Published in final edited form as:

Mol Biosyst. 2012 November ; 8(11): 2868–2882. doi:10.1039/c2mb25190f.

Integrated experimental and model-based analysis reveals the spatial aspects of EGFR activation dynamics

Harish Shankaran¹, Yi Zhang^{1,†}, William B. Chrisler², Jonathan A. Ewald^{2,‡}, H. Steven Wiley³, and Haluk Resat^{1,*}

¹Computational Biology and Bioinformatics Group, Pacific Northwest National Laboratory, Richland, WA, 99352, USA

²Cell Biology and Biochemistry Group, Pacific Northwest National Laboratory, Richland, WA, 99352, USA

³Environmental and Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, 99352, USA

Abstract

The epidermal growth factor receptor (EGFR) belongs to the ErbB family of receptor tyrosine kinases, and controls a diverse set of cellular responses relevant to development and tumorigenesis. ErbB activation is a complex process involving receptor-ligand binding, receptor dimerization, phosphorylation, and trafficking (internalization, recycling and degradation), which together dictate the spatio-temporal distribution of active receptors within the cell. The ability to predict this distribution, and elucidation of the factors regulating it, would help to establish a mechanistic link between ErbB expression levels and the cellular response. Towards this end, we constructed mathematical models to determine the contributions of receptor dimerization and phosphorylation to EGFR activation, and to examine the dependence of these processes on sub-cellular location. We collected experimental datasets for EGFR activation dynamics in human mammary epithelial cells, with the specific goal of model parameterization, and used the data to estimate parameters for several alternate models. Model-based analysis indicated that: 1) signal termination via receptor dephosphorylation in late endosomes, prior to degradation, is an important component of the response, 2) less than 40% of the receptors in the cell are phosphorylated at any given time, even at saturating ligand doses, and 3) receptor phosphorylation kinetics at the cell surface and early endosomes are comparable. We validated the last finding by measuring the EGFR dephosphorylation rates at various times following ligand addition both in whole cells and in endosomes using ELISAs and fluorescent imaging. Overall, our results provide important information on how EGFR phosphorylation levels are regulated within cells. This study demonstrates that an iterative cycle of experiments and modeling can be used to gain mechanistic insight regarding complex cell signaling networks.

Keywords

ErbB/HER/RTK signaling; mathematical modeling; signal transduction; phosphatase activity; parameter estimation

*Corresponding author: Haluk Resat, Computational Biology and Bioinformatics Group, Pacific Northwest National Laboratory, PO Box 999, MS J4-33, Richland, WA 99352, USA, haluk.resat@pnnl.gov, Phone: (509) 372-6340.

†Current address: Department of Medicine, University of Tennessee Health Science Center, Memphis, TN, 38103, USA

‡Current address: Department of Urology, University of Wisconsin School of Medicine and Public Health, Madison, WI 53792

Description: Location-dependent EGFR activation analysis revealed signal termination in late endosomes, and comparable dephosphorylation kinetics at the cell surface and endosomes.

INTRODUCTION

The human epidermal growth factor receptor (HER; also known as ErbB) family of receptor tyrosine kinases mediates a diverse set of physiological responses including cell proliferation, survival and motility.^{1,2} The importance of this receptor system to normal cell physiology and the involvement of ErbB overexpression in cancer pathophysiology²⁻⁹ have led to extensive studies on the mechanisms and downstream consequences of ErbB activation. The variety of ErbB ligands, differential trafficking of the individual receptors, formation of a range of dimers between receptors, and distinct signaling potencies of the various dimers are some of the aspects that contribute to the complexity of ErbB signaling.^{1,10-13} This complexity, and the availability of quantitative information, accumulated over the years, has led to the extensive use of mathematical models to understand the ErbB system.^{10,11,14}

ErbB activation is a multi-step process wherein ligand addition induces the formation of receptor-ligand complexes that dimerize to yield receptor homo- and hetero-dimers among family members.¹⁵⁻¹⁷ The dimerization pattern in a cell depends upon the receptor expression level as well as receptor trafficking.^{11,13,18} Dimerization promotes trans-phosphorylation of the receptor cytoplasmic tails via the tyrosine kinase activities of the partners,^{16,19} thereby leading to the generation of active receptors that are capable of signaling. In the field, the usual practice is to treat the phosphorylated receptors as activated receptors. Thus, in the remainder of the paper, we use the terms active and phosphorylated receptors in an interchangeable fashion. There is considerable evidence to suggest that the dimerization pattern is an important determinant of the consequences of ErbB signaling.^{20,21} Therefore, the ability to quantitatively predict the abundances of receptor dimers, and the effect of dimerization on receptor activation would help to relate cellular responses to receptor expression levels.

Receptor phosphorylation is a reversible dynamic process that reflects the balance between kinase and phosphatase kinetics. Since the receptors may have access to potentially different complements of adaptor molecules and signaling components, receptor phosphorylation levels can vary among cellular compartments.²²⁻²⁵ There is considerable evidence that phosphatases at different cellular locations may be differentially involved in tonic suppression of EGFR activation in the absence of ligand, and regulation of ligand-induced EGFR phosphorylation levels, and in signal termination via removal of the phosphate group.²⁶⁻²⁹ One of the consequences of receptor trafficking is that the intracellular location of receptors shifts over time. For ErbB receptors, ligand induced activation causes increased internalization and receptor degradation.¹³ Therefore, following ligand addition, the phosphorylated receptors are expected to shift from the cell surface to the internal compartments as time progresses. One of our aims was to quantify this time-dependent shift in the active receptor pool and to investigate whether receptor dephosphorylation kinetics varies among cellular compartments.

Mathematical models of varying degrees of complexity have been previously constructed to describe ErbB phosphorylation, trafficking and downstream signaling.^{11,14} Rate parameters for these models are usually set based on the literature for related systems, or are obtained by fitting to a limited set of experimental data. Parameters for fast processes such as (de)phosphorylation and association kinetics can be especially difficult to obtain. Thus, even though the rate constants for receptor association and activation for the ErbB system have been reported in the literature, there is considerable uncertainty surrounding these parameters and their location dependence. Here, we address this gap by using a model-based analysis to extract EGFR dimerization and activation parameters from experimental data

collected in a human mammary epithelial (HME) cell line that expresses EGFR (ErbB1) at much higher levels compared to the other ErbB molecules. We constructed a series of alternate models for EGFR activation at several complexity levels, and collected the appropriate experimental data to aid parameter estimation. The objectives of our study were to first determine the simplest mathematical model (parsimonious) that is consistent with the experimental data, and then use it to investigate if EGFR dimerization and phosphorylation dynamics varies with cellular location in HME cells.

We found that a model containing three cellular compartments (cell surface, early endosomes, and late stage multivesicular bodies), with terminal dephosphorylation occurring in a distinct compartment prior to receptor degradation, was necessary to explain the experimental data. Our results indicated that, even at saturating doses of ligand, a relatively small percentage (< 40%) of the receptors in the cell was phosphorylated at any given time. Further, our model-based analysis indicated that EGFR dephosphorylation rates at the cell surface and in the early endosomal compartments were comparable. We validated this prediction in ELISA and fluorescent imaging experiments. Our results for these two independent but complementary measurements confirmed that the activities of the phosphatases involved in dampening EGFR phosphorylation are comparable across cellular locations.

EXPERIMENTAL MATERIALS AND METHODS

Cell culture and treatment

The HME cell used in this study was originally provided by Martha Stampfer (Lawrence Berkeley Laboratory, Berkeley, CA) as cell line 184A1-1. It expresses approximately 200,000 molecules of EGFR/ErbB1/HER1.^{18,30} Cells were maintained in DFCI-1 medium supplemented with 12.5 ng/ml EGF (PeproTech, Rocky Hill, NJ) as described previously.³¹ At about 70–80% cell confluency, DFCI-1 medium was replaced with bicarbonate-free DFHB minimal medium lacking EGF, plus 0.1% bovine serum albumin. Cells were then brought to quiescence for 12–18 hours before treatment. To activate the EGFR, EGF at a known concentration was added to the culture medium, and cells were incubated at 37°C for fixed amounts of time from 5 to 120 min.

Phosphorylated receptor levels in the internal compartments were determined using an acid-stripping protocol, which selectively dephosphorylates cell surface receptors without altering the phosphorylation of internalized receptors.³² Following acid stripping, cells were washed 3X with ice cold PBS and incubated at room temperature for one minute to allow surface receptor dephosphorylation. After another round of cold PBS washing, cells were solubilized with ice cold lysis buffer (1% NP-40, 20 mM pH 8.0 Tris buffer, 137 mM NaCl, 10% glycerol, 2 mM EDTA, supplemented with 1 mM heat activated sodium orthovanadate and 1% protease inhibitor cocktail III; Calbiochem, La Jolla, CA) for 20 min. Un-stripped cells were also lysed using the same lysis protocol. Cell lysates were collected with a scraper. Lysates were centrifuged at 13,000 rpm for 10 min at 4°C, and the supernatants were transferred into fresh microtubes. Obtained cell lysates were either analyzed immediately or stored at a –80°C freezer until needed.

Receptor mass and phosphorylation measurements

ELISA assays to quantify the receptor mass and phosphorylation levels were performed using the R&D DuoSet IC ELISA kits (R&D Systems Inc., Minneapolis, MN). We have previously published EGFR receptor mass and phosphorylation levels in response to 100 ng/ml EGF – a saturating ligand concentration.³³ Our current analysis is based on recent data where we examine the response to sub-saturating ligand doses. Two types of ELISA data were collected as a function of time following ligand addition:

1. *The EGFR receptor mass (mR_i)* was quantified in total cell lysates using capture and probe antibodies specific to EGFR following the addition of 12 ng/ml EGF. These experiments measured the receptor degradation over time.
2. *The extent of EGFR tyrosine phosphorylation (pR_i)* was assayed by pulling total cell lysates down with antibodies specific to EGFR, and subsequently probing with a polyclonal phospho-tyrosine antibody. These measurements were done for various concentrations of EGF from 0.6 to 30 ng/ml, and measured the receptor activation levels as a function of time and ligand dose.
3. When measurement 2 above was repeated following ligand stripping, it enabled us to quantify *the extent of phosphorylation (pR_i) in the internal compartments*. Internal phosphorylation measurements were obtained as a detailed time course following the addition of 12 ng/ml EGF.

The ELISA results were normalized based on the total protein present in the cell lysate (measured using the Bicinchoninic Acid protein quantitation kit, Sigma, St. Louis, MO), and were expressed in units of picograms per microgram of total lysate protein.

Measurement of EGFR dephosphorylation rates

We measured EGFR dephosphorylation using two approaches, imaging and ELISA as follows:

Imaging—HME cells were plated onto sterile cover slips and were grown to 50% confluency. Cells were brought to quiescence by removing EGF for 12–18 hours before treatment, and were then stimulated for 15 minutes with 25 ng/ml Alexa 488-conjugated EGF (Invitrogen, Carlsbad, CA). After EGF treatment, cover slips were transferred to a 37°C PBS solution containing 1 μ M of the EGFR tyrosine kinase inhibitor AG1478 to initiate dephosphorylation. Since the use of an affinity purified site-specific antibody typically results in better signal to noise ratios in fluorescent imaging, we preferred to use an antibody that binds specifically to phospho-tyrosine 1173 (pY1173) of the EGFR cytoplasmic tail. At specific time points following AG1478 treatment, cells were fixed using ice cold 3.6% paraformaldehyde and stained with Alexa 647-conjugated sheep antibody against pY1173 of the EGFR cytoplasmic tail.²² Red (pY1173) and green (EGF) channel images for chosen image fields were captured using a Leica DMIRE confocal microscope using an HCX PL APO CS 100.0x1.40 NA Oil objective. Green channel images were used to create masks of cells and endosomes for determining phosphorylation levels in these respective objects, and to enable normalization of phosphorylation levels. Analysis of the collected image data was performed using MATLAB (Mathworks, Natick, MA) functions to automatically create masks of cells and endosomes based on the green channel intensities. For each of the cells/endosomes in a given image we computed a normalized phosphorylation level by dividing the red channel intensity by the green channel intensity. An average phosphorylation level was determined for each time point following AG1478 addition by averaging the normalized phosphorylation level over multiple cells/endosomes.

ELISA—For ELISA measurements of receptor dephosphorylation, cells were stimulated for 5, 10 or 15 min with 25 ng/ml EGF following which 1 μ M AG1478 was added for specified amounts of time. Cells were then lysed in the presence of phosphatase inhibitors and the pY1173 antibody was used to determine the phosphorylation level, which was normalized based on the total protein present in the lysate. We note that, to enable comparison, dephosphorylation ELISA experiments were performed using the same antibody that was used in the imaging experiments.

Analysis—We plotted the average phosphorylation level as a function of time following AG1478 addition and fitted an exponential decay function of the form $Y = Y_0 \exp(-k_{dephos} t) + C$ to the experimental results to determine the dephosphorylation rates k_{dephos} for cells and/or endosomes. The exponential decay function fit was performed using the *nlinfit* non-linear fitting algorithm of MATLAB.

MATHEMATICAL MODELS FOR EGFR ACTIVATION

General considerations

Mathematical models can be powerful tools to analyze the dynamics of signaling networks.^{11,14,34–38} A critical aspect of model development is the chosen level of complexity: The model should capture the fundamental biophysical/biochemical reactions in the system, and should be detailed enough to address the mechanistic question being asked. At the same time, the availability of model parameters, and the ability to identify parameters from the available experimental data should be taken into account. Models that are too simple may be devoid of the necessary details for addressing relevant questions, while too complex models may incorporate many unknown parameters. The utility of models at both of these extremes is questionable. Ideally, model scope and complexity should be addressed by integrating model construction and experimental design at the outset, and refined through a systematic experiment-modeling iteration cycle.

Our mathematical models (Fig. 1) for the EGFR system combined the essential biochemical reactions with receptor trafficking to predict EGFR receptor mass, dimerization and phosphorylation dynamics at the cell surface and the interior. Our general approach was to keep the models as simple as possible to avoid over-fitting the limited experimental data during parameter estimation, while capturing the important steps of the processes that we are investigating – the dynamics of EGFR phosphorylation and its dependence on spatial location within the cell.

Modeling receptor dimerization and phosphorylation

Our models include reaction steps explicitly if their rate constants have been previously measured in the used cell line. We use lumped parameters or scaling factors where detailed kinetic information is unavailable to avoid the use of too many undetectable parameters. As we have previously shown, endogenous receptor expression levels of ErbB2 and ErbB3 are much lower than that of EGFR in the parental HME cells used here, and EGFR phosphorylation occurs mainly through the formation of EGFR homodimers in these cells.^{30,33} Thus, only the EGFR interactions were considered in this study.

The biochemical reactions in the model (box at the top of Fig. 1) involve the reversible binding of ligand, L to free receptor monomers, RF to yield bound receptors, RB with forward rate k_{on} and reverse rate k_{off} . Bound receptors can form dimers, RD with forward rate k_c and reverse rate k_d . These biochemical reactions are allowed to occur at the cell surface and in early endosomes with reaction fluxes being modeled using mass action kinetics. For the sake of simplicity, we only simulate ligand-induced dimerization in the model. Although pre-formed receptor dimers may exist in the absence of ligand stimulation, we assume that such basal dimerization occurs at a low, insignificant level. This is consistent with the ideas that the conformational changes induced by ligand binding greatly enhance dimer stability and EGFR kinase activity.^{15,16}

Dimerization facilitates the trans-phosphorylation of receptor cytoplasmic tails at a number of tyrosine sites due to the kinase activities of the partners in the dimer. Since the association/dissociation and phosphorylation/dephosphorylation processes are expected to be rapid (sub-minute) compared to the time scale of experimental sampling (typically

minutes-to-hours), the lack of information prohibits reliable parameterization of the individual steps of these successive reactions. We therefore treat receptor activation as a lumped process: Our models explicitly keep track of the receptor dimers, while contributions of the dimers to the receptor phosphorylation signal is accounted for with a multiplicative phosphorylation factor (pf) to convert the number of receptor dimers to the phosphorylation levels measured in the ELISA experiments (Fig. 1). The pf factor can be thought of as a lumped phosphorylation affinity that combines the characteristics of all possible tyrosine sites in a dimer, and it represents how effectively a particular receptor gets activated and contributes to the measured phosphorylation signal. We therefore refer to the pf as the *phosphorylation efficiency factor*. We assume that the pf values are time-invariant biophysical constants that do not depend upon the EGF concentration because they account for molecular scale internal processes occurring in an already formed dimer. Since the pf value incorporates the phosphorylation affinity of the receptors, it is expected to be a function of the rate constants for trans-phosphorylation mediated by the EGFR kinase activity, and dephosphorylation mediated by phosphatases. Irrespective of whether a receptor is at the cell surface or in the EE, its kinase domain would face the cytoplasm and therefore the receptors would be exposed to similar pH. Assuming that local ATP availability is not limiting and since trans-phosphorylation in a dimer is an intrinsic biophysical property of the receptors, kinase activities are likely to remain the same across the surface and EE compartments. However, because the local availability of phosphatases can result in a location-dependent dephosphorylation rate,^{26,27,29} the pf values may vary between cell substructures.

To analyze the receptor mass ELISA data, we used a mass scaling factor mf to convert the total number of receptor molecules in the cell to the receptor mass ELISA reading. Since the parental HME cells express 2×10^5 EGFR,¹⁸ given the ELISA receptor mass reading in the absence of ligand addition, mR_0 , mf can be calculated as $mR_0/2 \times 10^5$ and hence it is a known constant. Note that analogous conversion factor for relating the number of phosphorylated receptors to the pR ELISA measurements is implicitly included as part of the pf factor.

Alternate models for receptor trafficking

With regard to receptor trafficking, there are several model formulations in the literature that characterize the dynamics of the ErbB system using two- (e.g., Refs. 33,39–41) or three-compartment (e.g., Refs. 32,42). Two-compartment mathematical models contain the cell surface and early endosomal (EE) compartments; in this formulation signal termination is achieved solely through receptor degradation. Three compartment models include the additional late endosomal (LE) compartment, and it is assumed that signal termination occurs via irreversible receptor dephosphorylation in the LE compartment prior to receptor degradation. In this study we examine if two compartments are enough, or if more compartments are necessary to model EGFR activation dynamics. The models are described in detail below.

In the two-compartment models, receptors enter the EE compartment via internalization (Fig. 1, left panel). The EE exit flux is captured via an exit rate constant k_x , which is partitioned into the recycling and degradation fluxes using a recycling fraction parameter f . The model also includes a receptor synthesis flux term Q_R which defines the rate at which free unbound receptors are added to the plasma membrane. The activity state of the receptor is known to play an important role in receptor trafficking.^{14,43,44} Since receptor dimers are the only active/phosphorylated species in our model, all monomers (unbound and ligand-bound) are assumed to have identical trafficking rates, while distinct trafficking rates are assigned to the dimers. Monomers and dimers are respectively internalized at rates $k_{e,m}$ and $k_{e,d}$, have endosomal exit rates $k_{x,m}$ and $k_{x,d}$, and have recycling fractions f_m and f_d . Free

ligand molecules are assumed to be recycled and degraded at the constitutive rates used for receptor monomers. For any species, the recycling rate k_r and the degradation rate k_d can be calculated as $k_x f$ and $k_d(1-f)$ respectively with appropriate values being used for k_x and f .

The more complex three-compartment model was built to allow for the finite degradation time of receptors following their exit from the EE. After exiting the EE, receptors are destined for degradation and they become part of pre-lysosomal compartments called multi-vesicular bodies (MVBs) wherein multiple endosomes are packaged together.^{14,44} These MVBs are also referred to as late endosomes. Receptors in MVBs are dephosphorylated by endoplasmic reticulum (ER) localized tyrosine phosphatases by virtue of membrane contacts between portions of the MVB and the ER⁴⁵ thereby inducing signal termination. Receptors are delivered from the MVBs to lysosomes where protein degradation occurs. Thus, when a phosphorylated receptor exits the EE, there is an intermediate step wherein receptor phosphorylation may be lost but the receptor itself stays intact. This process is built into the three-compartment model using an idealized LE compartment (Fig. 1, right panel). Receptor monomers and dimers enter the LE with rate $k_{l,m}$ and $k_{l,d}$ and exit the LE via degradation at rate $k_{d,m}$ and $k_{d,d}$ respectively. We assume, for the sake of simplicity, that receptors in the LE get dephosphorylated rapidly. Hence they do not contribute to receptor phosphorylation measurements, but these proteins contribute to receptor mass measurements. This assumption makes knowing the receptor dimerization levels in the LE unnecessary, and therefore biochemical reactions in the LE were omitted from our model.

Parameter values

One of our objectives was to estimate the unknown rate parameters by fitting the models to EGFR phosphorylation and receptor mass ELISA data. To facilitate parameter estimation, and to better ground the model in reality, we employ previously determined values for rate constants if those rates were measured directly in our cell line (Table 1). Internalization, recycling and degradation rates in the used HME cells have been previously reported for free and ligand-bound EGFR.^{18,41} The recycling and degradation rates were calculated by combining measurements of recycling and degradation fluxes with measurements of the number of internal molecules at steady-state. The analysis in these studies assumed that the entire recycling and degradation flux originated from the EEs. Thus, because of their equivalency, the experimentally determined parameter values^{18,41} (values for k_x and f) can be used as is in our two-compartment models. However, the internal receptor pool is partitioned between EEs and LEs in the three compartment models. In order to use the flux measurements to constrain the three compartment models we define the ratio of the entry to the exit rate into the LE as the parameter $\delta = k_l/k_d$. It should be noted that, at steady-state, δ is also the ratio of the number of molecules in the LE to that in the EE. We use values δ_m and δ_d respectively for monomers and dimers. We can express the individual rate constants k_r , k_l and k_d in the three compartment model as a function of the known trafficking parameters (k_x and f) and the unknown δ values (see the expressions in the right panel of Fig. 1). These relationships are derived by requiring that, for a given set of k_x and f values, the recycling and degradation fluxes in the three compartment model match those for a two-compartment model. Although we have distinct parameters δ_m and δ_d for monomers and dimers, these parameters can be constrained further by requiring that once in the LE, monomers and dimers should be degraded at the same rate. By setting $k_{d,m} = k_{d,d}$ we get the relationship $1/\delta_d = \{[k_{x,m}(1-f_m) \times (1/\delta_m + 1)]/[k_{x,d}(1-f_d)]\} - 1$ (see the expressions for $k_{d,m}$ and $k_{d,d}$ in the right panel of Fig. 1). Thus, given that δ_d can be calculated from δ_m , the three compartment model has only one extra parameter compared to the equivalent two-compartment version.

The unknown parameters in the two-compartment model are those related to dimer dissociation (k_u) and phosphorylation efficiency (pf). As discussed above, the three

compartment models have an additional unknown parameter, δ_m . In order to examine the spatial dependence of dimerization and phosphorylation, we constructed different versions of the two and three compartment models by allowing k_u and/or p_f to be different in the cell surface and EE compartments (Fig. 1).

Simulations of the mathematical models

The biochemical reactions and receptor trafficking processes described above together determine the dynamics of ligand and receptor species in the model. Governing equations describing the rates of changes in the levels of species are as follows:

Cell Surface

$$\begin{aligned} dRF_s/dt &= -k_{on,s}RF_sL_s + k_{off,s}RB_s - k_{e,m}RF_s + k_{r,m}RF_e + Q_R \\ dL_s/dt &= [-k_{on,s}RF_sL_s + k_{off,s}RB_s + k_{r,m}L_e(N_{av}V_e)]/(N_{av}V_{mc}) \\ dRB_s/dt &= k_{on,s}RF_sL_s - k_{off,s}RB_s - k_cRB_s^2 + 2k_{u,s}RD_s - k_{e,m}RB_s + k_{r,m}RB_e \\ dRD_s/dt &= (1/2)k_cRB_s^2 - k_{u,s}RD_s - k_{e,d}RD_s + k_{r,d}RD_e \end{aligned}$$

Early Endosome

$$\begin{aligned} dRF_e/dt &= -k_{on,e}RF_eL_e + k_{off,e}RB_e + k_{e,m}RF_s - (k_{r,m} + k_{l,m})RF_e \\ dL_e/dt &= (-k_{on,e}RF_eL_e + k_{off,e}RB_e)/(N_{av}V_e) - (k_{r,m} + k_{l,m})L_e \\ dRB_e/dt &= k_{on,e}RF_eL_e - k_{off,e}RB_e + k_{e,m}RB_s - k_cRB_e^2 + 2k_{u,e}RD_e - (k_{r,m} + k_{l,m})RB_e \\ dRD_e/dt &= (1/2)k_cRB_e^2 - k_{u,e}RD_e + k_{e,d}RD_s - (k_{r,d} + k_{l,d})RD_e \end{aligned}$$

Late Endosome

$$\begin{aligned} dRF_l/dt &= k_{l,m}RF_e - k_{d,m}RF_l \\ dL_l/dt &= k_{l,m}L_e - k_{d,l}L_l \\ dRB_l/dt &= k_{l,m}RB_e - k_{d,d}RB_l \\ dRD_l/dt &= k_{l,d}RD_e - k_{d,d}RD_l \end{aligned}$$

Here L , RF , RB and RD respectively denote ligand, free receptor monomers, ligand-bound receptor monomers, and receptor dimers. Subscript labels “s”, “e” and “l” are used for species concentrations at the cell surface, early endosome (EE) and late endosome (LE), respectively. Same letters are used as suffixes (letter after the “,”) for the biochemical reaction rate constants in these compartments. Suffixes “m” and “d” are used in trafficking rate constants to denote whether they apply to monomers (“m”) or dimers (“d”). We note that the above equations describe the 3 compartment models and that the LE compartment is absent in the 2 compartment models. For the two compartment models, the exit terms from the early endosome are suitably modified by replacing rates of the type $(k_r + k_l)$ with k_x . Relationships between k_r , k_b , k_d and the parameters k_x , f and δ can be found in Figure 1.

The initial conditions for the system can be determined by solving the system steady-state equations in the absence of ligand. This yields $RF_{s0} = k_{x,m}R_0/(k_{l,m} + k_{x,m})$; $RF_{e0} = (R_0 - RF_{s0})/(1 + \delta_m)$, $RF_{l0} = RF_{e0}\delta_m$ and $Q_R = R_0k_{l,m}k_{x,m}(1 - f_m)/(k_{l,m} + k_{x,m})$ where R_0 is the total receptor expression in the cell prior to ligand addition. The same equations apply to the two compartment models with the following modifications: $RF_{e0} = R_0 - RF_{s0}$; $RF_{l0} = 0$.

For any model, given the values of the fixed and unknown parameters (Table 1), we first computed the system steady-state in the absence of ligand to calculate the initial levels of receptors at the cell surface and in the EE and LE compartments, as well as the receptor synthesis flux Q_R as a function of the total receptor expression level and the receptor trafficking rates. We then simulated the system following the addition of a fixed amount of ligand to the extracellular space to predict the total receptor numbers and receptor dimer levels within each compartment. Model governing equations listed above were integrated using the MATLAB ordinary differential equation solver *ode15s*. Computed receptor numbers and dimer levels were respectively multiplied with the factors mf and pf to obtain the model predictions for total receptor mass (mR_i) and total (pR_i) and internal (pR_i) phosphorylation levels (see top box in Figure 1). These predictions were then compared with the experimental data to compute the error residuals.

Parameter estimation

The unknown parameters for each model were determined by simultaneously fitting the model to all of the experimental data. We computed scaled residual vectors (residual = model prediction – experimental data) for pR and mR predictions by dividing each of these residuals by the maximum values measured in the phosphorylation and receptor mass measurements, respectively. The use of scaled values apportions roughly equal importance to the receptor mass and receptor phosphorylation measurements during parameter estimation. We then concatenated the scaled pR and mR residual vectors and used *lsqnonlin* – the MATLAB nonlinear least squares regression function – to determine the optimal parameter values. For each model, we performed 100 optimization runs that were started with different, randomly chosen initial parameter values. Random initial values for the fit parameters were chosen from a uniform distribution within the upper and lower parameter bounds (Table 1). For each of the 100 optimized solutions, we computed the root-mean-squared-error (RMSE) between the model and the experiment based on the weighted residual vector, and retained those solutions with $RMSE \leq 2\%$ of the best RMSE. Each of these selected solutions offered a comparable fit to the experimental data, and upper and lower bounds for the parameter estimates were determined based on these solutions.

The parameter estimation described above was performed for each of the models with the fixed parameters being held constant at the values listed in Table 1. For a subset of the models, we examined the effect of relaxing the constraints on the fixed parameters. We sampled the biochemical parameters that were previously fixed at constant values ($k_{on,s}$, $k_{off,s}$, $k_{on,e}$, $k_{off,e}$, k_c , $k_{e,m}$, $k_{e,d}$, $k_{x,m}$, $k_{x,d}$, f_m and f_d ; see Table 1) from log-normal distributions with means given by the values in Table 1 and coefficients of variation, CV (=standard deviation/mean) of 20%. To avoid outliers at the tails of the distribution, we selected only those values that were within 2 standard deviations of the mean (i.e., parameters were allowed to vary by $\pm 40\%$ of the mean value). We note that for cases where the parameter standard deviations are listed in the literature, the CV values for the parameters are typically within 10–20%.^{18,46} We performed 100 optimizations with the fixed parameters being drawn from these log-normal distributions. The initial guesses for the unknown parameters were the previously determined optimum values (obtained with fixed parameters held at values in Table 1). As before, for each of these 100 optimizations, we minimized the scaled residual vector by allowing only the unknown parameters to vary during the optimization. Following optimization, we selected the best solution (solution with the lowest RMSE), and examined whether allowing flexibility in the fixed parameters in this manner led to an improvement in the match to the experimental data.

Model comparison

An objective model comparison analysis has to account for differences in the numbers of fitted parameters. We accomplished this by employing the Akaike information criterion (AIC).⁴⁷ The AIC value examines whether an increase in model complexity results in a sufficient increase in prediction accuracy to warrant selection of the complex model. We calculated the AIC as, $AIC = 2N_{par} - 2L$, where N_{par} is the number of model parameters. L is the log-likelihood function and can be calculated as $L = -(N_{pts}/2) \times [\log(2\pi \text{MSE}) + 1]$ where N_{pts} = total number of experimental data points (54 in our case) and MSE is the mean-squared-error between the model and the experiment. In comparing different models, the one with the lowest AIC value, AIC_{min} , is the preferred choice. The relative strength of evidence for model i with AIC value AIC_i , compared to the best model can be quantified as $\exp(-\Delta_i/2)$ where $\Delta_i = AIC_i - AIC_{min}$. For example, models with Δ_i values of 2 and 10, are roughly speaking, 0.368-times and 0.007-times as likely as the best model to explain the data. As a rule of thumb, models with $\Delta_i < 2$, are said to have substantial support, while those with $\Delta_i > 10$, are said to have essentially no support.⁴⁸ Here, we use these criteria to evaluate the various models.

RESULTS

Model-based analysis of EGFR dynamics

Our objective is to rigorously examine alternate hypotheses regarding the spatial dependence of EGFR dimerization and phosphorylation by systematically comparing various models with experimental data on EGFR activation. This analysis required collecting experimental datasets that provide the information necessary to determine the receptor dimerization affinity and phosphorylation efficiencies; and for examining the spatial dependence of these processes. Specifically, we: 1) measured EGFR phosphorylation levels as a function of time after stimulating HME cells with EGF at doses ranging from 0.6 to 30 ng/ml (0.1 nM to 5 nM) – this data provided the information about receptor dimerization and phosphorylation trends at multiple activation levels; 2) obtained detailed time courses for total and internal EGFR phosphorylation in response to 12 ng/ml (2 nM) EGF - measurement of the relative contribution of internal receptors to receptor phosphorylation provided the benchmark data regarding the spatial dependence of the phosphorylation affinity; and 3) measured EGFR receptor mass as a function of time at the 12 ng/ml EGF concentration - measurement of receptor degradation made it possible to examine the disparity in the time-dependent decay in receptor phosphorylation and in receptor mass. The entire experimental dataset is presented in Figures 2A and 2B (points in the figures).

We analyzed this dataset using various mathematical models that were schematically illustrated in Figure 1 and described in detail in the Methods section. Our approach is to fix the model parameters at previously reported values (see Table 1) if the values were determined via direct experimentation, while estimating the remaining parameters from our experimental data through parameter optimization.

Two compartment models—As the simpler case, we first examined the mathematical models involving two compartments (2C models) – the cell surface and early endosomes (Fig. 1, left panel). Such models have been widely used for modeling and analysis of the EGFR system (e.g.,^{33,39–41}). As discussed in the Methods section, the unknown parameters in the 2C models are the dimer dissociation rates (k_d) and the phosphorylation efficiency factors (pf) at the cell surface and in the endosomes.

We wanted to test if model-based analysis can be used to uniquely estimate the k_d and pf values given that these parameters jointly determine the measured receptor phosphorylation

levels in our experiments. Secondly, we wanted to test if the data supports differences in the k_u and/or pf values between the cell surface and the EE compartments. To this end, we examined various versions of the 2C model where the k_u and pf values were either set to be equal at the cell surface and EEs or allowed to vary between these compartments (Fig. 1, left panel). The fit statistics for the four possible combinations are presented in Table 2, and Figure 2C visually illustrates how well these models fit a subset of the experimental data. As seen, all 2C models (dotted lines in Fig. 2C) perform poorly in fitting the data and their poor performance is also reflected in the high AIC scores for these models (Table 2). The main reason for the poor fit is that the receptor phosphorylation levels start to decay beyond ~10 min at a rate that is faster than the rate of decline in the receptor mass (compare slopes of the decaying portions of pR_t and mR_t in Fig. 2B). At first glance, it may appear that an increased dimer dissociation rate ($k_{u,e} > k_{u,s}$) or decreased phosphorylation factor ($pf_e < pf_s$) in the interior compartment compared to the cell surface can lead to an apparent decay in the phosphorylation level over time through a mechanism that is independent of receptor degradation. However, an increase in $k_{u,e}$ leads to increased receptor recycling, which makes the response as a whole more sluggish and hence pR_t does not show the desired decay (Fig. 2C, green dotted line). A decrease in pf_e leads to earlier onset of decay in pR_t than the one seen experimentally (Fig. 2C, pink and blue dotted lines). In this modeling scenario the excess decay in pR_t over that of mR_t is achieved through receptor internalization, which is too rapid to account for the gradual decay in measured phosphorylation levels beyond 10 min. Overall, none of the 2C models accurately capture the data even if k_u and pf are allowed to take on different values in the two compartments.

Three compartment models—The 2C models assume that receptor degradation begins as soon as a non-recycling receptor exits the EE. In reality, receptors destined for degradation become part of multivesicular bodies (MVBs) before they are eventually routed to the lysosome where degradation occurs. Receptor dephosphorylation in MVBs occurs through the action of phosphatases localized to the endoplasmic reticulum.^{26,45} Thus, following exit from the EE, receptors can stay intact for some time but they quickly lose their phosphorylation signal. To account for this process we constructed a three-compartment model (3C model; Fig. 1, right panel) containing an idealized “late endosome (LE)” which, simply put, is a site for accumulation of dephosphorylated receptors. Specifically, we assume that the receptors in the LE do not contribute to receptor phosphorylation measurements but contribute to receptor mass measurements.

As was the case for the 2C models we examined various versions of the 3C models where k_u and/or pf were allowed to vary between the cell surface and the EE (Fig. 1). The 3C models fared significantly better than the 2C models (Fig. 2C solid lines and Table 2) in fitting the experimental data, even after accounting for the one extra fit parameter (see below). All 3C models yielded comparable fits to the data (Fig. 2C), and their predictions for receptor dimerization are similar (Fig. 2D). At saturating EGF concentrations all four models predict that at most 40% of the total cellular EGFR is dimerized at any instant of time (Fig. 2D). In contrast, the poorly performing 2C models predicted peak values of 30–95% for receptor dimerization (Fig. 2D), with dimerization predictions being quite disparate between the models. We note that predictions of the 3C models are in line with our previously reported experimental measurements for the fraction of total cellular EGFR that is phosphorylated.³³

Model comparison using the Akaike Information Criterion

Since the tested models differed in terms of the number of fit parameters (i.e., degrees of freedom) and their ability to fit the data, we compared their performance by computing the Akaike Information Criterion (AIC; see Methods section).⁴⁷ The 3C model where k_u and pf values were equal for cell surface and EE compartments (Model 5) had the lowest AIC

(Table 2). The difference in AIC between the 2C models and the best model was at least 10, thereby enabling us to reject the 2C models (see Methods). On the other hand, the difference between AIC values of the other 3C models and that of the best model was within 2. Thus, given that the AIC is a rough model selection measure, all of the 3C models can be acceptable representations of the system.⁴⁸

We examined whether allowing greater flexibility in the fixed model parameters would render the 2C models comparable in performance to the 3C models. To this end, we selected the two best 2C models – models 3 and 4 – and performed 100 additional optimizations for each, with the fixed parameters being sampled from log-normal distributions with means given by the values in Table 1, and coefficients of variation of 20% (see Methods). As expected, the additional flexibility resulted in an improvement in the fit, with the AIC values for models 3 and 4 improving to 6.50 (from 10.95) and 7.87 (from 12.34) respectively. However, these AIC values are still significantly greater than the AIC of the best 3C model by 7.1 and 8.5, respectively (Table 2). Hence, even in this scenario the 2C models are less than 0.03-times as likely as the best 3C model to explain the data (see Methods). It should be noted that the AIC for the 3C models would also improve if the fixed parameters were allowed a similar flexibility during the optimization. Thus, we conclude that the superiority of the 3C models is a robust finding that is not contingent upon the particular values chosen for the fixed parameters.

To further investigate the plausibility of the various models, we examined the parameter estimates for the 3C models (Table 3). We found that three of the 3C models (Models 5–7) yield comparable parameter estimates and suggest that pf values at the surface and EE are comparable. Interestingly, even in the case of Model 7 where the pf values were allowed to vary, parameter estimation resulted in comparable pf values at the surface and EEs (Table 3). Obtained parameters for Model 8, the model with the most degrees of freedom, differ from the other three 3C models: estimated dimer dissociation rate k_u at the surface is about 70 times faster than k_u at the EE. Since the dimer dissociation rate is a function of the biophysical interaction between the receptors, and since molecular crowding effects are accounted for in the models by explicitly including the compartment volumes (Table 1), it is unlikely that the dissociation rate would be significantly different between the two compartments. Based on these considerations, while acceptable in terms of fit statistics, Model 8 is biologically unlikely. Hence, our model-based analysis of the experimental data indicates that association and phosphorylation kinetics at the surface and in the endosomal compartments are comparable.

Receptor distribution predictions based on the best model

Model 5 is the best of the studied models in terms of AIC score and it is also the most parsimonious of the 3C models. Parameter estimation with this model yields unique estimates for all 3 fit parameters (see narrow ranges for parameters in Table 3). In Figure 3 we present predictions of the response to a saturating EGF concentration using all Model 5 parameter estimates that yielded a reasonable fit to the experimental data (see Methods). The tight overlap between the gray lines is indicative of the confidence in the predictions. Peak phosphorylation percentages of ~60% and ~50% are observed for the cell surface and EEs, respectively, but the overall percentage of phosphorylated receptors for the cell peaks at ~35% (Fig. 3A). This is because a significant fraction of the total receptor mass becomes localized to the LE (Fig. 3B), which does not contribute to receptor phosphorylation. We note that the distribution of dimers between the surface and EEs shifts with time (Fig. 3C and 3D); while ~63% of the dimers are at the cell surface 5 min after ligand addition, this percentage drops to ~50% by 15 min (Fig. 3D).

EGFR dephosphorylation rates

One of the key predictions of the best fitting models is that the pf value is comparable between cell surface and EE compartments. As discussed in the Model Construction section, the most likely source of variations in the pf value between cell surface and EE compartments is differences in phosphatase activities at these locations. We therefore validated the predictions of our model-based analysis by measuring the EGFR dephosphorylation rates to examine potential spatial dependence of phosphatase activities. Dephosphorylation rates were measured using two complementary approaches: imaging and ELISAs (Methods section). For both assays HME cells were treated with 25 ng/ml EGF for specified amounts of time to induce receptor phosphorylation and trafficking, following which a small molecule EGFR kinase inhibitor (1 μ M AG1478) was added to prevent further phosphorylation. Since the forward phosphorylation reaction was inhibited, the time-dependent decay in phosphorylation levels enabled us to quantify receptor dephosphorylation rates.

Imaging experiments enabled us to directly measure receptor dephosphorylation rates at the whole cell level and for the EE compartment. Cells were stimulated with Alexa 488-conjugated EGF for 15 min in these experiments. The use of fluorescently labeled EGF facilitated the identification of cells and endosomes, and enabled the normalization of the receptor phosphorylation measurements. The 15 min stimulus was chosen to ensure that a significant fraction of the receptors would be present in EEs. Figures 4A and 4B present representative green channel (EGF), red channel (pY1173), and merged images of two different cells taken at 0 and 80 sec following AG1478 addition, respectively. The lower relative levels of the red channel intensity in Fig. 4B compared to that in Fig. 4A is indicative of receptor dephosphorylation. We analyzed the green channel images to identify cells (Fig. 4C) as well as endosomes (Fig. 4D). The phosphorylation levels in cells and endosomes were fitted with an exponential decay function to determine the dephosphorylation rate (Fig. 4E). Based on imaging, we measured dephosphorylation rates of $7.8 \pm 4.3 \text{ min}^{-1}$ and $5.8 \pm 4.3 \text{ min}^{-1}$ for cells and endosomes, respectively (Fig. 4F, dark bars).

We have independently verified the optically measured cellular dephosphorylation rates in ELISA experiments. ELISA experiments took advantage of the observation that the dimer distribution between cellular compartments changes in time after stimulation with ligand (see, previous section), and that this change shifts the bias of the phosphorylation signal from cell surface to endocytic compartments. This suggests a novel means to indirectly validate if pf values for the cell surface and EE compartments are in fact similar: One can envision a cumulative pf factor for the whole cell, which will have a value in between the pf factors for cell surface and internal receptors. Its value would be determined by the relative dominance of the signal contributions coming from the two compartments. If pf is different for the cell surface and EE receptors, as the receptor population shifts from the cell surface to interior compartments, the cellular pf factor should vary over time. In contrast, an overall pf value that stays invariant as a function of time would indicate that pf values are comparable for the cell surface and EE compartments.

Cells were stimulated for 5, 10 or 15 min with EGF following which kinase inhibitor AG1478 was added. The stimulus times were chosen based on model predictions that the spatial distribution of phosphorylated receptors would be significantly different between 5 and 15 min (Fig. 3D). From the decay in the phosphorylation levels, dephosphorylation rates after 5, 10 and 15 min of EGF stimulation were determined to be $5.7 \pm 2.0 \text{ min}^{-1}$, $3.3 \pm 2.4 \text{ min}^{-1}$ and $5.0 \pm 3.8 \text{ min}^{-1}$, respectively (Fig. 4F, white bars). The dephosphorylation rate stays roughly constant over time while receptor trafficking shifts the signaling bias between the cellular compartments. As discussed above, this indirectly implies that the

dephosphorylation rates (and hence the pf factors) are comparable between cell surface and internal compartments.

We note that there was significant error associated with the measurements (Fig. 4F). While it may be possible that subtle differences in the dephosphorylation rates can be discerned using more sensitive assays, it is also possible that the errors were a true reflection of the inherent biological variability. Based on our results, we conclude that there are not significant differences in the dephosphorylation rates between cells and endosomes as measured in the imaging and ELISA experiments. These results support the modeling conclusion that the pf values are comparable between the cell surface and the endosomes, and thereby serve as a validation of the mathematical model. It should be pointed out that the dephosphorylation rates obtained in the two-independent, ELISA and imaging, experiments after 15 min of EGF stimulation were comparable. Equally importantly, none of the measured dephosphorylation rates were significantly different from each other based on ANOVA tests. These validations add further support to our conclusions.

As discussed above (cf., the Model Comparison section), although it has an acceptable fit score, Model 8 is biologically unlikely. Our dephosphorylation rate measurements provide further support to this argument – Model 8 predicts that pf values for the surface and EE compartments differ by a factor of 2.4. However, our results indicate comparable pf factors for the cell surface and internal compartments.

DISCUSSION

We analyzed EGFR phosphorylation dynamics in response to various doses of EGF using a mathematical model of EGFR activation and trafficking to extract information regarding receptor dimerization and phosphorylation. We found that: i) signal termination via receptor dephosphorylation in late endosomes is an important component of the response, ii) less than 40% of the receptors in the cell are phosphorylated at any instant of time even at saturating ligand doses (Figs. 2D, 3A) iii) we can treat the role of receptor dimerization and phosphorylation separately, and thus uniquely estimate dimer abundances (Figs. 2D, 3A) when the appropriate data is available, and iv) the phosphorylation efficiency is comparable at the cell surface and the EE. We validated the last finding by measuring the receptor dephosphorylation rates at various times following ligand addition both in whole cells and in endosomes.

The present study builds on our previous work where we used a simpler mathematical model to analyze EGFR phosphorylation.³³ In the earlier study, our experimental data consisted of the response to a single saturating (100 ng/ml) dose of EGF in the same HME cell line studied here. We chose not to fix any of the model parameters using prior information, and instead sought to estimate all of the parameters using the data. Hence, to keep the model size small enough to avoid over-fitting the data, our previous model contained a single lumped biochemical reaction step that led to the formation of “active” receptors, a pf scaling factor that was assumed to be the same for cell surface and EEs, and a receptor trafficking component that had lumped internalization, and degradation steps, with no recycling. In order to separately estimate the rate constant for receptor activation and the pf parameter, we needed to constrain the model by providing it with the percentage of cellular EGFR that was phosphorylated. The latter quantity was experimentally measured to have a peak value of ~20–30% in immunoprecipitation (IP) experiments, and the availability of this constraint then made it possible to complete the parameter estimation.³³ We note that quantification of the percent of phosphorylated receptors in IP experiments is subject to limitations such as the antibody pull-down efficiencies; and the measured low dimerization percentage was contrary to expectations. In the current study, the availability of dose response data and the

use of previously determined rate constants for receptor-ligand binding and trafficking allowed us to: i) use a more detailed model for the system, ii) separate the role of the dimerization affinities and pf values and iii) examine the spatial dependence of the k_u and pf values without additional IP data. Interestingly, the current model also predicted that less than 40% of the cellular EGFR are phosphorylated even at saturating doses of EGF. This prediction is in good agreement with our prior IP results and provides additional validation to our study. Our analysis also informs us that sorting of receptors into late endosomes where terminal dephosphorylation occurs is the main reason for the low levels of fractional phosphorylation.

There are several mechanisms that can potentially contribute to differences in receptor phosphorylation between the cell surface and the EE. For example, the differences in pH between these compartments has been shown to affect the on- and off- rates for receptor-ligand binding.⁴⁶ Further, the smaller volume of the EE can lead to an apparent increase in the ligand concentration thereby enhancing binding. Since we could quantitatively account for the changes in receptor-ligand binding in our mathematical model, we were able to specifically examine whether the phosphorylation efficiency was different between the cell surface and EE. This is an illustration of how mathematical modeling can be used to account for overlapping influences on the response, thereby enabling targeted investigation of particular features of the system.

There have been previous studies examining the EGFR phosphorylation dynamics using imaging. Offterdinger et al. constructed a single molecule FRET probe to assay EGFR phosphorylation in COS-7 cells.²⁷ They showed that long term EGF treatment resulted in the accumulation of dephosphorylated receptors in the perinuclear regions. This observation is also in agreement with other studies indicating that EGFR dephosphorylation precedes receptor degradation.^{22,45} Our three compartment model, where dephosphorylated EGFR can accumulate in the LE prior to degradation, was conceptually based on these studies. Further, as in our study, Offterdinger et al. monitored EGFR dephosphorylation at the cell membrane and in a central region of the cell by adding the tyrosine kinase inhibitor AG1478, 2 and 24 min after EGF stimulation²⁷. A rapid dephosphorylation on the order of 10s of seconds was observed in both instances. Both the time course of dephosphorylation decay and the observation of comparable dephosphorylation kinetics at 2 and 24 min are in agreement with our findings.

Hendriks et al.⁴⁰ analyzed ErbB phosphorylation dynamics in human lung carcinoma cells expressing comparable levels of EGFR, ErbB2 and ErbB3. Their study employed a two compartment model with a cell surface and a lumped internal compartment, and model parameters, including dimerization affinities, were kept fixed at constant values. Rather than optimizing model parameters against experimental data, Hendriks et al. visually compared the shapes of the model predictions for the fraction of phosphorylated receptors with experimental data on receptor phosphorylation. Their analysis assumed that the ratio of phosphorylated receptors to dimers (analogous to our pf value) was initially 1, but decayed exponentially in time as a consequence of dephosphorylation. Setting this apparent dephosphorylation rate to 0.18 min^{-1} in the endosomes enabled them to generate model predictions that matched the shape of the experimental data.⁴⁰ In the current study, we estimated the dimerization affinity and the pf scaling factor for the cell surface and endosomes by fitting a mathematical model to the experimental data. The pf value in our case encapsulates the activity of phosphatases that counter EGFR phosphorylation. Rather than modeling terminal dephosphorylation by allowing the pf value of a lumped interior compartment to decay with time, our model allowed dephosphorylated receptors to accumulate in a separate LE compartment. We believe that this conceptual representation is closer to the reality where phosphatases regulate EGFR phosphorylation levels and terminal

dephosphorylation at separate subcellular locations.^{26–29} That said, the conclusion of Hendriks et al. that the phosphorylation decay is localized to the internal compartment is consistent with our model. Further, we can use the late endosomal entry rate of dimers ($k_{l,d}$) as an estimate of the terminal dephosphorylation rate in our system. Using parameter estimates for Model 5 and the expression for $k_{l,d}$ in Fig. 1 we get a value of $\sim 0.15 \text{ min}^{-1}$ for the irreversible dephosphorylation step, which is comparable to the phosphorylation decay rate of 0.18 min^{-1} determined by Hendriks et al.⁴⁰

The long term objective of our research is to establish a quantitative link between the ErbB dimerization pattern and cell behavior.^{30,33,49} As we continue to accumulate quantitative information regarding cell signaling pathways, and seek to understand the operating principles of these complex networks, mathematical models have become an integral part of the endeavor. The long term goal of model development is the ability to predict the system response to perturbations in the absence of experiments, or in a sense, to replace experiments altogether. However, in order to get to that point, we need to go through an iterative cycle of model refinement, where existing data is used to improve a model, and new predictions from the model are experimentally validated. For this process to be effective, we need to pay attention to model construction – models in general need to be as simple as possible (parsimonious) to enable parameter estimation, while still representing the important events in the system. The process of model refinement, rather than merely being a stepping stone towards the construction of a predictive model, can offer up significant mechanistic insights about the system. Our current study represents an illustrative example wherein the analysis of experimental data using parsimonious models resulted in a novel insight regarding the system, which was then experimentally validated.

Acknowledgments

The research described in this paper was funded by the National Institutes of Health Grant 5R01GM072821-06 to H.R. Pacific Northwest National Laboratory is a multiprogram national laboratory operated by Battelle for the U.S. Department of Energy under Contract DE-AC06-76RL01830.

References

1. Yarden Y, Sliwkowski MX. Nat Rev Mol Cell Biol. 2001; 2:127–137. [PubMed: 11252954]
2. Holbro T, Hynes NE. Annu Rev Pharmacol Toxicol. 2004; 44:195–217. [PubMed: 14744244]
3. Holbro T, Civenni G, Hynes NE. Experimental cell research. 2003; 284:99–110. [PubMed: 12648469]
4. Hynes NE, MacDonald G. Curr Opin Cell Biol. 2009; 21:177–184. [PubMed: 19208461]
5. Normanno N, Bianco C, De Luca A, Maiello MR, Salomon DS. Endocr Relat Cancer. 2003; 10:1–21. [PubMed: 12653668]
6. Herbst RS. International journal of radiation oncology, biology, physics. 2004; 59:21–26.
7. Patel R, Leung HY. Current pharmaceutical design. 2012
8. Seshacharyulu P, Ponnusamy MP, Haridas D, Jain M, Ganti AK, Batra SK. Expert opinion on therapeutic targets. 2012; 16:15–31. [PubMed: 22239438]
9. Mitsudomi T, Yatabe Y. The FEBS journal. 2010; 277:301–308. [PubMed: 19922469]
10. Citri A, Yarden Y. Nat Rev Mol Cell Biol. 2006; 7:505–516. [PubMed: 16829981]
11. Lazzara MJ, Lauffenburger DA. Experimental cell research. 2009; 315:717–725. [PubMed: 19022246]
12. Olayioye MA, Neve RM, Lane HA, Hynes NE. Embo J. 2000; 19:3159–3167. [PubMed: 10880430]
13. Wiley HS. Experimental cell research. 2003; 284:78–88. [PubMed: 12648467]
14. Wiley HS, Shvartsman SY, Lauffenburger DA. Trends in Cell Biology. 2003; 13:43–50. [PubMed: 12480339]

15. Dawson JP, Berger MB, Lin CC, Schlessinger J, Lemmon MA, Ferguson KM. *Mol Cell Biol.* 2005; 25:7734–7742. [PubMed: 16107719]
16. Schlessinger J. *Cell.* 2002; 110:669–672. [PubMed: 12297041]
17. Lemmon MA. *Experimental cell research.* 2009; 315:638–648. [PubMed: 19038249]
18. Hendriks BS, Opresko LK, Wiley HS, Lauffenburger D. *Cancer research.* 2003; 63:1130–1137. [PubMed: 12615732]
19. Bose R, Zhang X. *Experimental cell research.* 2009; 315:649–658. [PubMed: 18761339]
20. Olayioye MA, Graus-Porta D, Beerli RR, Rohrer J, Gay B, Hynes NE. *Mol Cell Biol.* 1998; 18:5042–5051. [PubMed: 9710588]
21. Yen L, Benlimame N, Nie ZR, Xiao D, Wang T, Al Moustafa AE, Esumi H, Milanini J, Hynes NE, Pages G, Alaoui-Jamali MA. *Mol Biol Cell.* 2002; 13:4029–4044. [PubMed: 12429844]
22. Burke P, Schooler K, Wiley HS. *Mol Biol Cell.* 2001; 12:1897–1910. [PubMed: 11408594]
23. Hancock JF. *Nat Rev Mol Cell Biol.* 2003; 4:373–384. [PubMed: 12728271]
24. Haugh JM, Schooler K, Wells A, Wiley HS, Lauffenburger DA. *J Biol Chem.* 1999; 274:8958–8965. [PubMed: 10085141]
25. Kolch W. *Nat Rev Mol Cell Biol.* 2005; 6:827–837. [PubMed: 16227978]
26. Haj FG, Verveer PJ, Squire A, Neel BG, Bastiaens PI. *Science.* 2002; 295:1708–1711. [PubMed: 11872838]
27. Offterdinger M, Georget V, Girod A, Bastiaens PI. *J Biol Chem.* 2004; 279:36972–36981. [PubMed: 15215236]
28. Reynolds AR, Tischer C, Verveer PJ, Rocks O, Bastiaens PI. *Nature cell biology.* 2003; 5:447–453.
29. Yudushkin IA, Schleifenbaum A, Kinkhabwala A, Neel BG, Schultz C, Bastiaens PI. *Science.* 2007; 315:115–119. [PubMed: 17204654]
30. Zhang Y, Opresko L, Shankaran H, Chrisler WB, Wiley HS, Resat H. *BMC Cell Biol.* 2009; 10:78. [PubMed: 19878579]
31. Band V, Sager R. *Proc Natl Acad Sci U S A.* 1989; 86:1249–1253. [PubMed: 2919173]
32. Resat H, Ewald JA, Dixon DA, Wiley HS. *Biophys J.* 2003; 85:730–743. [PubMed: 12885624]
33. Shankaran H, Zhang Y, Opresko L, Resat H. *Biochem Biophys Res Commun.* 2008; 371:220–224. [PubMed: 18424261]
34. Kholodenko B, Yaffe MB, Kolch W. *Science signaling.* 2012; 5:re1. [PubMed: 22510471]
35. Bachmann J, Raue A, Schilling M, Becker V, Timmer J, Klingmuller U. *Journal of internal medicine.* 2012; 271:155–165. [PubMed: 22142263]
36. Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK. *Nature cell biology.* 2006; 8:1195–1203.
37. Hughey JJ, Lee TK, Covert MW. *Wiley interdisciplinary reviews. Systems biology and medicine.* 2010; 2:194–209. [PubMed: 20836022]
38. Ventura AC, Jackson TL, Merajver SD. *Cancer research.* 2009; 69:400–402. [PubMed: 19147549]
39. Chen WW, Schoeberl B, Jasper PJ, Niepel M, Nielsen UB, Lauffenburger DA, Sorger PK. *Mol Syst Biol.* 2009; 5:239. [PubMed: 19156131]
40. Hendriks BS, Cook J, Burke JM, Beusmans JM, Lauffenburger DA, de Graaf D. *Syst Biol (Stevenage).* 2006; 153:22–33. [PubMed: 16983832]
41. Hendriks BS, Orr G, Wells A, Wiley HS, Lauffenburger DA. *J Biol Chem.* 2005; 280:6157–6169. [PubMed: 15572377]
42. Shankaran H, Wiley HS, Resat H. *Biophys J.* 2006; 90:3993–4009. [PubMed: 16533841]
43. Wiley HS, Herbst JJ, Walsh BJ, Lauffenburger DA, Rosenfeld MG, Gill GN. *J Biol Chem.* 1991; 266:11083–11094. [PubMed: 1645724]
44. Sorkin A, Goh LK. *Experimental cell research.* 2008; 314:3093–3106. [PubMed: 18793634]
45. Eden ER, White IJ, Tsapara A, Futter CE. *Nature cell biology.* 2010; 12:267–272.
46. French AR, Tadaki DK, Niyogi SK, Lauffenburger DA. *J Biol Chem.* 1995; 270:4334–4340. [PubMed: 7876195]
47. Akaike H. *IEEE Trans Autom Control.* 1974; AC-19:716–723.

48. Burnham, KP.; Anderson, DR. Model selection and inference: A practical information-theoretic approach. Springer-Verlag; New York, NY, USA: 1998.
49. Zhang Y, Shankaran H, Opresko L, Resat H. IET Syst Biol. 2008; 2:273–284. [PubMed: 19045822]
50. Hendriks BS, Opresko LK, Wiley HS, Lauffenburger D. J Biol Chem. 2003; 278:23343–23351. [PubMed: 12686539]
51. Kholodenko BN, Demin OV, Moehren G, Hoek JB. J Biol Chem. 1999; 274:30169–30181. [PubMed: 10514507]
52. Hendriks BS, Wiley HS, Lauffenburger D. Biophys J. 2003; 85:2732–2745. [PubMed: 14507736]

\$watermark-text

\$watermark-text

\$watermark-text

Figure 1. Mathematical models for EGFR phosphorylation and trafficking

Illustrations are presented for the receptor trafficking component of two (left panel), and three (right panel) compartment models for the EGFR system. In the two compartment model, comprising of the cell surface and early endosomes (EEs), receptor monomers and dimers are internalized at rate $k_{e,m}$ and $k_{e,d}$, recycled from the EE at rates $k_{r,m}$ and $k_{r,d}$ and degraded at rates $k_{d,m}$ and $k_{d,d}$ respectively. In the three compartment model, monomers and dimers enter the late endosome (LE) from the EE at rates $k_{l,m}$ and $k_{l,d}$ respectively. Receptors are added to the cell surface with flux Q_R in both model types. Expressions relating the individual trafficking rates to the endosomal exit rate, k_x , the recycling fraction, f , and the parameter, δ (ratio of entry to exit rate for the LE), are provided. For both the two and three compartment models the biochemical reactions indicated in the box at the top of the figure are simulated at the cell surface, and EE, respectively. The biochemical reactions of receptor-ligand binding, and dimerization, when combined with receptor trafficking, enable prediction of the receptor dimer (RD_s , RD_e) as well as the total receptor ($R_{t,s}$, $R_{t,e}$, $R_{t,l}$) numbers at the cell surface (subscript s), EE (subscript e) and LE (subscript l). Note that the “ x ” in the expression for $R_{t,x}$ in the box should be substituted with “ s ”, “ e ” or “ l ” depending upon whether receptor masses are being calculated for the cell surface, EE or LE. Phosphorylation (pf), and receptor mass (mf) scaling factors are used to convert dimer and total receptor abundance predictions to the phosphorylation (total – pR_i ; internal pR_i), and receptor mass (mR_i) levels that would be measured in an ELISA experiment. The specific model predictions that are directly compared with experimental data are indicated in green. The unknown parameters that are estimated by data fitting (k_u , pf and δ_m) are shown in red. Although δ_d (shown in blue) is an unknown parameter, it is related to δ_m via an algebraic relation (see Methods) and is not an independent parameter. We tested various versions of the two and three compartment models by allowing k_u and/or pf to vary between the cell surface and the EE (boxes at the bottom of the left and right panels).

Figure 2. Model fits to experimental data and comparison of model predictions

A & B) The complete experimental dataset (markers in A and B) involving measurements of total EGFR phosphorylation as a function of EGF dose (panel A) and detailed time courses of receptor mass, and total (circles) and internal (squares) phosphorylation in response to 12 ng/ml EGF (panel B). The lines in the panels represent model predictions using Model 5, the best model in terms of the Akaike Information Criterion. **C)** Comparison of model fits to total receptor phosphorylation (top panel), internal phosphorylation (middle panel) and receptor mass (bottom panel) data collected following addition of 12 ng/ml EGF. Model fits are shown for two compartment models (dotted lines; red, green, blue and pink are used for models 1–4, respectively) and three compartment models (solid lines; red, green, blue and pink are used for models 5–8, respectively). **D)** Predictions for the receptor dimer fraction in the cell in response to a saturating EGF concentration (100 ng/ml) using each of the 8 models. Line styles and colors for the various models are as in panel C.

Figure 3. Model predictions for receptor dimerization and receptor distribution

Model predictions are shown for model 5 using the best fit parameter estimate (estimate with the lowest RMSE; dark colored lines) and for all estimates with RMSE within 2% of the minimum value (gray lines). Response predictions are shown as a function of time following the addition of a saturating EGF concentration (100 ng/ml). Model 5 is the best among the 8 models tested in terms of the AIC. **A)** Phosphorylated EGFR as a fraction of total EGFR is presented for the cell as a whole (red), cell surface (green) and EE (blue). Receptors are assumed to be phosphorylated when present as part of a dimer at the cell surface or the EE. **B)** Receptor numbers at the surface, EE and LE are presented as a fraction of total cellular receptor abundance. **C)** The absolute number of receptor dimers at the cell surface and EE. **D)** Dimers at the cell surface (red) and in the EE (green) as a percentage of the total dimer abundance in these compartments put together.

Figure 4. EGFR dephosphorylation rate measurements

HMECs were stimulated for specified amounts of time to induce receptor phosphorylation following which 1 μ M AG1478 was added to inhibit phosphorylation to monitor the subsequent dynamics to measure the dephosphorylation rate. Receptor phosphorylation levels were measured as a function of time following AG1478 addition using anti-phospho-Tyr 1173 (pY1173) antibody. **A&B**) Representative images of fixed cells stimulated with Alexa 488 conjugated EGF (green channel) and stained with Alexa 647 conjugated pY1173 (red channel) taken before (panel A), and at 80 sec (panel B) after AG1478 addition. Green channel (top left), red channel (top right) and merged (bottom left) images are shown in each case. **C&D**) Representative results of automated image analysis to identify cells (panel C) and endosomes (panel D) using green channel intensities. The green outlines in each of the panels represent the boundaries of objects determined using image analysis. **E**) Representative receptor dephosphorylation time course (markers) along with an exponential decay fit (solid line) used to determine the dephosphorylation rate. Results are shown for ELISA measurements of receptor phosphorylation levels following addition of AG1478. **F**) Dephosphorylation rates determined using ELISAs following addition of EGF for 5, 10 and 15 min (white bars), and rates determined for cells and endosomes using imaging following addition of EGF for 15 min (gray bars).

Table 1

Parameters used in the mathematical models *

Parameter	Description	Value/Range
Fixed parameters		
$k_{on,s}$	On rate for receptor-ligand binding at surface	$9.7 \times 10^7 / \text{M/min}^a$
$k_{off,s}$	Off rate for receptor-ligand binding at surface	$0.24 / \text{min}^a$
$k_{on,e}$	On rate for receptor-ligand binding in EE	$8.5 \times 10^6 / \text{M/min}^b$
$k_{off,e}$	Off rate for receptor-ligand binding in EE	$0.66 / \text{min}^b$
k_c	On rate for dimerization	$2 \times 10^{-4} (\#/\text{cell})^{-1} \text{min}^{-1}^c$
$k_{e,m}$	Monomer internalization rate	$0.07 / \text{min}^a$
$k_{e,d}$	Dimer internalization rate	$0.28 / \text{min}^a$
$k_{x,m}$	Monomer endosomal exit rate	$0.08 / \text{min}^d$
$k_{x,d}$	Dimer endosomal exit rate	$0.04 / \text{min}^a$
f_m	Monomer recycling fraction	0.8^a
f_d	Dimer recycling fraction	0.4^a
V_{mc}	Media volume per cell	$5 \times 10^{-10} \text{ liters}^e$
V_e	EE volume	$3 \times 10^{-14} \text{ liters}^f$
R_0	Receptor expression in the cell prior to ligand addition	$2 \times 10^5 \text{ molecules}^a$
mf	Receptor mass scaling factor	$5.5 \times 10^{-4} (\text{ELISA unit})/\text{receptor}^g$
Unknown parameters		
$k_{u,s}$	Dimer dissociation rate at surface	$10^{-3} - 10^3 / \text{min}$
$k_{u,e}$	Dimer dissociation rate in EE	$10^{-3} - 10^3 / \text{min}$
pf_s	Phosphorylation scaling factor for surface	$10^{-5} - 10^0 (\text{ELISA unit}/\text{dimer})$
pf_e	Phosphorylation scaling factor for EE	$10^{-5} - 10^0 (\text{ELISA unit}/\text{dimer})$
δ_m	Ratio of LE entry to exit rate for monomers	$0.1 - 2$

* Exact values are presented for fixed parameters; ranges used during parameter estimation are presented for the unknown parameters

^aBased on 18

^bBased on 46

^c k_c is set to the diffusion-limited value of $1 \times 10^{-3} (\#/\text{cell})^{-1} \text{min}^{-1}$ in 50. A k_c value of $\sim 3 \times 10^{-4} (\#/\text{cell})^{-1} \text{min}^{-1}$ is used in 51, with the assumption that the dimerization rate is lower than the diffusion-limited value. Here, we set k_c at 1/5th the diffusion-limited value.

^dBased on 41

^eBased on 52

^fThe media volume per cell was calculated assuming 4 ml of media for a plate containing 8×10^6 cells, based on the cell treatment conditions used for our experiments

^gThe mass scaling factor was calculated as $mf = mR_0 / 2 \times 10^5$ where mR_0 is the ELISA receptor mass reading prior to ligand addition (see Methods)

Table 2

Fit statistics for the various mathematical models*

Model	Type ¹	Assumptions	N _{par} ²	SSE ³	MSE ⁴	RMSE ⁵	L ⁶	AIC ⁷
1	2C	$k_{u,s} = k_{u,e}; pf_s = pf_e$	2	2.39	0.0443	0.2105	-12.00	28.01
2	2C	$k_{u,s}$ $k_{u,e}; pf_s = pf_e$	3	1.85	0.0343	0.1852	-9.01	24.02
3	2C	$k_{u,s} = k_{u,e}; pf_s$ pf_e	3	1.06	0.0197	0.1402	-2.47	10.95
4	2C	$k_{u,s}$ $k_{u,e}; pf_s$ pf_e	4	1.03	0.0192	0.1384	-2.17	12.34
5	3C	$k_{u,s} = k_{u,e}; pf_s = pf_e$	3	0.65	0.0120	0.1096	3.31	-0.62
6	3C	$k_{u,s}$ $k_{u,e}; pf_s = pf_e$	4	0.62	0.0115	0.1072	3.82	0.36
7	3C	$k_{u,s} = k_{u,e}; pf_s$ pf_e	4	0.65	0.0120	0.1095	3.33	1.34
8	3C	$k_{u,s}$ $k_{u,e}; pf_s$ pf_e	5	0.59	0.0109	0.1044	4.43	1.13

* Statistics are based on a weighted residual vector wherein residuals for phosphorylation level, and receptor mass measurements were divided by the respective maximum values for each of these measurements. There were a total of 54 experimental data points used for parameter estimation (see Figs. 2A & 2B). The model with the lowest AIC is highlighted in bold.

¹ Model type – 2C = two compartment model; 3C = three compartment model

² N_{Par} – Number of model parameters

³ SSE – Sum of squares error (or residual sum of squares)

⁴ MSE – Mean-squared error = SSE/N_{pts} where N_{pts} is the number of data points (= 54)

⁵ RMSE – root-mean squared error

⁶ L – log-likelihood function = $-(N_{pts}/2)[\log(2\pi \text{MSE})+1]$

⁷ AIC – Akaike information criterion = $2N_{Par} - 2L$

Table 3

Parameter estimates for the three compartment models *

Model	$k_{u,s}$	$k_{u,e}$	$pf_s \times 10^2$	$pf_e \times 10^2$	δ_m
5	2.95 (2.48 – 3.61)		2.97 (2.81 – 3.19)		1.27 (1.24 – 1.29)
6	8.08 (3.39 – 19.0)	1.43 (0.66 – 3.05)	3.95 (2.88 – 5.86)		1.42 (1.27 – 1.58)
7	2.45 (0.84 – 7.31)		2.94 (2.72 – 3.54)	2.44 (1.21 – 8.87)	1.17 (0.81 – 1.66)
8	6.83 (4.21 – 15.0)	0.094 (0.002 – 0.89)	4.42 (3.67 – 5.70)	1.82 (1.36 – 3.46)	1.10 (0.94 – 1.38)

* The parameter values that resulted in the smallest RMSE are presented for each model. The minimum and maximum values for each parameter, indicated within the parentheses, were determined by taking into account all parameter estimates that had an RMSE within 2% of the best fit RMSE.