SUPPORTING MATERIAL

Use of Mechanistic Models to Integrate and Analyze Multiple Proteomic Datasets

Edward C. Stites,[1,2*] Meraj Aziz,[1] Matthew S. Creamer,[1,3] Daniel D. Von Hoff,[1]
Richard G. Posner,[1,4*] and William S. Hlavacek[1,5*]

[1]Clinical Translational Research Division, Translational Genomics Research Institute, Phoenix, Arizona, USA; [2]Department of Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri, USA; [3]Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, Connecticut, USA; [4]Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona, USA; and [5]Theoretical Biology & Biophysics Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

TEXT S1

Specifying model structure

The structure of the (generic) model was specified using a rule-based approach (1), in which biomolecular interactions are represented in terms of rules. Rules were written using the BioNetGen language (BNGL) (2), which is a model-specification language compatible with a number of software tools, including BioNetGen (2, 3) and NFsim (4), which were used in this study to perform simulations. Most of the BNGL-encoded rules in the model were written for interactions between sites of autophosphorylation in EGFR and proteins containing SH2 and/or PTB domains. We considered all such interactions characterized by Hause et al. (5), who used a fluorescence polarization method to measure equilibrium dissociation constants (K_D's) for interactions between EGFR phosphopeptides and isolated SH2 and PTB domains. The K_D's reported by Hause et al. (5) are all less than 20 μM. For simulations, we reduced the generic model to a cell line-specific model, meaning a model consisting only of the subset of rules for interactions between proteins expressed in a particular cell line. The expressed proteins were defined by the protein copy number measurements of Kulak et al. (6) or Geiger et al. (7), as indicated, in the case of HeLa cells and by the measurements of Geiger et al. (7) in the case of the other 10 cell lines. For each version of the model (generic, HeLa cell specific, HEK 293 cell specific, etc.), a complete, executable model specification is provided in the form of a plain-text file with a .bngl filename extension (Software S1 in the Supporting Material). Furthermore, each cell line-specific file (e.g., EGFR_K_HeLa.bngl or EGFR_G_HeLa.bngl) is accompanied by a plain-text file with a .rnf filename extension, which provides a simulation protocol. The EGFR_generic.bngl file is provided mainly as a resource; this file includes rules for all of the EGFR interactions that were quantitatively characterized by Hause et al. (5). The library of rules provided as a resource in the EGFR_generic.bngl file could be valuable in future work aimed at further investigating cell line-specific differences in EGFR signaling.

As a simplification, we assumed that interactions of the SH2/PTB domain-containing proteins of interest with non-EGFR binding partners act as a uniform sink and therefore need not be considered explicitly in the model, given our focus on relative extents of recruitment. With use of data from more comprehensive assays of SH2/PTB interactions (8), this mean-field assumption could potentially be lifted in future work, provided that information is available

about the basal and EGF-regulated tyrosine phosphorylation levels of the non-EGFR binding partners of the relevant SH2/PTB domain-containing proteins. As a further simplification, we assumed that interactions are noncooperative. Thus, for example, we wrote the following BNGL-encoded rule for interaction of GRB2 (via its SH2 domain) with EGFR at Y1092 (UniProt numbering, entry P00533-1):

```
EGFR(Y1092~pY)+GRB2(SH2)<->EGFR(Y1092~pY!1).GRB2(SH2!1) ka_SH2PTB,kdErbB1_1092_GRB2
```

where `ka_SH2PTB` is an association rate constant and `kdErbB1_1092_GRB2` is a dissociation rate constant. This rule assumes that the interaction between EGFR and GRB2 is not affected, negatively or positively, by the molecular context in which the interaction occurs. The rate constants that appear in this rule were assigned values consistent with the K_D measured by Hause et al. (5) for this interaction, as described below. At bottom, we briefly explain how the rule given above should be interpreted under the conventions of BNGL (2). We also discuss the interpretation of additional rules.

Rules similar in form to the rule given above are included in our model to capture phosphotyrosine-SH2/PTB domain interactions characterized by Hause et al. (5). Recall that these interactions involve phosphotyrosines in EGFR, which are produced through ligand- and dimer-dependent autophosphorylation and removed through high, constitutive phosphatase activity and endocytosis. The SH2 and PTB domains considered in the model are found in direct binding partners of EGFR. These binding partners include signaling proteins that adopt autoinhibited conformations, that contain multiple (two) PTB/SH2 domains, that are tethered to the inner leaflet of the plasma membrane (viz., SFKs), and that are free to diffuse in the cytoplasm (i.e., cytosolic). As a simplification, we do not distinguish between cytosolic and plasma membrane-associated EGFR binding partners. In other words, all EGFR binding partners are taken to be cytosolic. Furthermore, we do not consider two-point attachment of signaling proteins to EGFR, which is possible in principle for the signaling proteins with multiple PTB/SH2 domains (e.g., SHC1, which contains a PTB domain and an SH2 domain). In other words, for these proteins, we only allow one of the two PTB/SH2 domains to interact with EGFR at a time. Finally, we do not consider autoinhibition.

In addition to rules for phosphotyrosine-PTB/SH2 domain interactions, the model consists of rules for EGF-EGFR binding, constitutive and EGF-induced EGFR dimerization, EGF-dependent autophosphorylation within dimers, and receptor dephosphorylation at phosphotyrosines that are not shielded by a binding partner. As a simplification, we only consider homodimers of EGFR, neglecting heterodimers of EGFR with other ErbB-family members. All phosphotyrosines that are not shielded by a binding partner are taken to be equally available for dephosphorylation. In the model, the phosphatases responsible for dephosphorylation of phosphotyrosines are implicit and are taken to be constitutively active. Thus, for example, activation of the protein tyrosine phosphatase PTPN11, which is a direct binding partner of EGFR included in the model, is not considered. This treatment of phosphatases is justified to some extent by high basal phosphatase activity, which can be expected to be the dominant fraction of phosphatase activity at early times after activation of EGFR signaling. Extension of the model to include activation of PTPN11, which is recruited to EGFR not only directly but also indirectly, would require consideration of interactions among the direct binding partners of EGFR, which is beyond the intended scope of the modeling effort. Rules for phosphatase activity and the other processes mentioned above are of the same form as rules included in the models of Creamer et al. (9) and Kozer et al. (10, 11). Other rules are included for receptor internalization, recycling, and ligand and receptor degradation, the

dynamics of which are intertwined with EGFR phosphorylation and recruitment dynamics. Trafficking and degradation were treated as in the model of Schoeberl et al. (12). This simplistic treatment captures the phenomenology of receptor internalization and related processes deemed relevant for our purposes but not the mechanistic details, such as receptor ubiquitination (13). A more careful treatment of the coupling between EGFR signaling and trafficking, as in the study of Shankaran et al. (14), would require consideration of multiple endosomal compartment types.

Setting parameter values

Parameter values are given and annotated in Table S1 in the Supporting Material, which is an Excel spreadsheet. They are also given in the executable model specifications (Software S1). Parameter values were set without systematic fitting and largely on the basis of proteomic data (5-7).

The parameters of each model include protein copy numbers that are specific for a particular cell type/line and rate constants for protein-protein interactions. These rate constants are parameters of rate laws associated with rules in the model. All reactions implied by a rule are taken to have the rate law assigned to the rule in the model specification. (The reactions implied by a rule involve a common transformation but may encompass many distinct reactants and products.) Thus, this simplification, which is characteristic of the rule-based approach to modeling (1), has the effect of coarse graining the chemical kinetics of a system, because in principle, each distinct individual reaction implied by a rule (with its particular set of reactants and products) may have a unique rate law. In the case of rules for phosphatase activity, phosphatases are not considered explicitly and the rate laws for these rules should be understood to represent approximations consistent with pseudo first-order decay of unprotected phosphorylated sites to their unphosphorylated forms. Such a rate law is appropriate if phosphatases are available in excess. The phosphatase activity in the EGFR signaling network is expected to be high (15). Rules for receptor internalization should also be understood to be associated with coarse-grained, approximate rate laws. The model does not attempt to account for the complicated details of the endocytic machinery responsible for ligand-induced receptor internalization, or lysosomal degradation.

For HeLa cells, we set copy numbers for EGFR and its cytoplasmic binding partners at the absolute protein abundances reported by Kulak et al. (6) or in the manner described below on the basis of the data of Geiger et al. (7), as indicated. Thus, we considered two HeLa cell-specific models, which were parameterized differently. For HEK 293 cells and other cell lines, we set copy numbers using the average of the intensity-based absolute quantification (iBAQ) values reported by Geiger et al. (7). In our simulations of responses of EGF-starved single cells to addition of a bolus of EGF, we considered a dose of 200,000 copies of EGF (6.2 kDa) in the volume of fluid surrounding a single cell, which is 1 nL for cells in suspension at a density of 10^6 cells/ml. For this volume, this dose corresponds to an EGF concentration of 2 ng/ml. To relate measured protein copy numbers to intracellular concentrations, we assumed that the cytoplasmic volume of a cell is 2 pL, the approximate volume estimated for a HeLa cell by Moran et al. (16).

For phosphotyrosine-SH2/PTB domain interactions, we set rate constants as follows. For each interaction, we set the forward (bimolecular association) rate constant, k_{on}, to 10^6 $M^{-1}s^{-1}$, which is a typical value for the interaction of an SH2/PTB domain with a phosphotyrosine docking site (17). We then set the reverse (unimolecular dissociation) rate constant to the value consistent with the corresponding K_D measured by Hause et al. (5). In other words, we set $k_{off} = k_{on}K_D$.

This approach ensures that interactions are parameterized consistently with measured equilibrium dissociation constants and assumes that variability in K_D values mainly arises from variability in dissociation rate constants, or equivalently lifetimes of protein-protein interactions. This assumption was deemed to provide reasonable first approximations.

The selection of other parameter values was guided by parameter estimates reported earlier in the literature. The reaction scheme considered for EGF-EGFR and EGFR-EGFR interactions is that of Wofsy et al. (18). The rate constants for the reactions in this scheme were chosen to be consistent with the estimates of equilibrium constants (and data) of Macdonald and Pike (19). There is also consistency with measurements/estimates of kinetic parameters for EGF-EGFR binding and EGFR dimerization (20). In the model, receptor dimers are taken to have a lifetime of 100 s. Although single-particle tracking results indicate that dimers bound to different numbers of ligands (0, 1, or 2) have different lifetimes (21), the differences are not dramatic. Rules for site-specific dimerization-dependent receptor autophosphorylation and site-specific receptor dephosphorylation were each assigned a (pseudo) first-order rate law with a rate constant estimated at 0.1 s^{-1}. This estimate is consistent with parameter values considered in the modeling studies of Kholodenko et al. (22) and Creamer et al. (9) and the empirical observations of high phosphatase activity made by Kleiman et al. (15). Although *in vitro* experiments indicate that the EGFR kinase has different k_{cat} and K_M values for different EGFR substrates (23), the differences do not appear to be dramatic. As a simplification, phosphatases are not considered to be inducible. As a further simplification, the model does not include EGFR phosphorylation in the absence of EGF. This latter assumption is more justified for a low (vs. high) level of EGFR expression (24). Parameters for ligand and receptor internalization and related processes were taken from Schoeberl et al. (12).

Interpretation of selected rules from the generic model for EGFR signaling

As noted above, in the generic model, we wrote the following BNGL-encoded rule for interaction of GRB2 (via its SH2 domain) with EGFR at Y1092:

```
EGFR(Y1092~pY)+GRB2(SH2)<->EGFR(Y1092~pY!1).GRB2(SH2!1) ka_SH2PTB,kdErbB1_1092_GRB2
```

where `ka_SH2PTB` is an association rate constant and `kdErbB1_1092_GRB2` is a dissociation rate constant.

This rule can be viewed as a generalized reversible reaction that defines forward and reverse classes of related but distinguishable reaction events arising from an interaction between two biomolecules named `EGFR` and `GRB2`. The rule indicates that this interaction is mediated by molecular components named `Y1092` (a constituent of `EGFR`) and `SH2` (a constituent of `GRB2`). Notably, the formal elements of the rule are in one-to-one correspondence with material entities (e.g., `GRB2` in the rule refers to GRB2 and `SH2` refers to the SH2 domain of GRB2). The nomenclature of rules is similar to that of standard chemical reactions; however, rules differ from standard chemical reactions in that they do not uniquely identify reactants (or products), which allows for concise model specification. For example, the rule above is silent about functional components of EGFR and GRB2 other than the SH2 domain in GRB2 and its phosphotyrosine docking site in EGFR. Thus, under a “don’t care, don’t write” policy, the rule applies to multiple phosphoforms of EGFR (and even distinct EGFR-containing signaling complexes) and consequently defines multiple reactions.

The rule above has left- and right-hand sides, which are separated by the symbol “`<->`.” This symbol indicates that the interaction represented by the rule is reversible. The plus sign on the

left-hand side indicates that the rule defines bimolecular (association) reactions when it is read from left to right. The absence of a plus sign on the right-hand side indicates that the rule defines unimolecular (dissociation) reactions when it is read in the opposite direction. When two rate constants are associated with a rule, as is the case here, two elementary mass-action rate laws are implied. In the rule above, one rate law, with rate constant `ka_SH2PTB`, is implied for all association reactions defined by the rule, and a second rate law, with rate constant `kdErbB1_1092_GRB2`, is implied for all dissociation reactions defined by the rule. The use of a single rate law for all association (or dissociation) reactions is a simplification, a type of coarse graining.

The left-hand side of the rule above defines necessary and sufficient conditions for GRB2 association with EGFR. Namely, association may occur if and only if the SH2 domain of GRB2 is free (indicated by an absence of a bond index appended to the component name `SH2`, which when present in a rule is prefixed by a "`!`" symbol) and Y1092 in EGFR is both free (again indicated by the absence of a bond index) and in an internal state labeled "`pY`." Internal states are convenient abstractions, which can be used to represent local properties of molecular components. Here, we use internal states to represent phosphorylation states of tyrosines. The label `pY` is used to represent a phosphorylated tyrosine, and the label `Y` is used to represent an unphosphorylated tyrosine. Internal state labels are prefixed by a tilde. Similarly, as already noted, bond indices are prefixed by an exclamation mark. The right-hand side of the rule defines the outcome of GRB2 association with EGFR: the formation of a bond between GRB2's SH2 domain and EGFR's pY1092 site. The bond is indicated through name sharing. In BNGL, associating a common bond index, such as "`1`" in the rule above, with a pair of component names indicates that the components are connected. The dot on the right-hand side of the rule serves as a separator; it also indicates that EGFR and GRB2 are connected (without specifying how). In the case of this particular rule, the dot is redundant, because sharing of the bond index `1` by `Y1092` and `SH2` also indicates that EGFR and GRB2 are connected. When read from right to left, the rule indicates that a bond between GRB2's SH2 domain and pY1092 in EGFR can be broken and that breaking of the bond causes dissociation of the proteins.

We will now discuss additional rules in the model, which characterize the following processes: ligand-receptor binding, receptor dimerization, autophosphorylation, dephosphorylation, signaling protein recruitment (in a case where the signaling protein includes two SH2 domains), internalization/recycling, and degradation. The rules presented below do not provide a complete characterization of these processes. The intention is to discuss the interpretation of selected rules from the model for readers who are not familiar with the conventions of BNGL.

The following rule characterizes reversible ligand-receptor binding:

```
EGFR(loc~m,lig,dim)+EGF(loc~e,rec)<->\
EGFR(loc~m,lig!1,dim).EGF(loc~e,rec!1) EGFkp1,EGFkm1
```

where `EGFkp1` is an association rate constant and `EGFkm1` is a dissociation rate constant. In this rule, a component named `loc` and its internal state are used to denote location. For the `EGF` molecule type, the possible internal states of the `loc` component are `e` (extracellular) and `int` (internalized), which can be seen by inspecting the `molecule types` block in the model-specification file (i.e., the .bngl file that can be processed by BioNetGen). For the `EGFR` molecule type, the possible internal states of the `loc` component are `m` (plasma membrane) and `int` (internalized). The `lig` component of `EGFR` represents the parts of the EGFR ectodomain responsible for ligand binding, which are domains I and III. The `dim` component of `EGFR` represents the part of the EGFR

ectodomain responsible for receptor dimerization, which is domain II. The `rec` component of EGF represents the part of EGF (in the mature soluble form) responsible for receptor binding, which is the EGF-like domain. This rule indicates that monomeric EGFR at the plasma membrane is able to reversibly bind soluble, extracellular EGF. The rule pertains to monomeric EGFR because the left-hand side stipulates that the `dim` component of `EGFR` be free, which is indicated by the absence of association of a bond index with the `dim` component. Note that the backslash character (`\`) indicates line continuation.

The following rule characterizes receptor dimerization:

```
EGFR(loc~m,lig,dim)+EGFR(loc~m,lig,dim)<->\
EGFR(loc~m,lig,dim!1).EGFR(loc~m,lig,dim!1) Dimkp1,Dimkm1
```

where `Dimkp1` is an association rate constant that characterizes all forward reactions implied by the rule and `Dimkm1` is a dissociation rate constant that characterizes all reverse reactions implied by the rule. This rule pertains to receptors that are not bound to EGF, because on the left-hand side the `lig` component in each reactant copy of `EGFR` is not attached to a bond index. The location tags in the rule restrict applicability of the rule to cases where receptors are localized to the plasma membrane. By convention, BioNetGen interprets all rate constants for bimolecular association reactions as single-site rate constants for reactions of the form $A + B \rightarrow$ product(s) (i.e., a user is expected to specify rate constants as if reactants in bimolecular association reactions are distinct and do not contain more than one site that mediates the interaction represented by a rule). BioNetGen automatically adjusts for deviations from this expectation. Here, the left-hand side of the rule is symmetric and therefore implies reactions of the form $A + A \rightarrow$ product whenever both qualifying reactants are identical. When reactants are identical, BioNetGen automatically adjusts for the reaction symmetry by multiplying `Dimkp1` by a factor of ½.

The following rule characterizes autophosphorylation of tyrosine residue Y998 in EGFR:

```
EGFR(lig!+,Y998~Y)->EGFR(lig!+,Y998~pY) k_kinase
```

where `k_kinase` is a pseudo first-order rate constant. As a simplification, autophosphorylation is taken to be a first-order process (vs. a process described by a complex reaction mechanism). This simplification is equivalent to assuming that the local substrate concentration (i.e., the local concentration of the unphosphorylated tyrosine residue) is much less than the Michaelis constant of the EGFR kinase domain. Thus, `k_kinase` is proportional to the ratio k_{cat}/K_M for the EGFR kinase and substrate and inversely related to the volume in which the EGFR kinase domain and the substrate are co-confined. In the rule, the notation `lig!+` indicates that the `lig` component of `EGFR` must be bound for a reaction to occur. The notation `Y998~Y` additionally indicates that the `Y998` component of `EGFR` must be in the `Y` (unphosphorylated) internal state for a reaction to occur.

The following rule characterizes dephosphorylation of Y998 in EGFR:

```
EGFR(Y998~pY)->EGFR(Y998~Y) k_phosphatase
```

where `k_phosphatase` is a pseudo first-order rate constant. In the model, as a simplification, we assume that unshielded phosphotyrosines are dephosphorylated with first-order kinetics whenever free of a binding partner. The rule presented here stipulates that the necessary and sufficient conditions for reaction are that the `Y998` component of `EGFR` be free (as indicated by the absence of an associated bond index) and furthermore that the `Y998` component of `EGFR` must be in the `pY` internal state (i.e., Y998 must be be phosphorylated).

The following pair of rules characterizes the reversible recruitment of a signaling protein containing two SH2 domains to a phosphotyrosine (pY1016) in EGFR:

```
EGFR(Y1016~pY)+PIK3R3(SH2C,SH2N)<->\
EGFR(Y1016~pY!1).PIK3R3(SH2C!1,SH2N) ka_SH2PTB,kdErbB1_1016_PIK3R3C
```

and

```
EGFR(Y1016~pY)+PIK3R3(SH2C,SH2N)<->\
EGFR(Y1016~pY!1).PIK3R3(SH2C,SH2N!1) ka_SH2PTB,kdErbB1_1016_PIK3R3N
```

where `ka_SH2PTB` is an association rate constant (used in both rules) and `kdErbB1_1016_PIK3R3C` and `kdErbB1_1016_PIK3R3N` are dissociation rate constants that are each specific for one of the two SH2 domains in PIK3R3 (the p85γ regulatory subunit of phosphatidylinositol 3-kinase). In the model, as a simplification, when a binding partner of EGFR contains more than one SH2 or PTB domain, we only allow one domain to be bound at a time. Thus, in accordance with this simplification, each of the above rules stipulates that both SH2 domains in PIK3R3 be free in order for an association reaction to occur.

The following rule characterizes internalization when read from left to right and recycling (the reverse process) when read from right to left:

```
EGF(rec!3,loc~e).EGFR(dim!1,lig!3,loc~m).EGFR(dim!1,lig,loc~m)<->\
EGF(rec!3,loc~int).EGFR(dim!1,lig!3,loc~int).EGFR(dim!1,lig,loc~int) kint,krec
```

where `kint` is the effective, pseudo first-order rate constant for internalization and `krec` is the effective, pseudo first-order rate constant for recycling. The effect of this rule, when read in either direction, is to change the location tags of a ligand and two receptors that are in a complex together.

The following rule characterizes degradation of internalized molecules:

```
EGF(rec!2,loc~int).EGFR(lig!2,dim!1,loc~int).EGFR(dim!1,lig,loc~int)->\
Trash() kdeg DeleteMolecules
```

where `kdeg` is the effective, pseudo first-order rate constant for degradation. The left-hand side of this rule refers to receptor dimers in which one receptor is liganded and the other receptor is unliganded. A requirement for prior internalization is indicated by the notation `loc~int`. The term `Trash()` on the right-hand side of the rule is a dummy variable, which serves as a counter. The counter is incremented each time there is a degradation event during a (stochastic) simulation. The term at the end of the rule, `DeleteMolecules`, is a keyword, which is used to control the behavior of a degradation rule (2). Here, it indicates that the rule should be applied even when EGFR is bound to one or more other molecules (e.g., SH2/PTB domain-containing binding partners). This keyword further indicates that only EGF and EGFR molecules are affected by this rule. Thus, SH2/PTB domain-containing binding partners of EGFR are not degraded when the rule is applied. Bonds connecting a copy of EGFR to be degraded to binding partners are removed, releasing the binding partners.

SUPPORTING REFERENCES

1. Chylek, L. A., L. A. Harris, C. S. Tung, J. R. Faeder, C. F. Lopez, and W. S. Hlavacek. 2014. Rule-based modeling: a computational approach for studying biomolecular site dynamics in cell signaling systems. Wiley Interdiscip Rev Syst Biol Med 6:13-36.
2. Faeder, J. R., M. L. Blinov, and W. S. Hlavacek. 2009. Rule-based modeling of biochemical systems with BioNetGen. Methods Mol Biol 500:113-167.

3. Blinov, M. L., J. R. Faeder, B. Goldstein, and W. S. Hlavacek. 2004. BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. Bioinformatics 20:3289-3291.
4. Sneddon, M. W., J. R. Faeder, and T. Emonet. 2011. Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. Nat Methods 8:177-183.
5. Hause, R. J., Jr., K. K. Leung, J. L. Barkinge, M. F. Ciaccio, C. P. Chuu, and R. B. Jones. 2012. Comprehensive binary interaction mapping of SH2 domains via fluorescence polarization reveals novel functional diversification of ErbB receptors. PLOS ONE 7:e44471.
6. Kulak, N. A., G. Pichler, I. Paron, N. Nagaraj, and M. Mann. 2014. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. Nat Methods 11:319-324.
7. Geiger, T., A. Wehner, C. Schaab, J. Cox, and M. Mann. 2012. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. Mol Cell Proteomics 11:M111 014050.
8. Tinti, M., L. Kiemer, S. Costa, M. L. Miller, F. Sacco, J. V. Olsen, M. Carducci, S. Paoluzi, F. Langone, C. T. Workman, N. Blom, K. Machida, C. M. Thompson, M. Schutkowski, S. Brunak, M. Mann, B. J. Mayer, L. Castagnoli, and G. Cesareni. 2013. The SH2 domain interaction landscape. Cell Rep 3:1293-1305.
9. Creamer, M. S., E. C. Stites, M. Aziz, J. A. Cahill, C. W. Tan, M. E. Berens, H. Han, K. J. Bussey, D. D. Von Hoff, W. S. Hlavacek, and R. G. Posner. 2012. Specification, annotation, visualization and simulation of a large rule-based model for ERBB receptor signaling. BMC Syst Biol 6:107.
10. Kozer, N., D. Barua, S. Orchard, E. C. Nice, A. W. Burgess, W. S. Hlavacek, and A. H. Clayton. 2013. Exploring higher-order EGFR oligomerisation and phosphorylation--a combined experimental and theoretical approach. "Molecular BioSystems" 9:1849-1863.
11. Kozer, N., D. Barua, C. Henderson, E. C. Nice, A. W. Burgess, W. S. Hlavacek, and A. H. Clayton. 2014. Recruitment of the adaptor protein Grb2 to EGFR tetramers. Biochemistry 53:2594-2604.
12. Schoeberl, B., E. A. Pace, J. B. Fitzgerald, B. D. Harms, L. Xu, L. Nie, B. Linggi, A. Kalra, V. Paragas, R. Bukhalid, V. Grantcharova, N. Kohli, K. A. West, M. Leszczyniecka, M. J. Feldhaus, A. J. Kudla, and U. B. Nielsen. 2009. Therapeutically targeting ErbB3: a key node in ligand-induced activation of the ErbB receptor-PI3K axis. Sci Signal 2:ra31.
13. Goh, L. K., and A. Sorkin. 2013. Endocytosis of receptor tyrosine kinases. Cold Spring Harb Perspect Biol 5:a017459.
14. Shankaran, H., Y. Zhang, W. B. Chrisler, J. A. Ewald, H. S. Wiley, and H. Resat. 2012. Integrated experimental and model-based analysis reveals the spatial aspects of EGFR activation dynamics. Mol BioSyst 8:2868-2882.
15. Kleiman, L. B., T. Maiwald, H. Conzelmann, D. A. Lauffenburger, and P. K. Sorger. 2011. Rapid phospho-turnover by receptor tyrosine kinases impacts downstream signaling and drug binding. Mol Cell 43:723-737.
16. Moran, U., R. Phillips, and R. Milo. 2010. SnapShot: key numbers in biology. Cell 141:1262-1262 e1261.
17. Panayotou, G., G. Gish, P. End, O. Truong, I. Gout, R. Dhand, M. J. Fry, I. Hiles, T. Pawson, and M. D. Waterfield. 1993. Interactions between SH2 domains and tyrosine-

phosphorylated platelet-derived growth factor beta-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach. Mol Cell Biol 13:3567-3576.
18. Wofsy, C., B. Goldstein, K. Lund, and H. S. Wiley. 1992. Implications of epidermal growth factor (EGF) induced egf receptor aggregation. Biophys J 63:98-110.
19. Macdonald, J. L., and L. J. Pike. 2008. Heterogeneity in EGF-binding affinities arises from negative cooperativity in an aggregating system. Proc Natl Acad Sci USA 105:112-117.
20. Zhou, M., S. Felder, M. Rubinstein, D. R. Hurwitz, A. Ullrich, I. Lax, and J. Schlessinger. 1993. Real-time measurements of kinetics of EGF binding to soluble EGF receptor monomers and dimers support the dimerization model for receptor activation. Biochemistry 32:8193-8198.
21. Low-Nam, S. T., K. A. Lidke, P. J. Cutler, R. C. Roovers, P. M. van Bergen en Henegouwen, B. S. Wilson, and D. S. Lidke. 2011. ErbB1 dimerization is promoted by domain co-confinement and stabilized by ligand binding. Nat Struct Mol Biol 18:1244-1249.
22. Kholodenko, B. N., O. V. Demin, G. Moehren, and J. B. Hoek. 1999. Quantification of short term signaling by the epidermal growth factor receptor. J Biol Chem 274:30169-30181.
23. Fan, Y. X., L. Wong, T. B. Deb, and G. R. Johnson. 2004. Ligand regulates epidermal growth factor receptor kinase specificity: activation increases preference for GAB1 and SHC versus autophosphorylation sites. J Biol Chem 279:38143-38150.
24. Endres, N. F., R. Das, A. W. Smith, A. Arkhipov, E. Kovacs, Y. Huang, J. G. Pelton, Y. Shan, D. E. Shaw, D. E. Wemmer, J. T. Groves, and J. Kuriyan. 2013. Conformational coupling across the plasma membrane in activation of the EGF receptor. Cell 152:543-556.

SUPPORTING TABLES

Tables S1–S3 are Excel spreadsheets, which are provided separately in a ZIP file.

Table S4 Sensitivity of peak EGFR binding to EGF dose

Simulation (1X EGF dose)	Simulation (0.1X EGF dose)	Simulation (10X EGF dose)
SHC1	**SHC1**	**SHC1**
GRB2	**GRB2**	**GRB2**
YES1	YES1	YES1
RASA1	RASA1	RASA1
VAV2	VAV2	VAV2
PTPN11	**PTPN11**	**PTPN11**
PIK3R2	PIK3R2	PIK3R2
PIK3R3	PIK3R3	PIK3R3
SH2B2	SH2B2	SH2B2
PTK6	PTK6	PTK6
SRC	**SRC**	**SRC**
PIK3R1	PIK3R1	PIK3R1
SHB	GRB7	SHB
GRB7	BLK	GRB7
BLK	TEC	BLK
TEC	SHB	TEC
HSH2D	HSH2D	HSH2D
VAV1	VAV1	VAV1

The first column lists the binding partners of EGFR considered in the HeLa cell-specific model of Table S1 by the maximum number bound to EGFR during the simulated response to an EGF dose of 200,000 copies per cell (1X). (This column is identical to the first column of Table 1.) The second column provides the same information, with the exception that a lower (0.1X) dose of EGF was considered in the simulation. The third column reports the results from a simulation in which a higher (10X) dose of EGF was considered.

SUPPORTING SOFTWARE

Software S1 is a ZIP file, which is provided separately. It contains the following plain-text files:

EGFR_G_A549.bngl	EGFR_G_A549.rnf
EGFR_G_GAMG.bngl	EGFR_G_GAMG.rnf
EGFR_G_HEK293.bngl	EGFR_G_HEK293.rnf
EGFR_G_HeLa.bngl	EGFR_G_HeLa.rnf
EGFR_G_HepG2.bngl	EGFR_G_HepG2.rnf
EGFR_G_Jurkat.bngl	EGFR_G_Jurkat.rnf
EGFR_G_K562.bngl	EGFR_G_K562.rnf
EGFR_G_LnCap.bngl	EGFR_G_LnCap.rnf
EGFR_G_MCF7.bngl	EGFR_G_MCF7.rnf
EGFR_G_RKO.bngl	EGFR_G_RKO.rnf
EGFR_G_U2OS.bngl	EGFR_G_U2OS.rnf
EGFR_generic.bngl	EGFR_generic.rnf
EGFR_K_HeLa.bngl	EGFR_K_HeLa.rnf