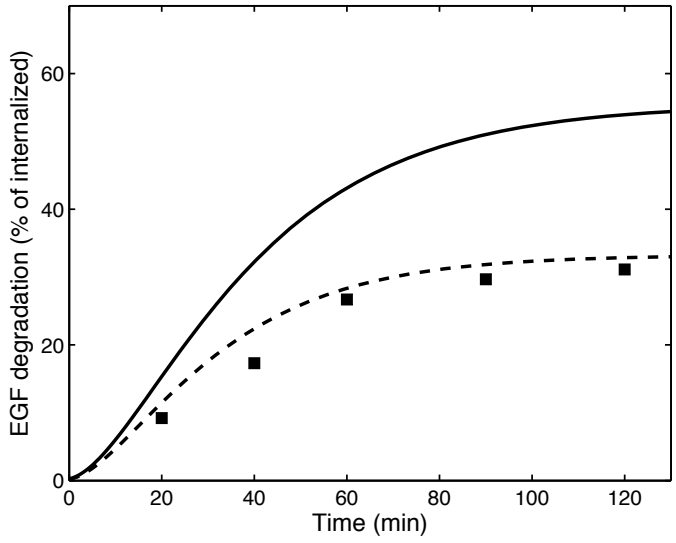


(b) Change in percentage of surface EGFR

Figure 4.7: Model simulation results show change in EGFR count with time, for low EGF stimuli (--) and high EGF stimulation continuous line.



(a) Recycling of EGF



(b) Degradation of EGF

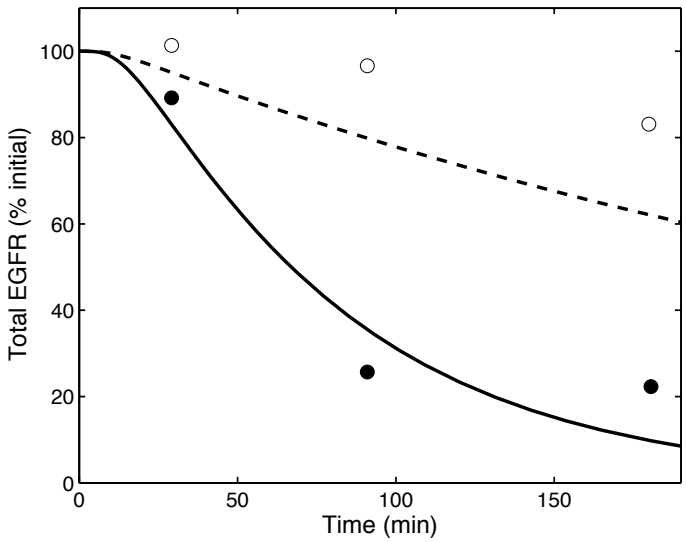
Figure 4.8: Model simulation results show change in EGFR count with time, for low EGF stimuli (--) and high EGF stimulation continuous line. Symbols represent data for low EGF (■) (Sigismund *et al*., 2008).

dose, ubiquitin binding is negligible and supports the assumption of sigmoidal behavior of ubiquitin binding as a function of EGF dose. From these results, it can be interpreted that very low receptor ubiquitination leads to low degradation and high recycling which keeps the receptor in circulation for longer time. This prolonged EGFR trafficking causes a sustained signaling (e.g. AKT and ERK signaling) as seen in findings of Sigismund *et al*. (2008).

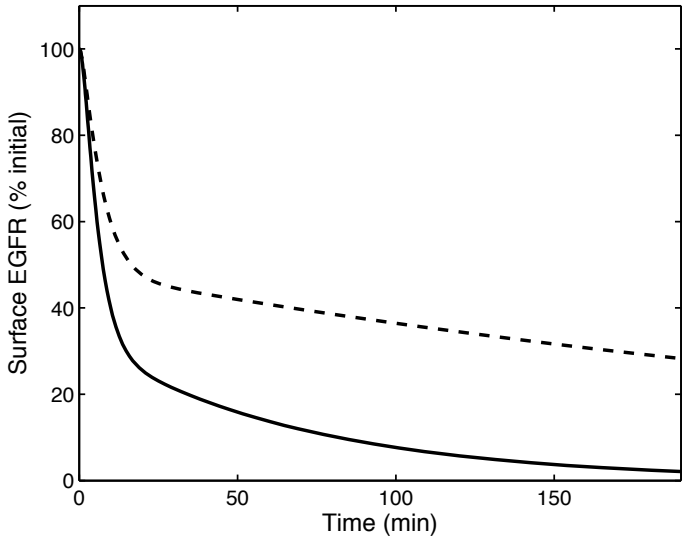
4.3 Dynamics of ubiquitination defective EGFR

Mutations in EGFR are known to cause dysfunctional EGFR trafficking leading to cancer and other cell cycle related defects (Gschwind *et al*., 2004). Defective ubiquitination in EGFR is considered responsible for prolonged EGFR signaling and is associated with cancer (Gschwind *et al*., 2004; Peschard and Park, 2003). Using our model, we aim to understand the effect of defective ubiquitination on the dynamics of EGFR transport. We also validate the model predictions against the experimental data of Y1045F EGFR mutant (supplementary text from Sigismund *et al*. (2008)). In these experiments, total EGFR and EGF recycling are measured for both wild type EGFR (wt-EGFR) and Y1045F mutant EGFR (mutant-EGFR). In order to avoid the interference from endogenous EGFR, the mutant-EGFR is expressed in NR6 fibroblasts lacking endogenous EGFR and for control measurements wt-EGFR is expressed.

To simulate ubiquitin defective mutant-EGFR, we set the ubiquitination rate to zero. Interestingly, our model qualitatively predicts the experimentally observed trafficking of both wt-EGFR and mutant-EGFR (Figure 4.9 and 4.10). Model predictions show that with no receptor ubiquitination, CIE route remains inactive forcing all the ligand-receptor complexes to internalize through CME route. Inside

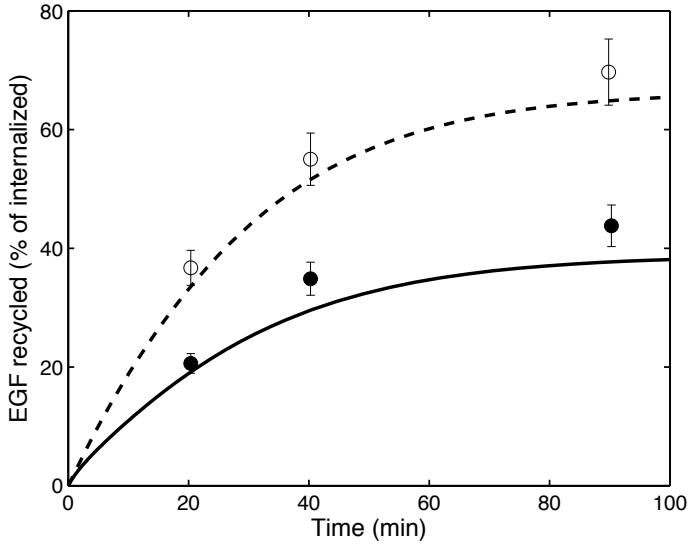


(a) Change in percentage of total EGFR

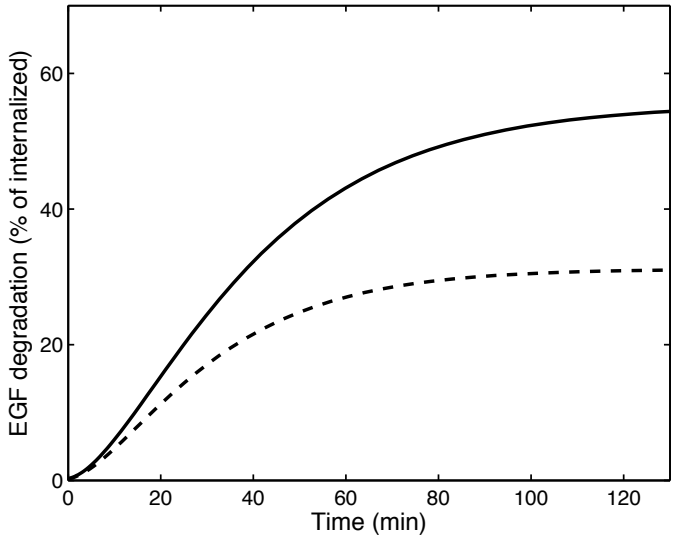


(b) Change in percentage of surface EGFR

Figure 4.9: Model simulation results show change in EGFR count with time, for mutant EGFR (--) and wt-EGFR continuous line. Symbols represent data for mutant (○) and wild-type (●) EGFR (Sigismund *et al*., 2008).



(a) Recycling of EGF



(b) Degradation of EGF

Figure 4.10: Model simulation results show change in EGFR count with time, for mutant EGFR (--) and wt-EGFR continuous line. Symbols represent data for mutant (○) and wild-type (●) EGFR and error bars represent 8% of mean (Sigismund *et al*., 2008).

the endosome, these ubiquitin-free complexes are sorted for recycling with higher probability than for degradation. Model predictions for the mutant EGFR show that after the initial decline, the surface EGFR level remains at higher level compared to the wt-EGFR (Figure 4.9(b)). Low degradation keeps mutant-EGFRs in circulation for longer time, causing an increase in half-life of the receptors. Our model predictions show a qualitative match with data of total mutant-EGFR (Figure 4.9(a)) and near quantitative match with recycled EGF data (Figure 4.10(a)). The drop in EGF degradation is also visible in model predictions (Figure 4.10(b)). Experiments with Y1045F mutant show no significant difference in the behavior of total EGFR and EGF recycling in the presence or absence of Filipin. The unaltered dynamics of mutant-EGFR in the presence of Filipin suggests a link between ubiquitination and the activation of CIE. With the ubiquitination rate set to zero in our model, any change in the CIE rate will not have an influence on the trafficking of mutant-EGFR. These results support the model assumption that receptor-ubiquitin binding is a prerequisite for active CI internalization.

Comparison between simulation results of wt-EGFR and mutant-EGFR clearly reflect the key function of receptor ubiquitination in down-regulating EGFR signaling. Model simulations show that ubiquitin binding initiates CI internalization and targets the ubiquitinated receptors for lysosomal degradation. Defect in ubiquitin binding leads to prolonged trafficking of mutant-EGFR due to low degradation and high recycling. Experimental data show an over-activation of AKT and ERK for Y1045F mutant (Sigismund *et al*., 2008). The model predictions can explain these results as high recycling, low degradation, and increased recirculation of signaling EGFRs could cause an over-stimulation of AKT and ERK pathway.

Chapter 5

Conclusions and discussion

Readjustments in the speed of EGFR recycling and degradation depending on the amplitude of stimuli can control overstimulation in signaling. In case of EGF induced signaling, shift from CME to CIE route changes the EGFR localization and modifies the signaling. In this work, we provide a mathematical model of EGFR trafficking with receptor ubiquitination as a key regulatory mechanism that alters the EGFR transport. Experimental studies using ubiquitin defective EGFR mutant show that receptor ubiquitination is not necessary for the receptor endocytosis (Jiang and Sorkin, 2003; Waterman *et al*., 2002). In addition, it is being proposed that receptor ubiquitination activates clathrin-independent route of endocytosis and ubiquitin binding is observed to be a prerequisite for the CI internalization of EGFR (Sigismund *et al*., 2005). These observations are in favor of the assumption that ubiquitin-free complexes internalize through CME and receptor ubiquitination at the cell membrane initiates the CI internalization.

The structure of our model allows us to study the effect of ubiquitination on the EGFR degradation and recycling. Unlike existing models of EGFR trafficking, our model specifically includes dose dependent receptor ubiquitination that func-

tions as a regulatory step and incorporates two modes of endocytosis. We calibrated the model using the experimental data from Sigismund *et al*. (2008). High correlation between sensitivities of different fits shows that the dynamics of the system remains unchanged among obtained fits (Figure 3.6 and 3.7). Considering the high correlations, we speculate that the obtained fits are in the neighborhood of a global optimal. In order to get an accurate estimate for the parameters, a large number of parameter sets are obtained by running the GA code multiple times. Using these fits we calculated mean and confidence interval over the parameters. We observe that the simulations carried out using mean parameter values can reproduce the experimental data (Figure 3.8 and 3.9). Thus, we reason that using mean parameter values our model can make accurate predictions.

The experimental observations suggest a switch-like behavior in ubiquitin binding in response to the rising extracellular EGF concentration (Sigismund *et al*., 2005). Different review articles also argue for a similar idea of concentration dependent ubiquitination initiating EGFR sorting (Acconcia *et al*., 2009; Woelk *et al*., 2007). This points to a regulatory mechanism responsible for the surface sorting of ligand-receptor complexes depending on the EGF concentration. Use of Hill kinetics in our model gives rise to a sigmoidal behavior in ubiquitin binding rate. Due to which model predicts an insignificant ubiquitin binding at low EGF dose. At high EGF dose, a sharp increase in ubiquitination divides receptors into ubiquitin-free subpopulation internalizing through CME and ubiquitin-bound subpopulation that uses CIE. Speedy degradation of the ubiquitin-bound receptors then causes rapid down-regulation in EGFR trafficking and associated signaling.

Another explanation about EGFR surface sorting comes from a computational study that assumes CME saturation at high EGF stimuli (Schmidt-Glenewinkel *et al*., 2008). Authors claim overloading of the CME route forces the receptors

into the CIE route reasoning that at high EGF concentration, depletion of limiting factors such as adapter proteins and/or cage proteins during CME causes saturation. However, an experimental study by Warren *et al*. (1997) shows saturation of transferrin receptor (TfR) (20 fold increase in TfR count), which exclusively uses CME, does not have a significant influence on the endocytosis of EGFR. These findings clearly indicate that internalization rate of EGFR remains unaffected, although there is an excessive consumption of clathrin molecules for CM internalization of overexpressed TfR. In the study of Sigismund *et al*. (2008) knockdown of adaptor protein-2 (AP2) is relatively less effective, in blocking CM internalization as compared to clathrin-KD. This observation signifies that limitation of adaptor protein has lesser influence on the CME probably due to the presence of alternative adaptor proteins. At low EGF dose, knockdown of CME components may induce saturation in CME and activation of CIE route. In contrast, knockdown experiments with low EGF show no significant utilization of CIE route (Sigismund *et al*., 2008, 2005). In addition, we argue that continuous recycling maintains a steady supply of CME component molecules at the cell membrane. Thus, we propose that CIE route is activated irrespective of CME saturation.

To study the role of different modes of endocytosis in EGFR trafficking, we block each of the endocytosis routes alternatively by setting corresponding internalization rate to zero (Figure 4.1, 4.3 and 4.2, 4.4). With the CIE rate set to zero, model predictions show increased recycling which leads to a prolonged EGFR trafficking and signaling. In contrast, restricting the CM internalization causes surface EGFR to undergo ubiquitination prior to internalization through CIE route. Increased degradation and reduced recycling lead to termination of EGFR transport and associated signaling. These simulation results show that EGF receptors meet different fate depending on their choice of internalization route. In a similar fashion,

transforming growth factor β (TGF-β) receptors meet different fate and activate different signaling pathways depending upon the mode of endocytosis (Di Guglielmo *et al*., 2003). The study also finds degradation of CIE routed receptor to be responsible for negative regulation of TGF-β signaling.

For low EGF concentration, our model predictions show that receptor-ubiquitin binding reaction occurs at a slow rate, producing insignificant amount of ubiquitinated receptors (Figure 4.7 and 4.8). CI internalization is minimal and CME appears dominant, keeping receptors in circulation for a longer time allowing a prolonged EGFR trafficking. The model predictions, for both high and low EGF dose endorse the existence of a switch like behavior in ubiquitination and subsequently support the assumption of EGF stimuli dependent change in the preference of internalization route. A similar shift in mode of endocytosis depending on the stimulus concentrations has been reported for platelet-derived growth factor induced receptor trafficking (De Donatis *et al*., 2008).

Ubiquitination deficient EGFR mutants are connected to various cancers. In our model, we simulate one such mutant by setting the ubiquitination rate to zero (Figure 4.9 and 4.10). Predictions show significant difference in the surface and total EGFR level, with mutant receptors showing a slower drop than wt-EGFR. Defective ubiquitination causes minimal lysosomal degradation leading to prolonged EGFR trafficking. In addition, the model predictions for ligand degradation and recycling agree with experimental data of Jiang and Sorkin (2003). We argue that the prolongation of EGFR transport is responsible for experimentally observed overstimulation in AKT signaling (Sigismund *et al*., 2008). These simulation results reflect the importance of receptor ubiquitination in controlling the EGF induced signaling.

The simulation results illustrate the regulatory function of ubiquitin binding

in EGFR trafficking and implications of defective ubiquitination. Yet, the mechanism that converts a graded extracellular signal into an all-or-none response of ubiquitin binding is unknown. Based on the nature of ubiquitin binding various theories have been proposed. One speculation is based on a possible Grb2 mediated cooperative Cbl binding (Acconcia *et al*., 2009). However, some other mechanisms such as the E1-E2-E3 enzyme reaction cascade and/or possibility for cooperative binding during poly-ubiquitination may give rise to ultrasensitivity. Details about the mechanism of EGFR-ubiquitin binding will help explore functioning of the sensing mechanism that cells use to detect change in the environment. Considering the limitations, we use Hill kinetics to replicate the proposed switch like response in receptor ubiquitination. This simplification allows us to develop a usable model, yet it limits the reach of our understanding about the mechanistic details of ubiquitin binding and their role in activation of CIE.

Inside the endosomes, sorting of receptors depending on the ubiquitination status plays a key role in deciding the fate of the signaling receptors. It is well recognized that ubiquitin bound receptors are more prone to lysosomal targeting. Details about the functioning of ubiquitin detection system involved in the endosomal sorting remain obscure, limiting further advancements in the model. In our model, difference in the rate constant values for the ubiquitin-bound and ubiquitin-free receptor populations is responsible for the proposed endosomal classification. It is of great interest to identify the key components that participate in the detection of ubiquitin-bound receptor. The answer to this query may lie in the choice of CIE. Further investigations are needed to gain a deeper understanding about the CIE activation and its influence on the endosomal configuration. Results of the model simulations approve of ubiquitin regulated surface and endosomal sorting of EGFR. Yet, it is too early to generalize the idea of ubiquitin regulated sorting, as

across different cell types and/or subtypes diverse mechanisms may exist that regulate the EGFR signaling. Hence, to avoid misinterpretation it will be appropriate to limit the scope of the model to experimental conditions under which the measurements are made (Sigismund *et al*., 2008).

Simulation results clearly indicate that our model captures the experimentally observed effects of endocytosis inhibition reflecting the regulatory role of ubiquitination in EGFR endocytosis and in the downstream sorting. Model predictions for low EGF dose emphasize the presence of a switch like behavior in receptor ubiquitin binding. Model simulations for mutant EGFR corroborate the importance of ubiquitin binding in the down-regulation of EGFR signaling. The various consequences of ubiquitin binding in the EGFR trafficking observed through model predictions redefine the role of receptor ubiquitination.

Considering the results obtained from model simulations, we propose that receptor ubiquitination activates the clathrin independent endocytosis and initiates surface as well as endosomal sorting responsible for the down-regulation of EGFR trafficking. In conclusion, our model provides insights into the underlying dynamics of dose-dependent EGFR trafficking and is able successfully explain the experimental observations. By adding various signaling pathways in it, our model can be extended to study the effects of altered EGFR localization on the signaling. This extended model can function as a framework that explains the dose-dependent modulation of EGF induced signaling. The knowledge gained from the extended model can be helpful in understanding how cells attenuate signaling and avoid overstimulation in order to escape cancer and other cell cycle related disorders.

Appendix A

Genetic algorithm

Genetic algorithms (GA) are the most popular evolutionary computational methods known for being the most powerful and broadly applicable stochastic optimization techniques (Gen and Cheng, 2000). GAs were invented and developed by John Holland and his students between 1960s and 70s in University of Michigan. Basis for the algorithm comes from the idea of natural selection observed during the evolution of species. Concept behind the GA is to evolve a population of solutions to any given problem using the operators responsible for inducing genetics variations during the natural selections (Mitchell, 1996).

Inspiration from biology

Cells in any living organism carry chromosomes (strings of DNA) containing genetic information of the organism. Each chromosome made up of small fragments called genes that encodes for a proteins. During the sexual reproduction together the two processes, crossover (or recombination) that causes exchange of genes between the two patents' single chromosomes and mutation which is a copying error,

introduce variations in the offspring's genome. This variation comes at cost of fitness, typically defined as the probability to survive and probability to reproduce.

GA mimics the sexual reproduction assuming a candidate solution to be the chromosome and a gene as a single entity in the solution. Crossover is the exchange of entities between different candidates and mutation denotes the random change in the entity. In our case the candidate solution (or individual or chromosome) is a set containing 13 values given to 13 unknown parameters that we intend to estimate and a single parameter value in the set hence becomes a gene. Crossover mixes different individuals and mutation modifies the value of parameters, introducing variations in newly produced individuals. Fitness of every individual is estimated using an objective function or a cost function. In our simulations objective function quantifies the difference between the solution provided by the individual and the experimental observations. Fitness of an individual is inversely proportional to the difference calculated by the objective function. Continuous cycle of crossover and mutation produces more and more fit individuals, which pushes the solution to the optimal value. Hence, factors governing the crossover and mutation in population are important to insure a speedy optimization. A typical genetic algorithm has following steps (Mitchell, 1996):

1. Generate a random population of desired size n as first guess.
2. Calculate the fitness for each individual in the population.
3. Continue the following process till offspring population reaches to n
    (a) Select individual based on the fitness with fittest given the high selection probability.
    (b) Crossover the selected individuals at randomly chosen points to form new offspring.

(c) Mutate the offspring at a random location.

4. Replace the old population with the new offspring population.

5. Return to step 2 until an optimal solution is obtained.

In the computational language one iteration in the GA is called a generation and the number of iterations vary depending on the requirements or based on the desired optimal of an objective function. The entire set of generation is termed as a run. GA is a stochastic algorithm hence each run may produce completely different solution sets. Therefore, an average over multiple GA runs is more reliable than single run. In the recent time the GAs improved to perform fast and converge in less iteration. MATLAB global optimization toolbox offers a GA function, which can be accessed to perform parameter estimation. We build our optimization code that calls the GA function in the toolbox. Using the optimization code we obtained numerous best fit solutions.

Optimization code in MATLAB

```
function mygacoderuninadvance
global datafil tot sur deg rec delow nh
%load data files
load sigismundetalfig1b.txt;
load sigismundetalfig1a.txt;
load sigismundetalfig3a.txt;
load sigismundetalfig5a.txt;
load sigismundetalfig4cLow.txt;
%parameter range file
load UbLb.txt
```

```
%Lower bound[0.0010 0.5000 0.0010 0.0010 0.0010 0.0010
%0.0010 0.0010 0.0010 0.0010 0.0010 0.0010 0.0010]
%Upper bound[0.1500 1.0000 0.1500 0.1500 0.5000 0.7000
%0.1500 0.1500 0.1500 1.0000 0.5000 0.1500 0.5000]
%Reset the random number generator to CPU clock
RandStream.setDefaultStream(RandStream('mt19937ar','seed',...
    sum(100*clock)));
sur=sigismundetalfig1b;
tot=sigismundetalfig1a;
deg=sigismundetalfig3a;
rec=sigismundetalfig5a;
delow=sigismundetalfig4cLow;

Lb=UbLb(1,:);
Ub=UbLb(2,:);

datafil=[sur; tot; deg; rec; 6 60; 6 40; 0 1; 0 1; 0 1; 0 1];
pops = 100;
gens = 500;
Elite=3;
iv=[Lb;Ub];
nh=2;
for i=1:100

options=...
gaoptimset('Display','off','Generations',gens,'StallGenLimit',...
20,'CrossoverFraction',0.8,'TolFun',1e-3,'PopulationSize',pops,...
'MutationFcn', {@mutationadaptfeasible},'PopInitRange',iv,...
'CrossoverFcn',{@crossoverscattered},'FitnessScalingFcn',...
{@fitscalingrank},...
'SelectionFcn',{@selectionstochunif},'EliteCount',Elite);
[p,fval,exitflag,output,final_pop,scores] =...
    ga(@modelobj, 13,[],[],[],[],...
      Lb,Ub,[],options);
```

```
  M=[i,p,fval];
dlmwrite('myresults.txt',M,'delimiter',...
    '\t','precision',6,'-append');
z=strcat('advpset',num2str(i));
if(fval<180)
dlmwrite('mybestfits.txt',M,'delimiter',...
    '\t','precision',6,'-append');
save(z);
end
end

end
%Objective function with constraints
function objective=modelobj(beta)
global datafil
xfil=datafil(:,1);
yfil=datafil(:,2);
yhatfil=mixmodelfil(beta,xfil);

objective=sum((yhatfil(1:10)-yfil(1:10)).^2)...
    +sum((yhatfil(16:19)-yfil(16:19)).^2)...
    +sum((yhatfil(11:12)-yfil(11:12)).^2)...
    +sum((yhatfil(14:15)-yfil(14:15)).^2);
if (yhatfil(20)<60)
    objective=1e6;
end
if (yhatfil(21)>40)
    objective=1e6;
end
if (yhatfil(22)>yhatfil(23))
    objective=1e6;
end
if (yhatfil(24)>yhatfil(25))
    objective =1e6;
end
```

```
if (yhatfil(26)>31)
    objective = 1e6;
end
if (yhatfil(27)<80)
    objective = 1e6;
end
if(yhatfil(28)<60)
    objective = 1e6;
end
if(yhatfil(29)>20)
    objective=1e6;
end
end

%Calculation of simulated data for objective function
function yhat=mixmodelfil(beta1,x)
Vx=1e-10;
Ve=1e-14;
Nav=6.02214E23;
R0=4.5e5;
%Total and Surface EGFR
L0=15.771;
I=zeros(17,1);
I(1)=L0;
I(2)=R0;
t1=x(1:6);
t2=x(7:10);
[T,Y] = ode15s(@devcelleqsfil,t1,I,[],beta1);
sur=100.*(Y(:,3)+Y(:,4)+Y(:,2))./R0;
[T1,Y1] = ode15s(@devcelleqsfil,t2,I,[],beta1);
total=100.*(Y1(:,3)+Y1(:,4)+Y1(:,2)+...
     Y1(:,5)+Y1(:,6)+Y1(:,7)...
    +Y1(:,8)+Y1(:,9)+Y1(:,10)+Y1(:,17))./R0;
%EGF degradation
L0=3.155;
```

```
R0=4.4e5;
t1=[0 6];
t2=[0; x(11:15)];
I=zeros(17,1);
I(1)=L0;
I(2)=R0;
%Pulse phase
[T,Y] = ode15s(@devcelleqsfil,t1,I,[],beta1);
I2=Y(end,:);
inito=(I2(5)+I2(7)+I2(11)+I2(14)+I2(6)+I2(8)+...
    I2(12)+I2(15));
intc=100*(I2(5)+I2(7)+I2(11)+I2(14))/inito;
intn=100*(I2(6)+I2(8)+I2(12)+I2(15))/inito;
I2(1:4)=0;
%Chase phase
[T1,Y1] = ode15s(@devcelleqsfil,t2,I2,[],beta1);
Int=Y1(:,5)+Y1(:,6)+Y1(:,7)+Y1(:,8)+Y1(:,11)...
    +Y1(:,12)+Y1(:,14)+Y1(:,15)...
    +(Nav*Vx).*Y1(:,1);
deg=Y1(:,14)+Y1(:,15);
pdeg=100.*deg(2:end)./Int(2:end);
pdc=sum(Y1(1:end,14)./(I2(5)+I2(7)+I2(11)+I2(14)));
pdn=sum(Y1(1:end,15)./(I2(6)+I2(8)+I2(12)+I2(15)));
pdrc=sum(Y1(1:end,13)./(I2(5)+I2(7)+I2(9)+I2(13)));
pdrn=sum(Y1(1:end,16)./(I2(6)+I2(8)+I2(10)+I2(16)));
pdce=100*Y1(end,14)./(I2(5)+I2(7)+I2(11)+I2(14));
pdne=100*Y1(end,15)./(I2(6)+I2(8)+I2(12)+I2(15));
%Recycling of EGF
%Pulse phase
[T,Y] = ode15s(@devcelleqsfil,[0 15],I,[],beta1);
I2=Y(end,:);
I2(1:4)=0;
I2;
%Chase phase
[T1,Y1] = ode15s(@devcelleqsfil,[0; x(16:19)],I2,[],beta1);
```

```
Y1;
wq=I2(5)+I2(7)+I2(11)+I2(14);
we=(Y1(end,5)+Y1(end,7)+Y1(end,11)+Y1(end,14));
prce=100*((I2(5)+I2(7)+I2(11)+I2(14))-(Y1(end,5)+...
     Y1(end,7)+Y1(end,11)...
    +Y1(end,14))./(I2(5)+I2(7)+I2(11)+I2(14));
prne=100*((I2(6)+I2(8)+I2(12)+I2(15))-(Y1(end,6)+...
     Y1(end,8)+Y1(end,12)...
    +Y1(end,15))./(I2(6)+I2(8)+I2(12)+I2(15));
Int=Y1(:,5)+Y1(:,6)+Y1(:,7)+Y1(:,8)+Y1(:,11)...
    +Y1(:,12)+Y1(:,14)+Y1(:,15)...
    +(Nav*Vx).*Y1(:,1);
rec=(Nav*Vx).*Y1(:,1)+Y1(:,3)+Y1(:,4);
prec=100.*rec(2:end)./Int(2:end);
yhat=[sur;total;pdeg;prec; intc; intn; pdc;...
    pdn; pdrc; pdrn; pdce; pdne;...
    prce; prne];
end
%Model equations
function dy=devcelleqsfil(t,y,beta1)
global nh
dy=zeros(17,1);
Vx=1e-10;
Ve=1e-14;
R0=4.5e5;
Nav=6.02214E23;
kb=0.063;
kub=0.16;
kp=4.9722;
ku=beta1(1)*((y(1))^nh)/((1.0)^nh+(y(1))^nh);
kun=0.0;
ken=beta1(2);
kec=beta1(3);
kee=beta1(4);
keb=0.0085;
```

```
keub=0.66;
krc=beta1(5);
krn=beta1(6);
klc=beta1(7);
kln=beta1(8);
kdrc=beta1(9);
kdrn=beta1(10);
kdlc=beta1(11);
kdln=beta1(12);
kun=0.0;
kdub=beta1(13);
dy(1)=(-kb*y(1)*y(2)+kub*y(3)+klc*y(11)+kln*y(12))/(Nav*Vx);%Lm
dy(2)=-kb*y(1)*y(2)+kub*y(3)+krc*y(9)+krn*y(10);%Rs
dy(3)=kb*y(1)*y(2)-kub*y(3)-kec*y(3)-ku*y(3)+krc*y(7)+kun*y(4);%Cs
dy(4)=ku*y(3)-ken*y(4)-kun*y(4)+krn*y(8);%Cu
dy(5)=kec*y(3)-kee*y(5);%Cic
dy(6)=ken*y(4)-kee*y(6);%Cin
dy(7)=kee*y(5)-keub*y(7)+keb*y(9)*y(11)/(Nav*Ve)-krc*y(7);%Cec
dy(8)=kee*y(6)-keub*y(8)+keb*y(10)*y(12)/(Nav*Ve)-krn*y(8);%Cen
dy(9)=keub*y(7)-keb*y(9)*y(11)/(Nav*Ve)-krc*y(9)-kdrc*y(9);%Rec
dy(10)=keub*y(8)-keb*y(10)*y(12)/(Nav*Ve)-krn*y(10)-kdrn*y(10);%Ren
dy(11)=keub*y(7)-keb*y(9)*y(11)/(Nav*Ve)-klc*y(11)-kdlc*y(11);%Lec
dy(12)=keub*y(8)-keb*y(10)*y(12)/(Nav*Ve)-kln*y(12)-kdln*y(12);%Len
dy(13)=kdrc*y(9);%Rdc
dy(14)=kdlc*y(11);%Ldc
dy(15)=kdln*y(12);%Ldn
dy(16)=kdub*y(17);%Rdn
dy(17)=kdrn*y(10)-kdub*y(17);%Rdu
end
```

Appendix B

Identifiability analysis

Idea behind the identifiability analysis is to investigate whether or not a unique solution for the unknown parameters of interest can be found using the available data. Volume of data available represent the amount of information at hand and quality indicates reliability of the data. Hence, for the purpose of parameter estimation knowledge about the quality and quantity of any available data is important.

Fisher information matrix

The Fisher information matrix (FIM) is a measure of the amount of information available in the data about the unknown parameters (Cobelli and DiStefano, 1980; Goodwin and Payne, 1977). Thus it is closely related to the identifiability of the parameters. Optimization can be defined as the process of minimizing the objective function, which can be defined as follows (Englezos, 2001) also defined in documentation of SensSB toolbox.

$$S(p) = \sum_{i=1}^{N}((\hat{y}_i - y_i(p))^T Q_i(\hat{y}_{ij} - y_{ij}(p)) \tag{B.1}$$

The expected value of the objective function for a differential change in the parameter from the optimal is given by:

$$E[S(p+\delta p)] \cong \delta p^T \left[\sum_{i=1}^{N} \left(\frac{\partial y_i}{\partial p}\right)^T Q_i \left(\frac{\partial y_i}{\partial p}\right)\right] \delta p + \sum_{i=1}^{N} tr(C_i Q_i) \qquad \text{(B.2)}$$

where, C_i represents the measurement error covariance matrix (typically Q_i is chosen as C_i^{-1}). In our case we use a unweighted least square hence Q_i is unity matrix. In order to minimize the objective function the term in the $[\cdot]$ should be positive. This termed is called the Fisher information matrix of the estimation problem (Rehm and Reed, 2001).

$$FIM = \sum_{i=1}^{N} \left(\frac{\partial y_i}{\partial p}\right)^T Q_i \left(\frac{\partial y_i}{\partial p}\right) \qquad \text{(B.3)}$$

Here, the term $\partial y_i/\partial p$ denotes the output sensitivity of variable y_i with respect to variations in parameter p . Non-singularity of the information matrix is regarded as the necessary and sufficient condition for identifiability of the systems (Goodwin and Payne, 1977; Cobelli and DiStefano, 1980). If the sensitivity equations show a linear dependence then FIM will become singular leading to unidentifiability (Rodriguez-Fernandez *et al*., 2006; Rehm and Reed, 2001).

Covariance matrix

The Fisher information matrix is also an approximation of the inverse of the parameter estimation error covariance matrix of the best linear unbiased estimator (Rehm and Reed, 2001).

$$C = FIM^{-1} = \left[\sum_{i=1}^{N} \left(\frac{\partial y_i}{\partial p}\right)^T Q_i \left(\frac{\partial y_i}{\partial p}\right)\right]^{-1} \qquad \text{(B.4)}$$

Correlation matrix

A correlation matrix provides information about the inter-relation between the estimated parameters. In this matrix the elements are the approximate correlation coefficients between two parameters and can be estimated from the Covariance matrix as follows (Miao *et al*., 2011):

$$r_{ij} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}},\ i \neq j, \tag{B.5}$$

$$r_{ij} = 1,\ i = j \tag{B.6}$$

Here, r_{ij} is the correlation coefficient between parameter p_i and p_j. If there exists a strong correlation between two parameters estimate p_i and p_j, values of r_{ij} is close to 1, parameters p_i and p_j are said to be indistinguishable. A strong correlation between the parameters indicates that the two parameters are strongly dependent on each other and hence cannot be separately estimated (Miao *et al*., 2011). A singular FIM indicates the presence of strong correlation between the parameters where the value of r_{ij} is greater than 0.99 (Rodriguez-Fernandez *et al*., 2006). In SensSB toolbox correlation coefficient of 0.95 and above is recommended as strong and if such strong correlation exists then systems is termed as unidentifiable.

References

Acconcia, F., Sigismund, S. and Polo, S. (2009) Ubiquitin in trafficking: the network at work. *Exp Cell Res* **315**, 1610–8.

Aguilar, R. C. and Wendland, B. (2005) Endocytosis of membrane receptors: two pathways are better than one. *Proc Natl Acad Sci U S A* **102**, 2679–80.

Aster, R. C., Thurber, C. H. and Borchers, B. (2005) *Parameter estimation and inverse problems*, volume ; v. 90. Elsevier Academic Press, Amsterdam.

Bellman, R. and Åström, K. (1970) On structural identifiability. *Mathematical Biosciences* **7**(3-4), 329 – 339.

Brown, K. S. and Sethna, J. P. (2003) Statistical mechanical approaches to models with many poorly known parameters. *Physical review. E, Statistical, nonlinear, and soft matter physics* **68**, 021904.

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. *Mol Syst Biol* **5**, 239.

Cobelli, C. and DiStefano, J. (1980) Parameter and structural identifiability con-

cepts and ambiguities: a critical review and analysis. *Am J Physiol* **239**(1), R7–24.

De Donatis, A., Comito, G., Buricchi, F., Vinci, M. C., Parenti, A., Caselli, A., Camici, G., Manao, G., Ramponi, G. and Cirri, P. (2008) Proliferation versus migration in platelet-derived growth factor signaling: the key role of endocytosis. *J Biol Chem* **283**, 19948–56.

Di Guglielmo, G. M., Le Roy, C., Goodfellow, A. F. and Wrana, J. L. (2003) Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. *Nat Cell Biol* **5**(5), 410–21.

Dikic, I. (2003) Mechanisms controlling EGF receptor endocytosis and degradation. *Biochem Soc Trans* **31**, 1178–81.

Dikic, I., Wakatsuki, S. and Walters, K. J. (2009) Ubiquitin-binding domains - from structures to functions. *Nat Rev Mol Cell Biol* **10**, 659–71.

Duan, L., Miura, Y., Dimri, M., Majumder, B., Dodge, I. L., Reddi, A. L., Ghosh, A., Fernandes, N., Zhou, P., Mullane-Robinson, K., Rao, N., Donoghue, S., Rogers, R. A., Bowtell, D., Naramura, M., Gu, H., Band, V. and Band, H. (2003) Cbl-mediated ubiquitinylation is required for lysosomal sorting of epidermal growth factor receptor but is dispensable for endocytosis. *J Biol Chem* **278**, 28950–60.

Englezos, P. (2001) *Applied parameter estimation for chemical engineers*, volume 81 of *Chemical industries*. Marcel Dekker, New York.

French, A. R., Sudlow, G. P., Wiley, H. S. and Lauffenburger, D. A. (1994) Postendocytic trafficking of epidermal growth factor-receptor complexes is mediated

through saturable and specific endosomal interactions. *J Biol Chem* **269**, 15749–55.

French, A. R., Tadaki, D. K., Niyogi, S. K. and Lauffenburger, D. A. (1995) Intracellular trafficking of epidermal growth factor family ligands is directly influenced by the pH sensitivity of the receptor/ligand interaction. *J Biol Chem* **270**, 4334–40.

Gen, M. and Cheng, R. (2000) *Genetic algorithms and engineering optimization*. Wiley, New York.

Goldbeter, A. and Koshland, D E, J. (1981) An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci U S A* **78**, 6840–4.

Goodwin, G. C. and Payne, R. L. (1977) *Dynamic system identification: experiment design and data analysis*, volume v 136. Academic Press, New York.

Gschwind, A., Fischer, O. M. and Ullrich, A. (2004) The discovery of receptor tyrosine kinases: targets for cancer therapy. *Nat Rev Cancer* **4**(5), 361–70.

Gutenkunst, R. N., Waterfall, J. J., Casey, F. P., Brown, K. S., Myers, C. R. and Sethna, J. P. (2007) Universally sloppy parameter sensitivities in systems biology models. *PLoS computational biology* **3**, 1871–78.

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. *Nat Cell Biol* **5**, 461–6.

Huang, C. Y. and Ferrell, J E, J. (1996) Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci U S A* **93**, 10078–83.

Huang, F., Goh, L. K. and Sorkin, A. (2007) EGF receptor ubiquitination is not necessary for its internalization. *Proc Natl Acad Sci U S A* **104**(43), 16904–9.

Jiang, X. and Sorkin, A. (2003) Epidermal growth factor receptor internalization through clathrin-coated pits requires Cbl RING finger and proline-rich domains but not receptor polyubiquitylation. *Traffic* **4**, 529–43.

Kholodenko, B. N., Demin, O. V., Moehren, G. and Hoek, J. B. (1999) Quantification of short term signaling by the epidermal growth factor receptor. *J Biol Chem* **274**(42), 30169–81.

Lund, K. A., Opresko, L. K., Starbuck, C., Walsh, B. J. and Wiley, H. S. (1990) Quantitative analysis of the endocytic system involved in hormone-induced receptor internalization. *J Biol Chem* **265**, 15713–23.

Mayor, S. and Pagano, R. E. (2007) Pathways of clathrin-independent endocytosis. *Nat Rev Mol Cell Biol* **8**, 603–12.

Miao, H., Xia, X., Perelson, A. S. and Wu, H. (2011) On identifiability of nonlinear ode models and applications in viral dynamics. *SIAM Rev Soc Ind Appl Math* **53**(1), 3–39.

Mitchell, M. (1996) *An introduction to genetic algorithms*. MIT Press, Cambridge, Mass.

Murphy, J. E., Padilla, B. E., Hasdemir, B., Cottrell, G. S. and Bunnett, N. W. (2009) Endosomes: a legitimate platform for the signaling train. *Proc Natl Acad Sci U S A* **106**(42), 17615–22.

Oda, K., Matsuoka, Y., Funahashi, A. and Kitano, H. (2005) A comprehensive

pathway map of epidermal growth factor receptor signaling. *Mol Syst Biol* **1**, 2005 0010.

Orth, J. D., Krueger, E. W., Weller, S. G. and McNiven, M. A. (2006) A novel endocytic mechanism of epidermal growth factor receptor sequestration and internalization. *Cancer research* **66**, 3603–10.

Peschard, P. and Park, M. (2003) Escape from Cbl-mediated downregulation: a recurrent theme for oncogenic deregulation of receptor tyrosine kinases. *Cancer Cell* **3**, 519–23.

Rehm, H.-J. and Reed, G. (2001) *Biotechnology: a multi-volume comprehensive treatise*. VCH, Weinheim, second edition.

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. *Biophys J* **85**(2), 730–43.

Ringerike, T., Blystad, F. D., Levy, F. O., Madshus, I. H. and Stang, E. (2002) Cholesterol is important in control of EGF receptor kinase activity but EGF receptors are not concentrated in caveolae. *J Cell Sci* **115**, 1331–40.

Rodriguez-Fernandez, M. and Banga, J. R. (2010) SensSB: a software toolbox for the development and sensitivity analysis of systems biology models. *Bioinformatics* **26**(13), 1675–6.

Rodriguez-Fernandez, M., Egea, J. A. and Banga, J. R. (2006) Novel metaheuristic for parameter estimation in nonlinear dynamic biological systems. *BMC Bioinformatics* **7**, 483.

Schmidt-Glenewinkel, H., Vacheva, I., Hoeller, D., Dikic, I. and Eils, R. (2008) An ultrasensitive sorting mechanism for EGF receptor endocytosis. *BMC Syst Biol* **2**, 32.

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. *Dev Cell* **15**, 209–19.

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. *Proc Natl Acad Sci U S A* **102**, 2760–5.

Sorkin, A. and Goh, L. K. (2008) Endocytosis and intracellular trafficking of ErbBs. *Exp Cell Res* **314**, 3093–106.

Sorkin, A. and von Zastrow, M. (2009) Endocytosis and signalling: intertwining molecular networks. *Nat Rev Mol Cell Biol* **10**, 609–22.

Starbuck, C. and Lauffenburger, D. A. (1992) Mathematical model for the effects of epidermal growth factor receptor trafficking dynamics on fibroblast proliferation responses. *Biotechnol Prog* **8**(2), 132–43.

Tzafriri, A. R. and Edelman, E. R. (2007) Endosomal receptor kinetics determine the stability of intracellular growth factor signalling complexes. *Biochem J* **402**, 537–49.

Wang, C.-C., Cirit, M. and Haugh, J. M. (2009) PI3K-dependent cross-talk interactions converge with Ras as quantifiable inputs integrated by Erk. *Mol Syst Biol* **5**, 246.

Warren, R. A., Green, F. A. and Enns, C. A. (1997) Saturation of the endocytic pathway for the transferrin receptor does not affect the endocytosis of the epidermal growth factor receptor. *J Biol Chem* **272**, 2116–21.

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. *The EMBO journal* **21**, 303–13.

Waters, C. M., Oberg, K. C., Carpenter, G. and Overholser, K. A. (1990) Rate constants for binding, dissociation, and internalization of EGF: effect of receptor occupancy and ligand concentration. *Biochemistry* **29**(14), 3563–9.

Weissman, A. M. (2001) Themes and variations on ubiquitylation. *Nat Rev Mol Cell Biol* **2**, 169–78.

Wiley, H. S. and Cunningham, D. D. (1981) A steady state model for analyzing the cellular binding, internalization and degradation of polypeptide ligands. *Cell* **25**(2), 433–40.

Woelk, T., Sigismund, S., Penengo, L. and Polo, S. (2007) The ubiquitination code: a signalling problem. *Cell Div* **2**, 11.

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

Ye, Y. and Rape, M. (2009) Building ubiquitin chains: E2 enzymes at work. *Nat Rev Mol Cell Biol* **10**, 755–64.