A reaction-diffusion model predicts the intracellular length scale over which

EGFR-initiated GAB1-SHP2 complexes persist

Paul J. Myers[1,†], Christopher M. Furcht[2,#a,†], William M. Deen[3], and Matthew J. Lazzara[1,4] *

[1]Department of Chemical Engineering, University of Virginia, Charlottesville, Virginia, United States of America

[2]Department of Chemical and Biomolecular Engineering, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America

[3]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America

[4]Department of Biomedical Engineering, University of Virginia, Charlottesville, Virginia, United States of America

[#a]Current Address: Bristol Myers Squibb, Cambridge, Massachusetts, United States of America

[†]These authors contributed equally to this work.

[*]Corresponding author

E-mail: mlazzara@virginia.edu

1

Abstract

Activation of receptor tyrosine kinases (RTKs) leads to the assembly of multi-membered protein complexes connected by phosphotyrosine-SH2 domain linkages. However, these linkages are relatively weak and reversible, which allows complex disassembly to occur on a time scale that permits phosphatases to dephosphorylate complex members and thereby regulate complex persistence. Here, we generated a computational reaction-diffusion model to predict the length scale over which membrane-bound RTKs can regulate the maintenance of such protein complexes through the intermediary action of diffusible cytoplasmic kinases. Specifically, we show that the RTK EGFR can activate SRC family kinases (SFKs) to maintain the association of SHP2 with phosphorylated GAB1, which activates SHP2, throughout the entire cell volume. This finding is dependent on the ability of SFKs to be activated by EGFR at the plasma membrane and subsequently diffuse through the cytosol, as altering the model topology to permit only SFK activation at the plasma membrane reduces the length scale of GAB1-SHP2 association. Modifying the model topology to neglect GAB1 binding to cytosolic and EGFR-bound GRB2 had little effect on this length scale. Indeed, a model sensitivity analysis identified protein diffusion, SFK inactivation, and GAB1 dephosphorylation as the processes that most strongly control the distance over which GAB1-SHP2 persists distal from EGFR. A model scaling analysis likewise predicted that the length scale of GAB1-SHP2 association is greatly extended compared to that of SFK activation and that GAB1-SHP2 complexes persist throughout the cell volume. Furthermore, the same processes identified in the model sensitivity analysis appeared in the length scale estimate for GAB1-SHP2 association. *In vitro* experiments using proximity ligation assay and immunofluorescence against GAB1-SHP2 and EGFR, respectively, suggested that GAB1-SHP2 complexes are distributed throughout cells and exist distally from EGFR during EGF stimulation.

Overall, our results suggest that GAB1-SHP2 complexes—and thus active SHP2—can persist distally from EGFR due to re-phosphorylation of GAB1 throughout the cytosol by EGFR-activated SFKs.

Introduction

Receptor tyrosine kinases (RTKs), such as the epidermal growth factor receptor (EGFR), promote the activity of downstream signaling pathways by initially recruiting SH2 and PTB domain-containing cytosolic adapter proteins to RTK phosphotyrosines. However, receptor-adapter and other phosphorylation-dependent protein complexes exhibit rapid dissociation kinetics (*1*, *2*), allowing protein tyrosine phosphatases (PTPs) to promote irreversible complex disassembly on the time scale of seconds without re-phosphorylation of the appropriate tyrosine residue by a nearby kinase (*3*, *4*). In addition to directly binding cytosolic proteins to facilitate the formation of phosphorylation-dependent multi-membered protein complexes (*5*), RTKs can also activate cytosolic kinases to promote complex assembly, which may serve to counteract the action of cytosolic phosphatases and thereby extend the length scale of complex association distal from RTKs. However, the precise cellular length scale over which RTKs can remotely control the association of these complexes throughout the cytosol is unknown. Here, we use EGFR as a model RTK to investigate the distance from the plasma membrane over which EGFR maintains the association of the cytosolic protein tyrosine phosphatase SRC homology 2 domain-containing phosphatase 2 (SHP2) with the adapter protein GRB2-associated binder 1 (GAB1).

SHP2 is a key signaling intermediate responsible for promoting the activity of ERK downstream of many RTKs and is basally auto-inhibited by intramolecular interactions between its N-SH2 and PTP domains (*6*). Activation of SHP2 downstream of EGFR primarily results from binding of SHP2 to phosphorylated GAB1, which is maintained by EGFR-activated Src family kinases (SFKs) (*7–9*). GAB1-SHP2 association occurs through the binding of SHP2's N- and C-SH2 domains to phosphorylated GAB1 tyrosines 627 and 659, thereby disrupting auto-inhibitory interactions and activating SHP2 (*6*). Binding of SHP2 to GAB1 also permits the redistribution of cytosolic SHP2 to the plasma membrane either through an EGFR-GRB2-GAB1-SHP2 protein

linkage (*10*) or through an alternative mechanism involving GAB1 binding to PIP$_3$(*11*). Similarly, other receptor tyrosine kinases such as Ret, HER2, and c-MET can also recruit SHP2 to the plasma membrane (*12–14*). However, some RTKs may preferentially promote the association of GAB1 with SHP2 at the plasma membrane more than others, as HGF but not EGF promotes substantial recruitment of GAB1 to the plasma membrane (*15*). Indeed, in a lung adenocarcinoma cell line, EGF-initiated GAB1-SHP2 complexes exist primarily in the cytosol, where SFKs can drive GAB1 phosphorylation distal from EGFR (*9*).

As SFKs can be activated at the plasma membrane through events including binding of SRC's SH2 domain to EGFR (*16*), they are inactivated in the cytosol through c-SRC kinase-mediated phosphorylation of negative regulatory tyrosines on SFKs (*17*). Thus, the number of active SFKs able to phosphorylate GAB1 may rapidly decline distal from the plasma membrane. Furthermore, because GAB1 and SHP2 continuously dissociate and re-associate throughout the cytosol due to relatively fast rates of phosphotyrosine-SH2 domain binding/unbinding (*1, 9*)—thus also permitting GAB1 to be dephosphorylated by phosphatases between binding cycles (*9*)—the local cellular concentration of GAB1-bound SHP2 may decline as this complex becomes inaccessible to active SFKs. Ultimately, the precise distance of GAB1-SHP2 complex persistence distal from EGFR depends on how fast SFKs and GAB1-SHP2 complexes diffuse into the cytosol before becoming irreversibly inactivated and dissociated, respectively. Processes such as receptor internalization may also help extend the length scale of GAB1-SHP2 complexes distal from the plasma membrane. This possibility is supported by others who have shown that receptors including EGFR and TrkA remain active within endosomes and can serve as nucleation sites for signaling complex formation (*18–21*).

Previous findings from our group suggest that an important aspect of SHP2-mediated signaling involves the subcellular localization of active, GAB1-bound SHP2 (*22, 23*), although no

prior studies have attempted to explore the spatial distribution of GAB1-SHP2 complexes within a cell. Here, we predict the temporal and spatial profiles of GAB1-SHP2 association within a representative EGF-treated cell using a computational model composed of coupled reaction-diffusion equations that describe the processes that regulate GAB1-SHP2 complex assembly at the plasma membrane and throughout the cytosol. The model predicted that GAB1 and SHP2 were able to remain highly associated distal from EGFR, with the concentration of GAB1-SHP2 complexes at the cell center being 77% of the concentration of GAB1-SHP2 complexes at the plasma membrane. This finding was dependent on EGFR-activated, diffusible SFKs phosphorylating GAB1 throughout the cytosol. Modifying the model topology to permit GAB1 to be phosphorylated only at the plasma membrane resulted in 60% of GAB1-SHP2 complexes being retained near the cell center. A model sensitivity analysis revealed that the total concentration of GAB1-SHP2 complexes was most sensitive to perturbations in the rate constants for SHP2 binding and unbinding to GAB1, while the distance over which GAB1-SHP2 complexes persist distal from EGFR was most sensitive to perturbations in protein diffusivity and the rate constants for SFK inactivation and GAB1 dephosphorylation. Order-of-magnitude estimates likewise predicted that the length scale of GAB1-SHP2 association is greatly extended compared to that of SFK activation, and *in vitro* experiments using proximity ligation assay and immunofluorescence showed that GAB1-SHP2 complexes can be found distal from EGFR. Collectively, our findings suggest that, in the absence of kinase activity distal from the plasma membrane, the extent of GAB1-SHP2 association rapidly decays due to GAB1-SHP2 complex dissociation and GAB1 dephosphorylation.

Model Development

General Model Considerations and Topology

Governing equations for cytosolic species. The model consists of a set of coupled reaction-diffusion equations which describe the processes of protein diffusion, protein binding, and protein phosphorylation that occur among EGFR, SFKs, GRB2, GAB1, and SHP2 in a representative spherical cell with a radius $R = 10$ μm. The resulting ordinary and partial differential equations for surface-associated proteins and cytosolic proteins, respectively, were solved using finite difference methods. Boundary conditions were also approximated using finite difference methods to solve for relevant protein concentrations at the cell surface and center. All reactions were assumed to follow the law of mass action. A general schematic of the model is given in Fig. 1, and all model parameters are summarized in Table 1.

Concentrations of cytosolic species were solved as a function of time (t) and space (r) using governing equations of the following form in spherical coordinates, assuming full rotational symmetry:

$$\frac{\partial C_i}{\partial t} = D_i \frac{1}{r^2} \frac{\partial}{\partial r}\left(r^2 \frac{\partial C_i}{\partial r} \right) + R_{Vi} \tag{1}$$

where C_i corresponds to the concentration of a given cytosolic protein, D_i corresponds to the diffusivity of that protein, and R_{Vi} corresponds to the net volumetric reaction rate of cellular processes leading to the generation or consumption of the cytosolic species designated by C_i.

Boundary conditions for cytosolic species. Neumann boundary conditions at the cell center were implemented for cytosolic species using equations of the form

$$\frac{\partial C_i}{\partial r}(0,t) = 0 \tag{2}$$

7

which corresponds to no flux of the cytosolic species C_i at the cell center. Robin boundary conditions at the plasma membrane were implemented for cytosolic species using equations of the form

$$D_i \frac{\partial C_i}{\partial r}(R,t) = R_{Si} \tag{3}$$

where C_i corresponds to the concentration of a given cytosolic protein, D_i is the diffusivity of that protein, and $R_{S,i}$ corresponds to the net reactive flux of cellular processes leading to the generation or consumption of C_i at the cell surface boundary. The sign on reactive terms comprising R_{Si} depends on whether C_i is being depleted or generated. For SFKs the inward flux of inactive or active SFKs is equal to the reaction for EGFR-catalyzed conversion of inactive SFK to active SFK. Thus, the inward flux for inactive SFKs will be negative because inactive SFKs are being depleted at the surface, while the inward flux for active SFKs will be positive because active SFKs are being generated at the surface. The boundary conditions for all other cytosolic proteins at the cell surface consist of flux conditions where the outward flux is equal to the reaction rate for these proteins binding to EGFR at the cell surface and the inward flux is equal to the reaction rate for these proteins unbinding from EGFR at the cell surface.

Governing equations for membrane-bound species. Concentrations of membrane-associated species were solved for as a function of time (t) only, based on the assumption that membrane-associated species do not diffuse through the plasma membrane or endocytose, using equations of the following form:

$$\frac{dC_{mi}}{dt} = R_{mi} \tag{4}$$

where C_{mi} corresponds to the concentration of a given membrane-associated species and R_{mi} corresponds to the net surface reaction rate of cellular processes leading to the generation or consumption of that membrane species.

Finite difference numerical solution approach. At a given time point, all partial differential equations for cytosolic proteins were discretized using an explicit finite difference method, where the first-order time and second-order space derivatives were approximated using forward and central differences, respectively. The system of algebraic equations which results from discretizing over the entire space domain was then explicitly solved for using the known concentrations from the previous time point. Next, the equations for surface-associated proteins and the boundary conditions for cytosolic proteins, which are coupled through the cell surface boundary, were simultaneously solved using an implicit finite difference method. The discretized boundary conditions for cytosolic proteins were evaluated first, where initial guesses for the concentrations of surface-associated proteins were provided. Subsequently, the discretized equations for surface-associated proteins were evaluated using the concentrations of cytosolic proteins previously solved for. This scheme was iterated by updating the guesses for the concentrations of surface-associated species until convergence was reached.

This system of equations was then advanced to the next time point, where this process was repeated until the final time point was reached. The general method is summarized according to the following steps:

1) All partial and ordinary differential equations for cytosolic and membrane-associated species, respectively, were discretized using finite difference methods.

2) Concentrations of all protein species were defined for the initial time.

3) The model was advanced one time step, and the concentrations of all cytosolic species were solved for in the bulk (i.e., everywhere except the cell center and cell surface boundaries).

4) At the same time step, the concentrations of all cytosolic species at the cell center were solved for using the appropriate boundary conditions at the cell center.

5) At the same time step, the concentrations of all cytosolic species at the cell surface were solved for using the appropriate boundary conditions at the cell surface, where initial guesses for concentrations of all membrane-associated species were provided.

6) The concentrations of all membrane-associated species were solved for using the concentrations of cytosolic species at the cell surface calculated in Step 5.

7) Steps 5 and 6 were repeated, where the initial guesses for the concentrations of membrane-associated species in Step 5 were updated using the concentrations calculated in the previous iteration of Step 6. This process was iterated until the concentrations obtained in Steps 5 and 6 converged.

8) Steps 3-7 were repeated until the final timepoint was reached.

For additional details and examples regarding the finite difference approach outlined here, please refer to S1 Appendix.

Simplified, Steady-State Model Description

Additional model calculations were performed using simplified sets of model equations in which EGFR and GRB2 binding events were neglected. Steady-state solutions for active and inactive SFKs were obtained by setting the concentration of phosphorylated EGFR at the cell membrane to the value obtained from the full set of reaction-diffusion equations after 5 minutes of simulation time and then solving the resulting system of equations for active and inactive SFKs. The final solution for active SFKs is given by

$$C_{SFK,a}(r) = \frac{k_{S,a} C_{pE,tot} C_{o,SFK}}{D_{SFK}} \left[\frac{\cosh mR}{R} + \frac{\sinh mR}{mR} \left(\frac{k_{S,a} C_{pE,tot}}{D_{SFK}} - \frac{1}{R} \right) \right]^{-1} \frac{\sinh mr}{mr} \qquad (7)$$

where $k_{S,a}$ is the rate constant of SFK activation, D_{SFK} is the diffusivity of SFKs, C_{pE} is the concentration of phosphorylated EGFR, $C_{o,SFK}$ is the total concentration of SFKs, R is the cell radius, and $m = (k_{S,i}/D_{SFK})^{1/2}$, where $k_{S,i}$ is the rate constant of SFK inactivation. The solution for inactive SFKs is given by $C_{SFK,i} = C_{o,SFK} - C_{SFK,a}$.

Without binding to EGFR and GRB2 at the cell membrane, the boundary conditions for the four remaining cytosolic species—GAB1, pGAB1, pGAB1-SHP2, and SHP2—are all of the form

$$\frac{\partial C_i}{\partial r}(0,t) = \frac{\partial C_i}{\partial r}(R,t) = 0 . \qquad (8)$$

Using Eq. 8, the reaction-diffusion equations for these four species were solved using a modified, steady-state finite difference method without time steps (see S2 Appendix for more details).

The governing equations for these four species were then simplified further to two equations for pGAB1 and pGAB1-SHP2 using conserved scalar quantities. The resulting equations were again solved using the modified finite difference scheme. Finally, the system was reduced to a single governing equation for pGAB1 by assuming that GAB1 and SHP2 binding had reached equilibrium and substituting to eliminate all other species from the differential equation. The resulting second-order differential equation for pGAB1 was once again solved using the modified finite difference scheme. See S2 Appendix for additional details regarding the model topology simplifications, the associated governing equations, and the steady-state finite difference approach.

Derivations of Length Scales for SFK Activation and GAB1-SHP2 Complex Association

11

Order-of-magnitude estimates were obtained for the length scales over which large changes in the concentrations of active SFKs and GAB1-SHP2 complexes are expected to occur. The analysis described here was performed assuming the system had reached equilibrium and steady state. An order-of-magnitude equality, denoted by the symbol "~", means that the quantities in an order-of-magnitude expression can be expected to deviate by roughly a factor of 3 up or down from exact equality. The scale of a quantity denotes the maximum order of magnitude that quantity is expected to take on. Thus, the length scales sought here represent the maximum distances—accurate to within a factor of 3—over which large changes in the concentrations of active SFKs and GAB1-SHP2 complexes are expected to occur. Additionally, the length scale obtained for GAB1-SHP2 was assumed to apply to both GAB1-SHP2 and GRB2-GAB1-SHP2 since the diffusivities of these two species are nearly identical (Table 1) and because GRB2 binding does not prevent GAB1-SHP2 association in the current model topology. See S3 Appendix for additional details regarding these derivations.

Derivation of SFK activation length scale. The length scale of SFK activation can be derived from the steady-state governing equation for active SFKs, which is written in spherical coordinates assuming full rotational symmetry as

$$D_{SFK} \frac{1}{r^2} \frac{d}{dr}\left(r^2 \frac{dC_{SFK,a}}{dr} \right) - k_{S,i} C_{SFK,a} = 0 \tag{9}$$

where D_{SFK} is the diffusivity of SFKs, $C_{SFK,a}$ is the concentration of active SFKs, $k_{S,i}$ is the rate constant of SFK inactivation, and r is the radial coordinate denoting the distance from the cell center. $C_{SFK,a}$ and r can be rendered dimensionless as

$$\theta = \frac{C_{SFK,a}}{C_{o,SFK}} \tag{10}$$

12

$$\eta = \frac{r}{R} \tag{11}$$

where $C_{o,SFK}$ is the total concentration of SFKs, R is the cell radius, θ is the dimensionless concentration of active SFKs, and η is the dimensionless radial coordinate. Substituting Eqs. 10 and 11 and their differentials into Eq. 9 and rearranging gives

$$\frac{1}{\eta^2}\frac{d}{d\eta}\left(\eta^2\frac{d\theta}{d\eta}\right) = \mathrm{Da}\,\theta . \tag{12}$$

Da, the Damköhler number, describes the competition between SFK inactivation and diffusion and is given by $\mathrm{Da} = \left(k_{S,i}R^2 / D_{SFK}\right)$.

In scaling analyses, equations are considered properly scaled when all field variables and their derivatives have been divided by their scales. The concentration of active SFKs cannot exceed the total concentration of SFKs, so the dimensionless concentration θ is properly scaled. However, for large Da, large changes in the concentration of active SFKs will occur over a much shorter distance than the cell radius R. This means that the dimensionless radial coordinate η has not been properly scaled. Proper scaling can be achieved by eliminating Da from Eq. 12 by rescaling η as

$$\chi = \eta\,\mathrm{Da}^{1/2} = r\left(\frac{k_{S,i}}{D_S}\right)^{1/2} \tag{13}$$

where χ is now the properly scaled dimensionless radial coordinate. From Eq. 13, the true length scale of SFK activation, denoted $L_{SFK,a}$, is given by

$$L_{SFK,a} \sim \left(\frac{D_S}{k_{S,i}}\right)^{1/2} . \tag{14}$$

Derivation of GAB1-SHP2 association length scale. While $L_{SFK,a}$ is readily derived by scaling the governing equation for active SFKs, the length scale for GAB1-SHP2 persistence is significantly more difficult to obtain with this approach due to the presence of multiple kinetic

13

terms in the GAB1-SHP2 governing equation. However, an alternative approach is to develop length scales associated with the reaction-diffusion processes that contribute to the formation and dissolution of GAB1-SHP2 complexes and combine those length scales. As will be verified with model calculations, these processes are SFK-mediated GAB1 phosphorylation, GAB1 and SHP2 dissociation, and GAB1 dephosphorylation. Over $L_{SFK,a}$, GAB1 is phosphorylated and subsequently binds to SHP2. Beyond $L_{SFK,a}$, GAB1 cannot be rephosphorylated once it unbinds from SHP2. Using the form of $L_{SFK,a}$ as a guide, the length scale associated with SHP2 unbinding from phosphorylated GAB1 can be estimated by

$$L_{uncoup} \sim \left(\frac{D_{G1S2}}{k_{S2,off}} \right)^{1/2} \tag{15}$$

where D_{G1S2} is the diffusivity of GAB1-SHP2 and $k_{S2,off}$ is the unbinding rate constant of SHP2 from GAB1. As long as GAB1 remains phosphorylated, it can rebind with SHP2 to reform GAB1-SHP2. The length scale of complex association is thus extended further by the length scale of GAB1 dephosphorylation, which can be estimated as

$$L_{dephos} \sim \left(\frac{D_{G1}}{k_{G1dp}} \right)^{1/2} \tag{16}$$

where D_{G1} is the diffusivity of GAB1 and k_{G1dp} is the rate constant of GAB1 dephosphorylation. Thus, the length scale of GAB1-SHP2 association, L_{G1S2}, is given by the sum of Eqs. 14-16,

$$L_{G1S2} \sim \left(\frac{D_{SFK}}{k_{S,i}} \right)^{1/2} + \left(\frac{D_{G1S2}}{k_{S2,off}} \right)^{1/2} + \left(\frac{D_{G1}}{k_{G1dp}} \right)^{1/2}. \tag{17}$$

Model Implementation

Model codes were written and compiled in MATLAB 2020b and Julia v1.6.1 and are available on GitHub under the username "pauljmyers". For all finite difference calculations, except

those for sensitivity analyses, the spatial dimension was discretized using a spatial step of 0.1 µm. For sensitivity analysis calculations, the spatial dimension was discretized with a spatial step of 0.2 µm to reduce computation times. For all finite difference calculations, the time step was set based on the explicit finite difference stability criterion described in Eq. 8 by Bieniasz (*24*) for diffusion problems with homogeneous reaction terms.

An additional implementation of the model was made using Virtual Cell (VCell), a freely available application designed for modeling cellular biology (*25*, *26*). The VCell implementation of the model, "Furcht&Myers _spatial_GAB1_SHP2", was constructed as a BioModel in VCell 7.1 and is available in the public domain through the Public BioModels portal in VCell and at http://vcell.org/vcell-models under the shared username "pmyers1995". In total, the VCell model consists of 18 distinct protein species (10 volumetric and 8 membrane-bound) and 25 reactions. Membrane species were defined as well-mixed since the diffusion of proteins on the cell membrane was not considered in this study. All model parameters were implemented as described in Table 1. Solution geometries were constructed in VCell by defining the cytoplasmic volume compartment as a sphere with a radius of 10 µm or 50 µm. All calculations were made for the first five minutes of EGF stimulation using VCell's fully-implicit, adaptive time step finite volume solver.

Representative cell. EGFR, GRB2, GAB1, SHP2, and SFKs were assumed to be at cellular concentrations of 6×10^5 species per cell, as described in our previous reaction-only model (*9*). The cell radius was assumed to be 10 µm, as described previously for other epithelial cells (*27*), for all calculations except when length scales for protein species persistence were directly calculated for comparison against order-of-magnitude length scale estimates. For those simulations the cell radius was set to 50 µm. Total concentrations of EGFR, SFKs, GRB2, GAB1, and SHP2 were converted to units of molecules/µm^3 using the appropriate volume and surface area conversion factors such

15

that the total number of molecules of each protein per cell did not change between calculations whether the cell radius was set to 10 μm or 50 μm.

EGFR phosphorylation. The process of EGFR phosphorylation was simplified to sequentially depend on the processes of EGF binding, EGFR dimerization, and EGFR phosphorylation. EGF binding at the plasma membrane was modeled as a reversible process characterized by association (*28*) and dissociation (*29*) rate constants. EGF was modeled at a constant concentration of 10 ng/mL. The EGFR dimerization rate constant was calculated as described previously (*30*). Dimer uncoupling rate constants in the presence of EGF were described previously (*9*). All dimer species were assumed to be symmetric. EGFR phosphorylation was modeled as a process which occurs between EGF-bound EGFR dimers, where both receptors are simultaneously phosphorylated at a representative tyrosine (Y1068) that is able to bind GRB2. EGFR dephosphorylation was modeled as a zeroth order process with a previously described rate constant (*9*).

GAB1 phosphorylation. GAB1 phosphorylation at a representative tyrosine (Y627) which is able to bind SHP2 was modeled as a process catalyzed by active SFKs throughout the cytosol. GAB1 dephosphorylation was modeled as zeroth order and occurred throughout the cytosol.

SFK activation. Similar to our previous model (*9*), the process of SFK activation was modeled as a first-order rate equation where phosphorylated EGFR activates SFK, which only occurs at the cell surface boundary for this model. SFK inactivation was modeled as zeroth order and occurred throughout the cytosol.

GRB2, GAB1, and SHP2 binding. Reactions for GRB2, GAB1, and SHP2 association/dissociation were described previously (*9*).

Phosphatase activity. EGFR and GAB1 dephosphorylation were modeled as zeroth order with respect to protein tyrosine phosphatases, which precludes the requirement for considering distinct phosphatase species.

Protein diffusivity. Diffusivities for each cytosolic protein monomer or complex were calculated based on the diffusivity of tubulin (*31*), which was adjusted based on differences in the hydrodynamic radii of tubulin and a given protein monomer or complex (*32*). See S4 Appendix for additional details.

Parameter fitting. Due to topological differences between this model and our previously developed model, which did not include spatial considerations, we chose to refit parameters for SFK activation/inactivation and GAB1 phosphorylation/dephosphorylation to match our previous model's predictions for concentrations of phosphorylated GAB1 and active SFKs in response to EGF treatment. The four relevant parameters (k_{G1p}, k_{G1dp}, $k_{S,a}$, $k_{S,i}$) were determined by computing the spatial average of [pGAB1] and [aSFK] in response to EGF and fitting these concentrations to prior predictions for [pGAB1] and [aSFK]. Particle swarm optimization was used to fit the unknown rate constants using the *particleswarm* function in MATLAB to minimize the total error between the current reaction-diffusion model's outputs and results from our previous model. The error was computed as the square of the difference between the outputs from the reaction-diffusion model output and the purely dynamic model. The best-fit results are included in Table 1.

Sensitivity analysis. Model sensitivity to simultaneous changes in parameter values was computed using a Latin hypercube sampling-based (LHS) global sensitivity analysis (*33*). LHS parameter combinations were generated in Julia using the QuasiMonteCarlo.jl package by allowing parameters to vary by up to a factor of 10 above or below their base values. Model solutions were then generated for each LHS-generated parameter combination, and model sensitivities were calculated as the Spearman rank correlation coefficient between model outputs and the LHS-

17

generated distributions for each parameter. The resulting Spearman ranks for a given model output

were then divided by the absolute value of the highest-magnitude rank for that output to obtain

normalized model sensitivities (i.e., normalized sensitivities on the range of -1 to 1 for each output).

Results

Base Model Predictions

We developed a computational model to predict the concentration of GAB1-SHP2 complexes as both a function of time and position for the first 5 minutes of EGFR activation following EGF binding within a spherical cell with an assumed radius of $R = 10$ μm. We limited our simulations to the first 5 minutes of EGFR activation since the peak of receptor activation and adaptor binding occurs within this early time period (*34*, *35*) and to avoid the intricacies involved in modeling EGFR internalization, trafficking, and degradation processes that become prominent after this time period(*5*, *36–38*). The model considers cellular processes including EGFR phosphorylation, SFK activation, GAB1 phosphorylation, GAB1-SHP2 binding, and diffusion of cytosolic proteins, the full details of which are described in Model Development. Model parameters were taken from a previous model we developed (*9*) except for parameters for SFK activation and inactivation and GAB1 phosphorylation and dephosphorylation, which were fit due to topological differences between the models. The rate constants for these processes were fit to predictions from our previous model to preserve similar concentrations of active SFKs and phosphorylated GAB1. The model thus fit is able to accurately recapitulate the previously predicted dynamics of SFK activation and GAB1 phosphorylation in response to EGF (Fig. 2A). Consistent with our previous model's finding of amplification of SFK activity downstream of EGFR, this model predicts that each active EGFR molecule leads to the phosphorylation of ~2 GAB1 molecules on average throughout a 5-minute EGF treatment simulation (Fig. 2B). When calculating protein concentrations as a function of distance from the cell surface, we found that the concentration of active SFKs rapidly decayed throughout the cytosol while the concentration of phosphorylated GAB1 remained relatively unchanged (Figs. 1C-D). We also note that the concentration of GAB1-

SHP2 complexes was essentially unchanged as a function of position within the cell, suggesting that EGFR-activated SHP2 retains its activity distal from the plasma membrane (Fig. 2E). We tested the robustness of this prediction by creating random parameter sets in which each parameter was randomly perturbed by up to an order of magnitude above or below its base value. For 91%, 96%, and 97% of the 1,000 parameter sets, the length scale for the concentration of GAB1-SHP2 complexes to reach 50%, 20%, and 10% of the maximal GAB1-SHP2 concentration at the cell surface, respectively, was greater than the cell radius R (Fig. 2F). Therefore, the model prediction of GAB1-SHP2 complexes being maintained distal from the cell surface appears robust even if the model parameters are not well estimated.

Steady-state model calculations. Model predictions were also made using several sets of simplified model equations in which GRB2 and EGFR binding were neglected. These calculations were made at steady-state after setting the total concentration of phosphorylated EGFR to the value from the full reaction diffusion model after 5 minutes of EGF treatment. Analytical solutions were obtained for active and inactive SFKs, and the concentrations of the remaining cytosolic species were calculated numerically. See Materials and Methods for additional details. Analytical solutions for active and inactive SFKs matched the solutions from the full reaction-diffusion model nearly identically (Fig. S1). Steady-state solutions from the simplified model topologies for GAB1-SHP2, pGAB1, GAB1, and SHP2 agreed with results from the full reaction-diffusion model to within ~15% error or less for each of these species (Fig. S1). These steady-state results further suggest that GAB1 binding to GRB2 has little influence on cytosolic GAB1-SHP2 formation since neglecting this interaction had very little effect on the length scale of complex persistence from the cell surface. Collectively, the predictions in Figures 1 and S1 predict that GAB1-SHP2 complexes are rapidly formed and distributed throughout the cell volume in the absence of EGFR downregulation by internalization and degradation.

Model predictions using Virtual Cell. Calculations were made with a Virtual Cell (VCell) implementation of the full reaction-diffusion model to further verify the prediction of GAB1-SHP2 maintenance. The VCell predictions of SFK activation and GAB1 phosphorylation again agreed well with results from the purely dynamic model (Fig. S2) (*9*). VCell predictions for the percentage of phosphorylated EGFR and the distribution of GAB1-SHP2 complexes were also essentially identical to those obtained from the MATLAB implementation of the model (Fig. S2). Importantly, both versions of the model predict that GAB1-SHP2 complexes are expected to persist at nearly a constant concentration throughout the volume of the cell.

Model Sensitivity Analysis

To identify model processes which strongly influence both the concentration and spatial distribution of GAB1-SHP2 complexes, we performed a Latin hypercube sampling-based (LHS) global sensitivity analysis for GAB1-SHP2 complex formation in response to 5 minutes of EGF treatment when allowing model parameters to simultaneously vary by up to a factor of ten in the LHS generation step (Fig. 3A). We found that the average GAB1-SHP2 concentration was most sensitive to perturbations to the rate constants for SHP2 binding to ($k_{S2,f}$) and unbinding from GAB1 ($k_{S2,r}$). To determine parameters which control the spatial distribution of GAB1-SHP2 complexes, we calculated the ratio of the concentration of GAB1-SHP2 complexes at the cell center to the cell surface and defined the length scale $r_{1/2}$ as the distance from the cell surface where the concentration of GAB1-SHP2 reaches 50% of its membrane value. Model predictions for $r_{1/2}$ were most sensitive to changes in the diffusivity of SFKs and the diffusivities of protein complexes containing GAB1 or SHP2, changes in the kinetic rate constants for GAB1-SHP2 binding and unbinding, and changes in the kinetic rate constants for GAB1 dephosphorylation (k_{G1dp}) and SFK inactivation ($k_{S,i}$). The ratio of GAB1-SHP2 at the cell center versus cell surface was similarly sensitive as $r_{1/2}$ to the same perturbations. In agreement with our reduced model calculations (Fig. S1), perturbations to the

21

kinetic rate constants for reactions involving EGFR and GRB2 had little effect on GAB1-SHP2 complex abundance or persistence in comparison to perturbations of other model processes. Overall, the model was most sensitive to coordinated changes in protein diffusivities compared to the other important processes identified by the global sensitivity analysis (Fig. 3B-E).

We next examined the effects of simultaneously altering the initial concentrations of protein species and kinetic parameters on the model output. Perturbations to the concentration of SFKs alone resulted in relatively small changes to the distribution of GAB1-SHP2 complexes (Fig. 3F). Given that we identified the processes of SFK inactivation and GAB1 dephosphorylation as strong determinants of the GAB1-SHP2 length scale, we sought to determine the relationship between SFK concentration and the rates of these two processes in determining the GAB1-SHP2 length scale. As noted previously, a ten-fold increase in the rate constant for GAB1 dephosphorylation greatly reduces the GAB1-SHP2 length scale, but a simultaneous ten-fold increase in SFK concentration somewhat negated the effect of an increased GAB1 dephosphorylation rate constant (Fig. 3G). Similarly, a ten-fold increase in the rate constant for SFK inactivation also greatly reduces the GAB1-SHP2 length scale, but an increase in SFK concentration was unable to overcome a faster rate of SFK inactivation.

Based on previous findings from our lab and others (9, 39), we also sought to determine if SHP2's binding to GAB1 could protect GAB1 from being dephosphorylated throughout the cytosol. While our previous model, which did not consider spatial distributions of proteins, found very little effect of SHP2 depletion or overexpression on GAB1 phosphorylation, the current model predicts that both SHP2 depletion and overexpression can greatly change the extent of GAB1 phosphorylation in the cytosol (Fig. 3H). The total percentage of phosphorylated GAB1 varied greatly in these simulations, ranging from ~98% phosphorylated when SHP2 was overexpressed to as low as ~48% phosphorylated when SHP2 was depleted. This occurs because unbound and

dephosphorylated GAB1 molecules near the cell center are in proximity of a relatively low concentration of active SFKs compared to the cell surface, which magnifies the protective effect of SHP2's binding on phosphorylated GAB1 since there fewer SFKs available to rephosphorylate GAB1 after its dephosphorylation upon dissociation from SHP2.

EGFR Pulse-Chase Dynamics

To determine the consequences of EGFR inhibition on the spatiotemporal profile of GAB1-SHP2 complexes, we simulated a pulse-chase experiment where EGF was added for 5 min, and afterwards the rate constant for EGFR phosphorylation was set to zero. We verified that this pulse-chase topology produced EGFR dephosphorylation kinetics similar to those achieved by our previous model, which explicitly included the EGFR inhibitor gefitinib in the model topology (Fig. 4A; (9)). When predicting the concentration of GAB1-SHP2 complexes as a function of cell position throughout a 2-minute EGFR inhibition chase, we noted that this concentration decayed most rapidly near the cell surface throughout the chase (Fig. 4B). A similar and even more pronounced trend was observed for the rate of SFK inactivation (Fig. 4C). Since our previous predictions suggest that both GAB1-SHP2 and active SFK concentrations are highest near the cell surface (Figs. 1B and D), these data suggest that the net rate of GAB1-SHP2 dissociation and SFK inactivation are largest where these proteins and protein complexes are most highly concentrated. This result has interesting implications related to protein complex cycling, as this suggests that the total number of GAB1-SHP2 dissociation and re-association events, as well as GAB1 dephosphorylation and re-phosphorylation events, is largest closer to the cell surface over a given time interval.

Model Extensions

While our previous experimental data suggests that a significant fraction of SFKs exist within the cytosol (9), where in some cases SFKs have been shown to be phosphorylated at the

active site (32), other reports suggest that SFKs may only exist in an active state primarily at the plasma membrane due to SFK palmitoylation (33). To see how this consideration would affect our model's prediction of the cellular GAB1-SHP2 length scale, we altered our model topology to only allow SFKs to exist in an active state at the plasma membrane by setting the diffusivity of active SFKs to nearly zero. This change caused GAB1-SHP2 association to drop by ~40% at the cell center compared to the cell surface (Fig. 5A), although the difference compared to the cell center concentration of GAB1-SHP2 with the base model (Fig. 1E) was modest. Still, the distribution of GAB1-SHP2 complexes more closely mirrored that for active SFKs from the base model topology (Fig. 2C). Additionally, the model parameters can be adjusted for this topology to still generate the same GAB1-SHP2 concentration gradient predicted by the base model, such as by increasing protein diffusivity or decreasing the GAB1 dephosphorylation rate constant (Figs. 4B-C).

Order-of-Magnitude Length Scale Estimates

To further validate our current model's predictions, we developed an order-of-magnitude estimate for the length scale of GAB1-SHP2 association based on adding individual length scales for the reaction-diffusion processes that are most important for complex association (Fig. 6A). These processes, which were identified independently of model calculations, included the same parameters to which the model predictions for the persistence of GAB1-SHP2 were most sensitive (Fig. 3). We compared the order-of-magnitude length scale estimates for active SFKs (Eq. 14) and GAB1-SHP2 (Eq. 17), denoted collectively as L_{OM}, against the model-predicted length scales $r_{1/2}$ and $r_{1/10}$, the latter of which is the length at which the concentration of a protein or protein complex falls to 10% of its maximal value at the cell membrane. Because $r_{1/2}$ and $r_{1/10}$ values were greater than or equal to the cell radius ($R = 10 \, \mu m$) in the base model, the cell radius was set to 50 μm for calculations of the model-predicted length scales so that values for $r_{1/2}$ and $r_{1/10}$ greater than the base cell radius R could be obtained. Calculations of $r_{1/2}$ and $r_{1/10}$ were also made with the VCell

implementation of the model to confirm the predictions made with the finite difference version (Fig. S2). L_{OM} and $r_{1/2}$ values were calculated with base model parameters and when the parameters appearing in Eq. 17 were individually varied by a factor of ten up or down from their base values (Fig. 6B). L_{OM} and $r_{1/2}$ predictions agreed well for both active SFKs and GAB1-SHP2, and in all scenarios the L_{OM} for active SFKs was less than the base cell radius R. There was less agreement between L_{OM} and $r_{1/10}$ values, but the error between any pair of L_{OM} and $r_{1/10}$ values was less than or equal to a factor of 3 in all cases (Fig. 6C). This was true even when model parameters were varied by a factor of ten, thereby indicating that Eqs. 14 and 17 accurately describe the cellular processes that are most important for determining the length scales of SFK activation and GAB1-SHP2 association. Furthermore, the scaling analysis predicts L_{OM} values of ~5.4 µm for GAB1-SHP2 and ~1.6 µm for active SFKs. The three-fold difference between these values agrees with the model prediction that GAB1-SHP2 complexes persist over a much greater distance from the cell membrane relative to active SFKs (Fig. 2, C and E). Since the curvature of the cell is not explicitly considered in our length scale estimates, the L_{OM} values calculated here represent the minima for the length scales of SFK and GAB1-SHP2 persistence. For cells smaller than 50 µm in radius, these length scales can be expected to be even higher as curvature effects of the cell membrane become more apparent, as predicted with the base model (Figs. 1 and 2). Thus, for a representative cell of radius 10 µm, the length scale estimates predict that GAB1-SHP2 complexes, and therefore active SHP2, will persist from the cell membrane over an intracellular distance that is at minimum on the same order of magnitude as the cell radius during EGFR-initiated cell signaling.

Visualizing GAB1-SHP2 complexes with proximity ligation assay

We next sought to visualize the localization of GAB1-SHP2 complexes in cells relative to both the cell membrane and EGFR. To do this, we used previously described HeLa cells that have

25

been gene-edited to express endogenous levels of the fusion protein mVenus-HRAS (referred to hereafter as HeLa/mV-HRAS) (*36*), which allowed us to visualize cell membranes without additional staining of membrane components. Model calculations using HeLa cell abundances (*42*) suggested that GAB1-SHP2 complexes should be distributed throughout the interior of these cells, albeit at lower concentrations compared to the base model predictions (Fig. S3). To test this prediction and to visualize GAB1-SHP2 complexes in HeLa/mV-HRAS cells, we used proximity ligation assay (PLA)—which has been previously used to study the spatiotemporal regulation of SHP2 oxidation (*43*)—coupled with standard immunofluorescence (IF) to analyze endogenous levels of GAB1-SHP2 and EGFR without engineering fluorescent-protein fusions of these proteins.

We first validated the use of PLA for determining the subcellular localization of protein signaling complexes in our system by staining EGF-treated HeLa/mV-HRAS cells for EGFR-GRB2 complexes via PLA and EGFR via IF (Fig. S4A). Since EGFR and GRB2 associate upon EGFR phosphorylation, we reasoned that EGFR-GRB2 PLA signals should be almost entirely colocalized to independent EGFR signals observed via IF. EGF treatment resulted in a robust induction of EGFR IF signal and EGFR-GRB2 PLA signals that were visible throughout cells (Fig. S4B-C). PLA signals seen with primary antibody controls indicated that some of the apparent basal association of EGFR and GRB2 observed in serum-starved cells could be attributed to non-specific binding of the PLA probes. Most importantly, colocalization analysis of EGFR-GRB2 PLA and EGFR IF puncta indicated that nearly all EGFR-GRB2 were colocalized with EGFR signals (Fig. S4D-E). This result gave us confidence in using PLA to visualize the subcellular localization of protein complexes relative to other fluorescent signals in our cell system.

We then performed PLA and IF against GAB1-SHP2 complexes and EGFR, respectively, to visualize their localization within HeLa/mV-HRAS cells (Fig. 7A). As expected, EGF treatment resulted in a robust induction of EGFR IF and GAB1-SHP2 PLA signals that were

26

visible throughout the cells (Fig. 7B-C). Primary antibody controls indicated that the presence of

PLA signal in serum-starved (untreated) cells is likely due to a combination of low basal

association of GAB1 and SHP2 and non-specific binding of the PLA anti-rabbit probe to epitopes

in proximity to the SHP2 primary antibody. Furthermore, the increase in GAB1-SHP2 PLA

signal as a result of EGF treatment is consistent with our previous biochemical measurements

that showed that GAB1 and SHP2 associate in response to EGF (*9*). We then determined the

fraction of GAB1-SHP2 PLA puncta colocalized with EGFR signal and saw that the fraction of

colocalized puncta decreased in response to EGF (Fig. 7D). This difference was most pronounced

after 15 minutes of EGF treatment when the majority of GAB1-SHP2 signals were not

colocalized with EGFR signals, which is consistent with our previous observations that GAB1-

SHP2 complexes are primarily cytosolic (*9*). Additionally, GAB1-SHP2 PLA signals can be seen

both near the cell membrane (near mV-HRAS signal) and throughout the cell interior.

Furthermore, the PLA results presented here provide experimental support for our model and

length scale predictions that GAB1-SHP2 complexes are distributed throughout the cytosol of

cells in response to EGF.

Discussion

Our analysis reveals that GAB1 and SHP2 are predicted to remain associated throughout the cytosol to a significant degree due to the capacity for SFKs to maintain cytosolic GAB1 phosphorylation throughout GAB1 phosphorylation/dephosphorylation cycles. We note that this conclusion is based in part on this model's indirect parameterization to previous experimental data gathered from a representative lung adenocarcinoma cell line expressing wild-type EGFR (*9*) and that this conclusion may not be applicable to EGFR-mediated induction of GAB1-SHP2 association in all cell backgrounds. A previous study noted impaired phosphorylation of the SFK member SRC at Y418 (*44*), an autophosphorylation site required for SRC activity, in an NSCLC cell line expressing a kinase-activated EGFR mutant, suggesting that SFK activity may be impaired in this setting. Interestingly, these cells still exhibited GAB1 phosphorylation on par with wild-type EGFR-expressing cells that require SFKs for GAB1 phosphorylation (*9*, *44*). Thus, there may be a different mode of GAB1 phosphorylation in the context of EGFR mutation, potentially direct phosphorylation of GAB1 by constitutively active mutant EGFR. Based on our model predictions which resulted when GAB1 phosphorylation only occurred at the plasma membrane (Fig. 5A), this mechanism would decrease the length scale over which phospho-GAB1 and GAB1-SHP2 complexes extend into the cytosol. This possibility may offer additional insight into the impairment of SHP2's function observed in cells with EGFR mutation, as a previous study from our group revealed that SHP2 function is impaired in mutant EGFR-expressing cells through apparent sequestration of active SHP2 at the plasma membrane with internalization-impaired EGFR mutants (*23*). In addition, EGFR internalization could also be required to initiate the activity of SFKs from within the cytosol at endosomes containing EGFR, as SFKs have been shown to traffic with activated EGFR in response to EGF (*45*).

Our analysis revealed that the GAB1-SHP2 length scale was sensitive to perturbations in the rate constants for GAB1-SHP2 binding and unbinding, SFK inactivation, and GAB1 dephosphorylation, but interestingly was much less sensitive to the rate constants for SFK activation and GAB1 phosphorylation (Fig. 3A). In fact, the rate constant for SFK activation had essentially no control over the steady state cellular distribution of GAB1-SHP2 complexes. This is likely the result of the model topology only permitting SFKs to be activated at the cell surface by membrane-bound EGFR, as SFKs will still be rapidly inactivated throughout the cytosol even if the surface concentration of active SFKs is increased or reaches steady state more quickly. The appearance of the rate constant for SFK inactivation but not activation in our order-of-magnitude estimate for active SFKs (Eq. 15) further supports this hypothesis. This is consistent with our finding that EGFR only weakly controls the spatial gradient of GAB1-SHP2 complexes within the cell (Fig. 3A), despite the requirement for surface-associated EGFR to activate cytosolic SFKs. Our parameter analysis also revealed that an increased concentration of SFKs can offset a reduction in the GAB1-SHP2 length scale caused by an enhanced rate of GAB1 dephosphorylation but cannot offset a reduction in this length scale caused by an enhanced rate of SFK inactivation, highlighting the complex relationships among cellular processes that influence the spatial distribution of GAB1-SHP2.

Characteristic GAB1-SHP2 reaction-diffusion time scales suggest that GAB1 needs to be phosphorylated ~3-4 times by SFKs throughout the cytosol to permit an individual GAB1-SHP2 complex to diffuse to the cell center from the cell surface (9). While our model predicts that this does occur—since the concentration of GAB1-SHP2 complexes remains relatively unchanged at the cell center relative to the cell surface—the concentration of active SFKs is substantially reduced distal from the plasma membrane. However, even this relatively low concentration of active SFKs is sufficient to permit the necessary number of GAB1 phosphorylation events required to maintain

29

a much higher extent of GAB1 phosphorylation and GAB1-SHP2 association in proximity to the cell center. This observation is consistent with findings from our previous study which showed that SFKs amplify EGFR activity to maintain GAB1-SHP2 association at low levels of EGFR phosphorylation (*9*).

While our model utilized parameters either taken from literature or fit to predictions from a model we previously developed (*9*), the refinement and verification of this model's parameters will be aided by the generation of experimental data specifying the spatial distribution of GAB1-SHP2 complexes within a cell, such as data which could be obtained from FRET microscopy. While no such data currently exists, our previous findings show that GAB1-SHP2 complexes do exist primarily in the cytosol rather than membrane compartments (*9*). It would also be beneficial to experimentally measure the diffusivity of all the cytosolic proteins included in the model by microscopy techniques such as fluorescence recovery after photobleaching (FRAP), as current predictions rely on calculating the diffusivities for each cytosolic protein monomer or complex by adjusting the diffusivity of tubulin (*31*) based on differences between the hydrodynamic radii of tubulin and a given protein monomer or complex (*32*). In addition to the model verification through refining model parameters, additional cellular processes can be incorporated into the model to more completely recapitulate all cellular processes relevant to cytosolic protein complex assemblies. Some of these processes to consider which can control the spatial organization of multiprotein complexes include receptor internalization (*18–20*) and coordinated protein complex assembly via cytosolic scaffold proteins (*46*).

Beyond providing an estimate of the length scale of GAB1-SHP2 association, our model predictions suggest that membrane-bound EGFR may be able to regulate intracellular signaling events over the entire intracellular length scale through the intermediate action of SFKs and diffusible GAB1-SHP2 complexes (Fig. 8). This paradigm of receptor-mediated signaling stands

30

in contRASt to the more conventional view that receptor-bound complexes, such as GAB1-SHP2, access different subcellular compartments through directed membrane localization and trafficking processes. Such an understanding may help generate new hypotheses to explain apparently contradictory observations. For example, recent work from the Sorkin Lab has shown that RAS activation and ERK phosphorylation paradoxically remain dependent on EGFR activity after EGFR and RAS have become physically separated during endocytosis of the receptor (*34*, *36*, *47*). One possible explanation for these observations is that a small pool of active, plasma membrane-localized EGFR is sufficient to maintain signaling through the RAS-MAPK cascade. Based on our model predictions and experiments in the same HeLa/mV-HRAS cells, it is also possible that GAB1-SHP2 complexes—which may be generated by active, endosomal EGFR via the mechanism described here—may be able to diffuse back to the cell membrane and thereby regulate RAS activity. SHP2 has been suggested to sustain RAS activation via dephosphorylation of GAB1 at Y317, a p120-RASGAP (RASA1) binding site (*48*), or by dephosphorylating RASGAP binding sites directly on EGFR, though other mechanisms of RAS regulation by SHP2 have been proposed (*49–52*), including the possibility of SHP2 acting as a scaffold that recruits GRB2-SOS complexes to the cell membrane (45, 46). Besides regulating RAS activity, SHP2 has also been reported to regulate the phosphorylation of paxillin (*55*, *56*) and the transmembrane adaptor protein PAG/CBP (*50*). Thus, there are several signaling targets of SHP2 that our analysis suggests EGFR can distally regulate via the reaction-diffusion mechanism described here.

More generally, our model predictions suggest that similar mechanisms may exist for other RTKs and cellular receptors like EGFR that can generate diffusible non-receptor kinases and signaling complexes. To this end, the model described in this study offers a platform to study the spatial distribution of cytosolic phospho-proteins or protein complexes initiated by the activity of membrane-associated receptors. In cases where experimental data is unavailable to compare

31

against model predictions, the model can still be used both to identify individual cellular processes that strongly control a protein's spatial distribution and to study how multiple processes interact with or compensate for one another. The capacity for relevant model parameters to be fit to experimental data obtained by FRET microscopy, for example, will enable a more quantitative understanding of how the dynamics of processes such as protein dephosphorylation compete with diffusion to control the cellular length scale over which a receptor tyrosine kinase can orchestrate the assembly of cytosolic protein complexes and thereby influence signaling processes distal from the receptor.

Materials and Methods

Cell culture

HeLa/mVenus-HRAS cells were supplied by A. Sorkin (University of Pittsburgh).and were cultured in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum (FBS), 1% 10,000 µg/mL penicillin-streptomycin, and 1% 200 mM L-glutamine.

Proximity ligation assay and immunofluorescence

HeLa/mV-HRAS cells were plated on 12-mm coverslips in six-well tissue culture plates at 3×10^5 cells per well and incubated for at least 24 hours in Dulbeco's Modified Eagle's Medium (DMEM) supplemented with 10 % FBS prior to serum starving. Cells were serum-starved by switching them to DMEM containing 0.1 % FBS for at least 16 hours prior to treatment with EGF. Serum-starved cells were then treated with recombinant EGF at 10 ng/mL (PeproTech), fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) for 20 minutes, and permeabilized with 0.25% Triton X-100 in PBS for 5 minutes. Nonspecific binding was blocked by incubation with Duolink Blocking Solution for 1 hour at 37 °C. All primary antibodies were incubated overnight at 4 °C. For EGFR-GRB2 experiments, mouse antibody targeting EGFR (CST D381B, Cell Signaling Technology, 1:100 dilution) and rabbit antibody targeting GRB2 (BD 610111, BD Biosciences, 1:100 dilution) diluted in Odyssey Blocking Buffer (OBB) were used. For GAB1-SHP2 experiments, mouse antibody targeting SHP2 (sc-7384, Santa Cruz Biotechnology, 1:50 dilution) and rabbit antibody targeting GAB1 (HPA049599, Sigma-Aldrich, 1:50 dilution) diluted in OBB were used. PLA probes were anti-rabbit PLUS and anti-mouse MINUS and were detected using Duolink In Situ Detection Reagents Far Red (Sigma-Aldrich). After staining with PLA probes, coverslips were incubated overnight at 4 °C with rat primary antibody targeting EGFR diluted in OBB (ab231, Abcam, 1:200 dilution). EGFR was detected by incubating coverslips in

33

OBB with anti-rat Alexa Fluor 594 (A-11007, Invitrogen, 1:750 dilution). Coverslips were mounted on glass slides using Duolink In Situ Mounting Medium with DAPI (Sigma-Aldrich). Epifluorescence images were obtained with a Zeiss AxioObserver Z1 using 20x, 40x oil immersion, and 63x oil immersion objective lenses.

Fluorescence images were analyzed using the FIJI distribution of ImageJ (*57*). Quantification of PLA puncta and the colocalization of puncta with EGFR signals was performed using custom ImageJ macros that were written using a modified combination of the puncta quantification processing steps described by Horzum et al. and the object-based colocalization analysis steps described by Moser et al. (*58*, *59*). The custom ImageJ macros are available on GitHub under the username "pauljmyers".
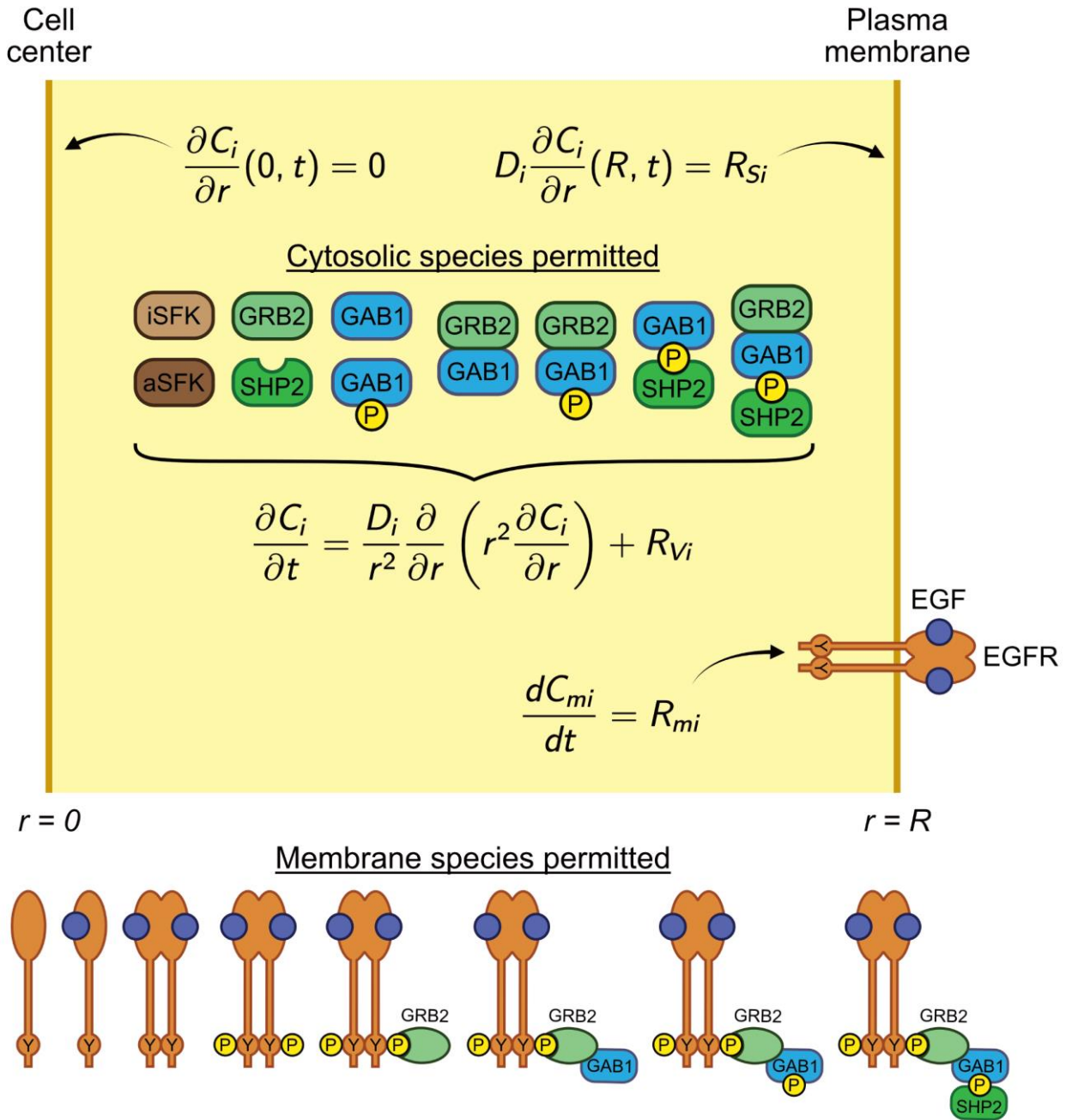
Acknowledgements

Figures

Fig. 1. Model schematic. A representative cell was modeled in spherical coordinates as a sphere of radius R. Cytosolic proteins and complexes were modeled with the standard reaction-diffusion equation in time and space. Symmetry (no-flux) boundary conditions were imposed at the cell center, and reactive Robin boundary conditions were imposed at the cell membrane. Membrane-

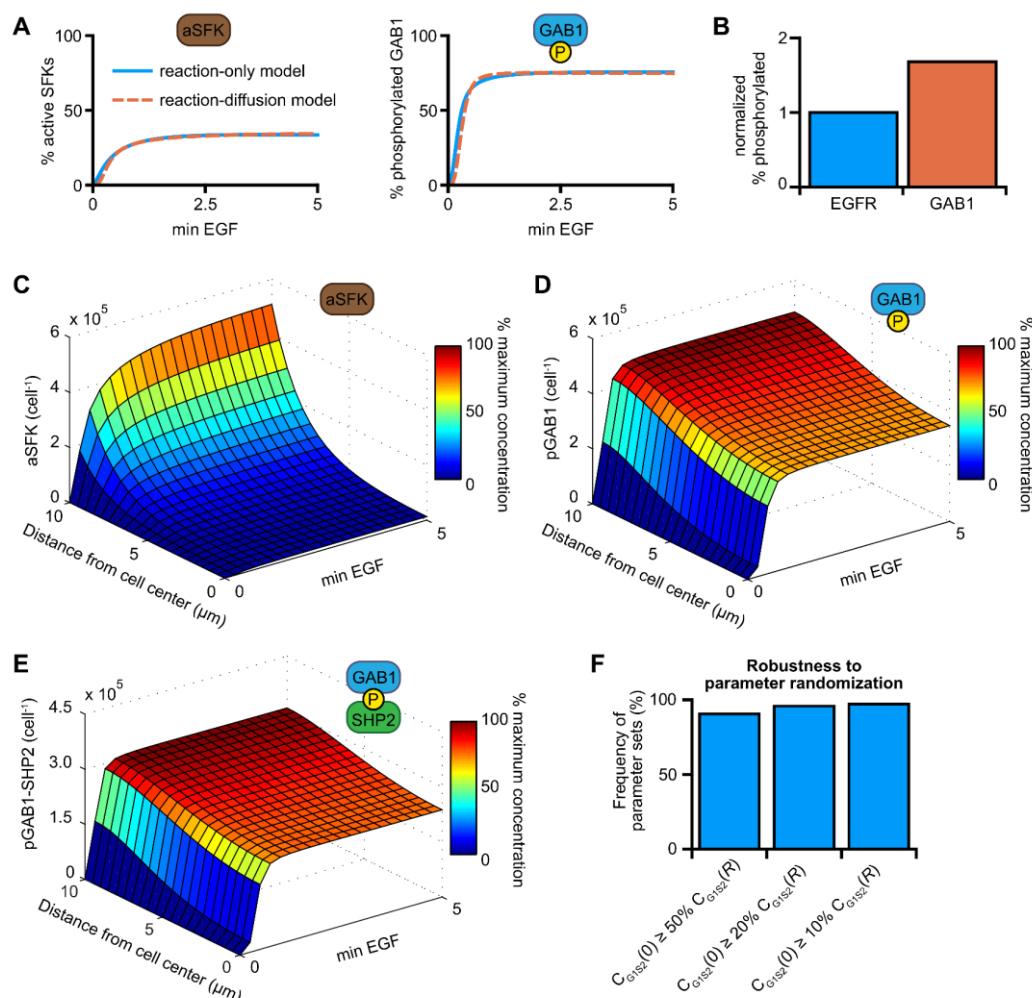bound species were assumed to be well-distributed across the membrane and were only modeled

in time.

Fig. 2. Base model predictions and fits. (**A**) Model predictions for the spatially averaged concentrations of active SFKs and phosphorylated GAB1 were fit to our previous model's predictions, as described in *Materials and Methods*. (**B**) Average concentrations of phosphorylated EGFR and GAB1 throughout 5 min of EGF treatment were calculated and normalized to the average concentration of phosphorylated EGFR. (**C-E**) Model predictions for the concentrations of active SFKs, phosphorylated GAB1, and GAB1-SHP2 complexes were plotted as a function of distance from the cell center and time following EGF treatment. (**F**) Model predictions for whether the concentration of GAB1-SHP2 complexes at the cell center reached 50%, 20%, or 10% of the maximal concentration of GAB1-SHP2 complexes at the cell surface throughout the cell radius, *R*,

39

were calculated for 1,000 random parameter sets, where each parameter was randomly varied by up to a factor of 10 above or below its base value. Protein diffusivities were collectively perturbed by the same factor in each simulation.
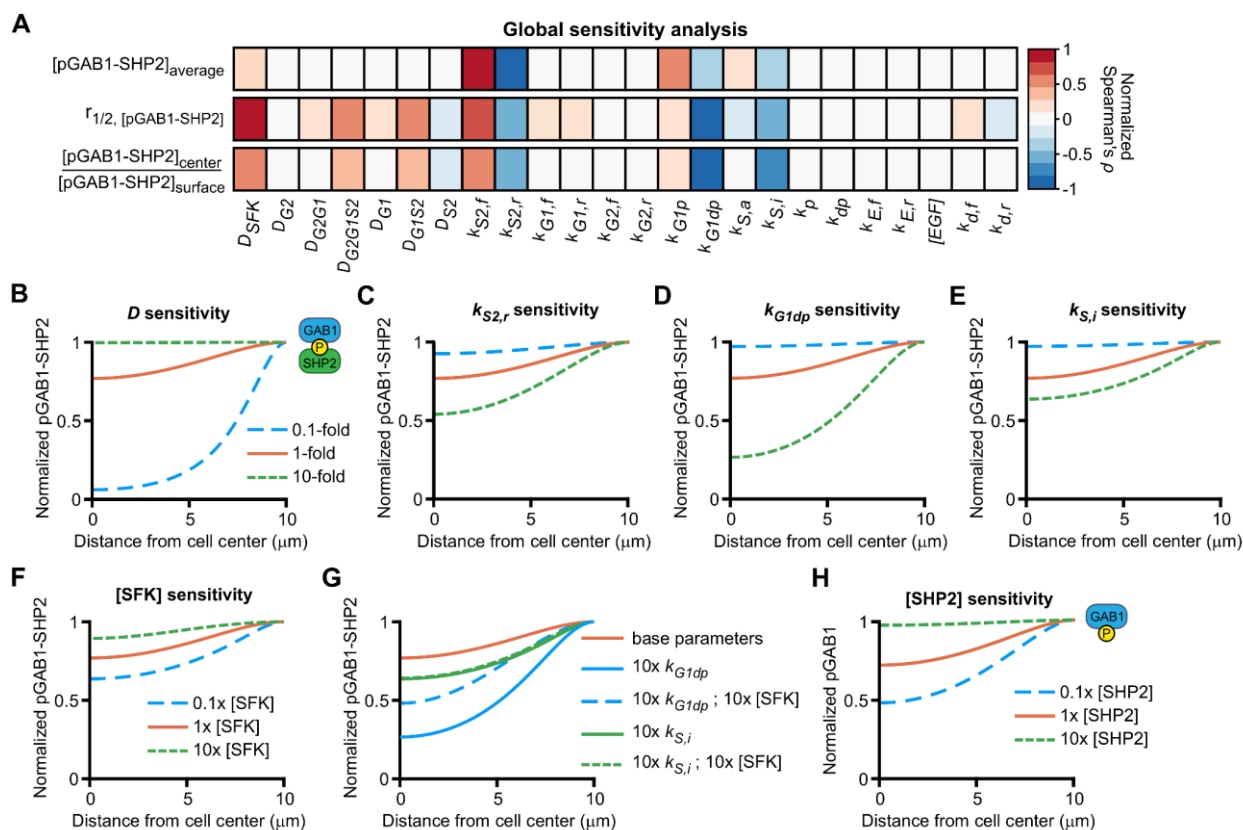
Fig. 3. Model sensitivity analysis. (**A**) Global sensitivity analysis of model predictions for the total concentration of GAB1-SHP2 complexes, distances over which the GAB1-SHP2 concentration decays to 1/2 of the maximum GAB1-SHP2 concentration at the cell surface, and the ratio of GAB1-SHP2 concentrations at the cell center to the cell surface for a 5 min EGF treatment. Latin hypercube sampling (LHS) was used to generate 3,000 combinations of the indicated model parameters when allowing them to vary by up to a factor of 10 above or below their base values, followed by computing model predictions for each LHS-generated set of parameters. Sensitivity was calculated as the Spearman rank between each parameter and each output, normalized to the highest magnitude rank for each output. (**B-E**) Model predictions for the normalized concentration of GAB1-SHP2 after 5 min EGF treatment were compared with predictions when all protein diffusivities (D), $k_{S2,r}$, k_{G1dp}, or $k_{S,i}$ were varied by a factor of 10. (**F-G**) Model predictions for the normalized concentration of GAB1-SHP2 after 5 min EGF treatment

41

were compared with predictions when the indicated parameters or protein concentrations were varied by a factor of 10. (**H**) Model predictions for the normalized concentration of phosphorylated GAB1 after a 5 min EGF treatment were compared with predictions when the concentration of SHP2 was either increased 10-fold or set to zero.
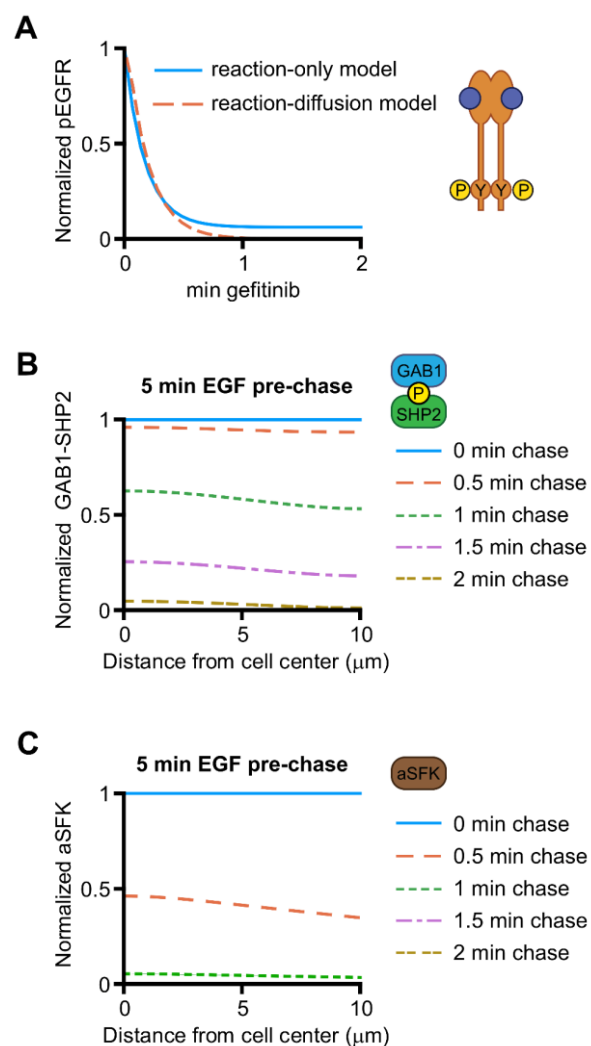
Fig. 4. EGF-pulse EGFR inhibition-chase predictions. (**A**) The normalized concentration of phosphorylated EGFR (spatiotemporal) was predicted for an EGF-pulse EGFR inhibitor-chase simulation, where a 5 min EGF treatment was followed by a 2 min EGFR inhibitor chase where the rate constant for EGFR phosphorylation was set to zero. The concentration of phosphorylated EGFR as a function of time following EGFR inhibitor-chase (spatiotemporal) was compared to predictions from our previous model where the EGFR inhibitor gefitinib was explicitly included in the model topology (temporal). The concentrations of phosphorylated EGFR were normalized by the maximum value within each time course. (**B-C**) The concentrations of GAB1-SHP2 complexes

and active SFKs at the indicated times following EGFR inhibition were normalized by their

concentrations prior to EGFR inhibition and plotted as a function of distance from the cell center.
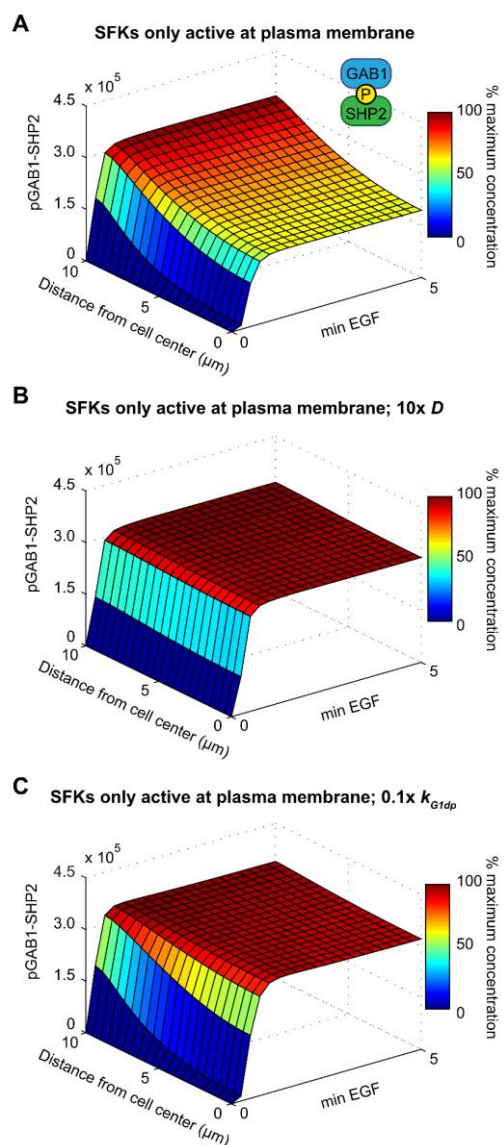
Fig. 5. Model predictions for SFKs only active at the cell surface. (**A-C**) Model predictions for the concentrations GAB1-SHP2 complexes were plotted as a function of distance from the cell center and time following EGF treatment when SFKs were only permitted to be active at the cell surface for either the base model parameters (**A**), a 10-fold increase in D (**B**), or a 10-fold decrease in k_{G1dp} (**C**).
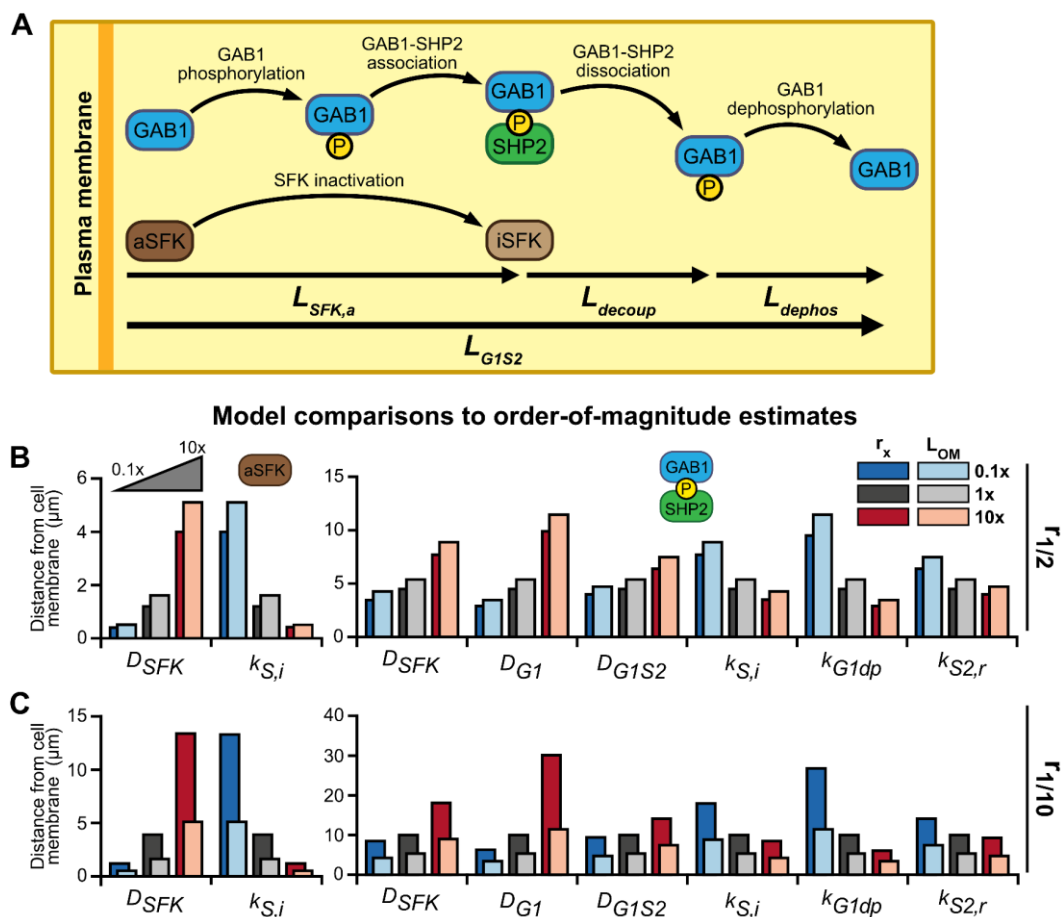
Fig. 6. Comparison of model and order-of-magnitude estimates for the length scales of protein and protein complex persistence. (A) Conceptualization showing the reaction-diffusion processes used to estimate the order-of-magnitude length scale estimate for GAB1-SHP2 association. Over the length scale of SFK activation and inactivation, GAB1 is phosphorylated and subsequently binds to SHP2. Beyond the length scale of SFK activity, the length scale of GAB1-SHP2 association is extended by the length scales over which GAB1-SHP2 dissociation and GAB1 dephosphorylation occur since GAB1 and SHP2 may repeatedly bind and unbind as long as GAB1 remains phosphorylated. **(B)** Model predictions for the distance over which the concentrations of active SFKs and GAB1-SHP2 fall to 1/2 of their maximum concentrations at the cell surface ($r_{1/2}$) compared to order-of-magnitude estimates for the length scales (L_{OM}) over which these species persist from the cell surface when D_{SFK}, D_{G1}, D_{G1S2}, $k_{S,i}$, k_{G1dp}, and $k_{S2,off}$ are varied by a factor of

10 from their base values. **(C)** Model predictions for the distance over which the concentrations of active SFKs and GAB1-SHP2 fall to 1/10 of their maximum concentrations at the cell surface ($r_{1/10}$) compared to order-of-magnitude estimates for the length scales over which these species persist from the cell surface when D_{SFK}, D_{G1}, D_{G1S2}, $k_{S,i}$, k_{G1dp}, and $k_{S2,off}$ are varied by a factor of 10 from their base values. Model predictions for $r_{1/2}$ and $r_{1/10}$ were made here by setting the cell radius to 50 µm since length scales values greater than 10 µm could not be obtained with the base cell radius of 10 µm. In (B) and (C), $r_{1/2}$ and $r_{1/10}$ are denoted by darker bars shifted slightly to the left; L_{OM} values were calculated according to Eqs. 14 and 17 and are denoted by lighter, right-shifted bars.
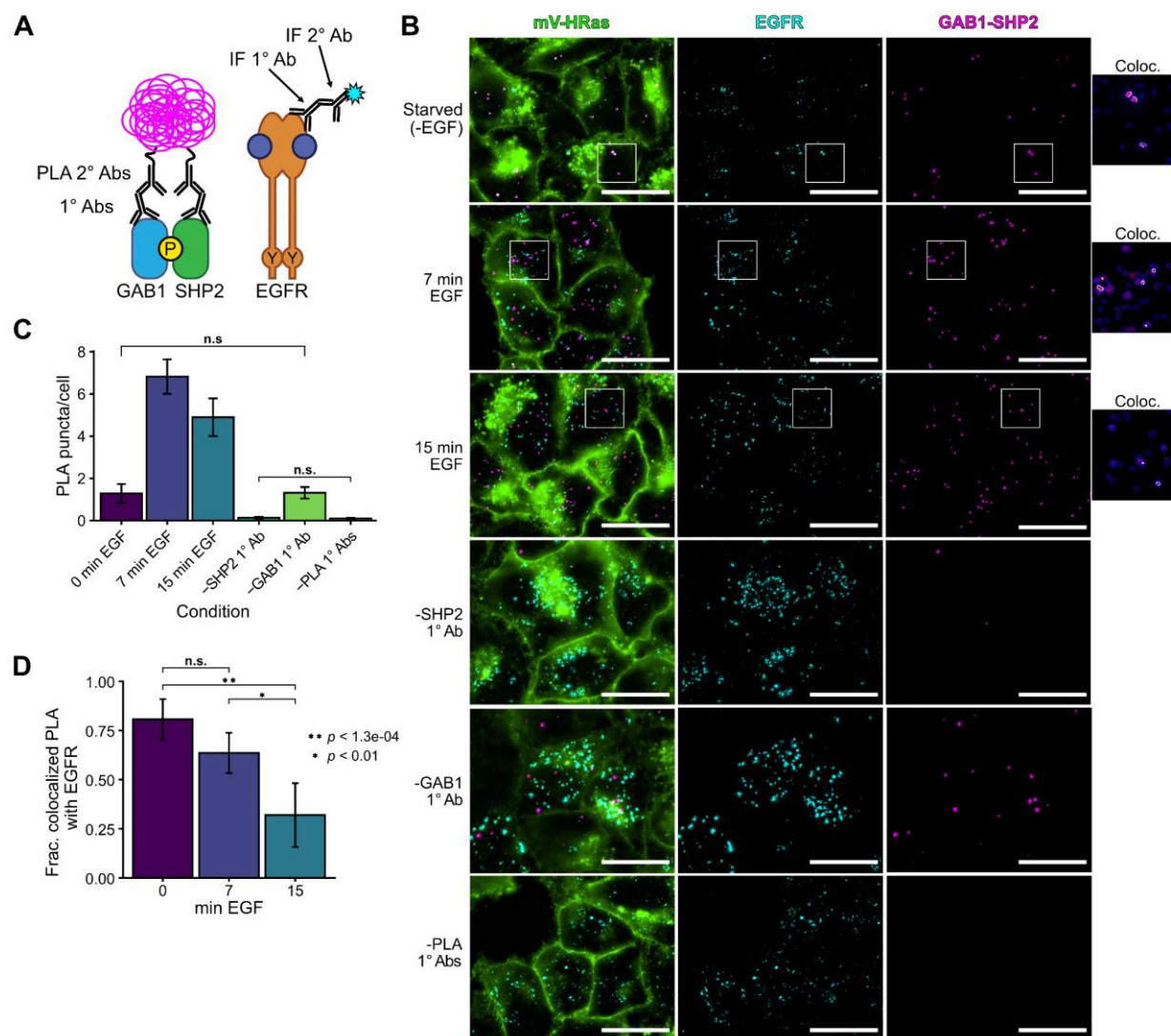
Fig. 7. Proximity ligation assay of GAB1-SHP2 complexes. (A) Schematic representation of the proximity ligation assay (PLA) and experimental setup. EGFR was stained via standard immunofluorescence (IF), and GAB1-SHP2 complexes were stained via PLA to determine their localization within cells relative to EGFR. (B) Maximum-intensity z-projections of HeLa/mVenus-HRAS that were treated with 10 ng/mL EGF as indicated, fixed, and stained for EGFR and GAB1-SHP2 by IF and PLA, respectively. Insets show 2.5x magnification of the region marked by the white square. Arrows point out examples of colocalization between EGFR IF signal and GAB1-SHP2 PLA signal. PLA antibody control samples were treated with EGF for up to 15 minutes. The

insets (far right) show 2.5x magnification of the regions marked by white rectangles, labeled as "Coloc." Blue outlines indicate detected EGFR and GAB1-SHP2 puncta, and white outlines indicate colocalization (overlap) between EGFR and GAB1-SHP2 signals. (C) Quantification of PLA signals per cell for each of the conditions shown in (C). Welch's ANOVA indicated a statistically significant difference among groups (F = 90.876, p < 2.2e-16). The Games-Howell post-hoc test was used to determine between-group significance levels (p < 0.05 unless otherwise indicated). (D) Fraction of GAB1-SHP2 PLA puncta co-localized with EGFR puncta in treated HeLa/mV-HRAS cells stained using the full PLA reaction (Welch's ANOVA: F = 12.153, p < 1.5e-04). Statistical significance was determined as in (C). $\alpha = 0.05$. Error bars, 95% confidence intervals. Scale bars, 20 μm. Images are representative of three independent sets of at least two coverslips per condition.
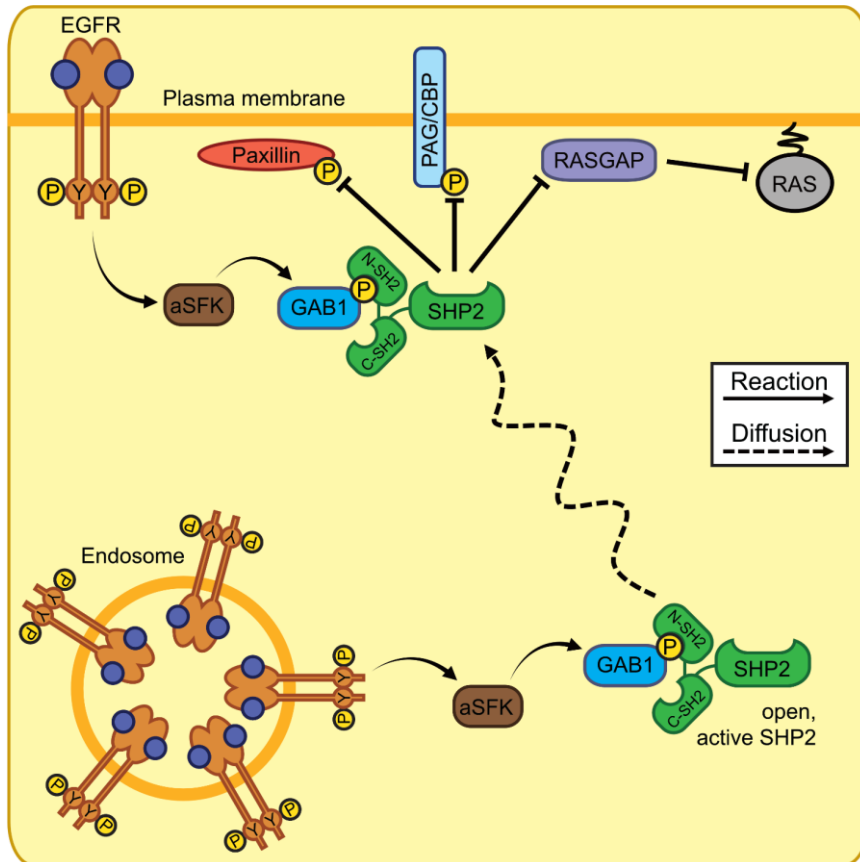
Fig. 8. Diffusible GAB1-SHP2 complexes enable EGFR to regulate signaling processes distal from the receptor. Through the intermediate action of SFKs, EGFR promotes the formation of diffusible GAB1-SHP2 complexes, and hence active SHP2, which can then translocate to different cellular compartments where it acts on its signaling targets. In cell backgrounds where the majority of EGFR is rapidly internalized after receptor activation, EGFR may still be able to influence signaling processes at the plasma membrane and in other cellular compartments via the activation of diffusible signaling complexes like GAB1-SHP2.

Table 1. Model parameters

Parameter (units)	Description	Value	Source
$k_{E,f}$ (μM^{-1} min^{-1})	EGF binding to EGFR, forward	3.1×10^2	(28)
$k_{E,r}$ (min^{-1})	EGF binding to EGFR, reverse	8.0×10^{-1}	(29)
$k_{dE,f}$ (μm^2 molec^{-1} min^{-1})	EGFR dimerization, forward	1.2×10^0	(9)
$k_{dE,r}$ (min^{-1})	EGFR dimerization, reverse	1.0×10^{-1}	(49)
k_{catE} (min^{-1})	EGFR phosphorylation, EGF-occupied dimer	1.3×10^1	(50)
k_{dp} (min^{-1})	EGFR dephosphorylation	8.0×10^0	(9)
$k_{S,a}$ (μm^3 molec^{-1} min^{-1})	SFK activation	7.7×10^{-1}	Fit
$k_{S,i}$ (min^{-1})	SFK inactivation	3.2×10^1	Fit
$k_{G2,f}$ (μm^3 molec^{-1} min^{-1})	GRB2 binding to EGFR, forward	1.6×10^0	(2)
$k_{G2,r}$ (min^{-1})	GRB2 binding to EGFR, reverse	4.6×10^2	(2)
$k_{G1,f}$ (μm^3 molec^{-1} min^{-1})	GAB1 binding to GRB2, forward	1.0×10^{-1}	(51)
$k_{G1,r}$ (min^{-1})	GAB1 binding to GRB2, reverse	6.0×10^1	(51)
k_{G1p} (μm^3 molec^{-1} min^{-1})	GAB1 phosphorylation	2.7×10^{-1}	Fit
k_{G1dp} (min^{-1})	GAB1 dephosphorylation	8.4×10^0	Fit
$k_{S2,f}$ (μm^3 molec^{-1} min^{-1})	SHP2 binding to phosphorylated GAB1, forward	1.6×10^0	(52)
$k_{S2,r}$ (min^{-1})	SHP2 binding to phosphorylated GAB1, reverse	4.6×10^2	(52)
$C_{o,EGF}$ (μM)	Extracellular EGF concentration	1.7×10^{-3}	See text
$C_{o,EGFR}$ (cell^{-1})	EGFR molecules per cell	6.0×10^5	(9)
$C_{o,GRB2}$ (cell^{-1})	GRB2 molecules per cell	6.0×10^5	(51)
$C_{o,GAB1}$ (cell^{-1})	GAB1 molecules per cell	6.0×10^5	(51)
$C_{o,SHP2}$ (cell^{-1})	SHP2 molecules per cell	6.0×10^5	(51)
$C_{o,SFK}$ (cell^{-1})	SFK molecules per cell	6.0×10^5	(51)
D_S (μm^2 min^{-1})	Diffusivity of SFK molecules	8.2×10^1	(31, 32)
D_{G2} (μm^2 min^{-1})	Diffusivity of GRB2 molecules	1.3×10^2	(31, 32)
D_{G2G1} (μm^2 min^{-1})	Diffusivity of GRB2-GAB1 molecules	6.1×10^1	(31, 32)
D_{G2G1S2} (μm^2 min^{-1})	Diffusivity of GRB2-GAB1-SHP2 molecules	5.5×10^1	(31, 32)
D_{G1} (μm^2 min^{-1})	Diffusivity of GAB1 molecules	6.6×10^1	(31, 32)
D_{G1S2} (μm^2 min^{-1})	Diffusivity of GAB1-SHP2 molecules	5.6×10^1	(31, 32)
D_{S2} (μm^2 min^{-1})	Diffusivity of SHP2 molecules	7.8×10^1	(31, 32)
R (μm)	Cell radius	1.0×10^1	(27)

References

1. B. J. Mayer, Perspective: Dynamics of receptor tyrosine kinase signaling complexes. *FEBS Lett*. **586**, 2575–2579 (2012).

2. Morimatsu, Takagi, Ota, Iwamoto, Yanagida, Sako, Multiple-state reactions between the epidermal growth factor receptor and Grb2 as observed by using single-molecule analysis. *Proceedings of the National Academy of Sciences of the United States of America*. **104**, 18013–18018 (2007).

3. M.-M. Zhou, J. E. Harlan, W. S. Wade, S. Crosby, K. S. Ravichandran, S. J. Burakoff, S. W. Fesik, Binding Affinities of Tyrosine-phosphorylated Peptides to the COOH-terminal SH2 and NH2-terminal Phosphotyrosine Binding Domains of Shc (∗). *Journal of Biological Chemistry*. **270**, 31119–31123 (1995).

4. M. Offterdinger, V. Georget, A. Girod, P. I. H. Bastiaens, Imaging Phosphorylation Dynamics of the Epidermal Growth Factor Receptor*. *Journal of Biological Chemistry*. **279**, 36972–36981 (2004).

5. A. Sorkin, M. McClure, F. Huang, R. Carter, Interaction of EGF receptor and grb2 in living cells visualized by fluorescence resonance energy transfer (FRET) microscopy. *Curr Biol*. **10**, 1395–1398 (2000).

6. B. G. Neel, H. Gu, L. Pao, The 'Shp'ing news: SH2 domain-containing tyrosine phosphatases in cell signaling. *Trends Biochem Sci*. **28**, 284–293 (2003).

7. H. Daub, C. Wallasch, A. Lankenau, A. Herrlich, A. Ullrich, Signal characteristics of G protein-transactivated EGF receptor. *EMBO J*. **16**, 7032–7044 (1997).

8. H. Gu, B. G. Neel, The "Gab" in signal transduction. *Trends Cell Biol*. **13**, 122–130 (2003).

9. C. M. Furcht, J. M. Buonato, M. J. Lazzara, EGFR-activated Src family kinases maintain GAB1-SHP2 complexes distal from EGFR. *Science signaling*. **8**, ra46 (2015).

10. A. Montagner, A. Yart, M. Dance, B. Perret, J. P. Salles, P. Raynal, A novel role for Gab1 and SHP2 in epidermal growth factor-induced Ras activation. *Journal of Biological Chemistry*. **280**, 5350–5360 (2005).

11. G. A. Rodrigues, M. Falasca, Z. Zhang, S. H. Ong, J. Schlessinger, A Novel Positive Feedback Loop Mediated by the Docking Protein Gab1 and Phosphatidylinositol 3-Kinase in Epidermal Growth Factor Receptor Signaling. *Molecular and Cellular Biology*. **20**, 1448 (2000).

12. M. Incoronato, A. D'Alessio, S. Paladino, C. Zurzolo, M. S. Carlomagno, L. Cerchia, V. de Franciscis, The Shp-1 and Shp-2, tyrosine phosphatases, are recruited on cell membrane in two distinct molecular complexes including Ret oncogenes. *Cell Signal*. **16**, 847–856 (2004).

13. X. Zhou, Y. M. Agazie, Molecular mechanism for SHP2 in promoting HER2-induced signaling and transformation. *J Biol Chem*. **284**, 12226–12234 (2009).

14. U. Schaeper, N. H. Gehring, K. P. Fuchs, M. Sachs, B. Kempkes, W. Birchmeier, Coupling of Gab1 to c-Met, Grb2, and Shp2 mediates biological responses. *J Cell Biol*. **149**, 1419–1432 (2000).

15. C. R. Maroun, M. A. Naujokas, M. Park, Membrane targeting of Grb2-associated binder-1 (Gab1) scaffolding protein through Src myristoylation sequence substitutes for Gab1 pleckstrin homology domain and switches an epidermal growth factor response to an invasive morphogenic program. *Mol Biol Cell*. **14**, 1691–1708 (2003).

16. C. L. Abram, S. A. Courtneidge, Src family tyrosine kinases and growth factor signaling. *Exp Cell Res*. **254**, 1–13 (2000).

17. M. Okada, H. Nakagawa, A protein tyrosine kinase involved in regulation of pp60c-src function. *J Biol Chem*. **264**, 20886–20893 (1989).

18. Y. Wang, S. Pennock, X. Chen, Z. Wang, Endosomal signaling of epidermal growth factor receptor stimulates signal transduction pathways leading to cell survival. *Mol Cell Biol*. **22**, 7279–7290 (2002).

19. M. Miaczynska, L. Pelkmans, M. Zerial, Not just a sink: endosomes in control of signal transduction. *Curr Opin Cell Biol*. **16**, 400–406 (2004).

20. H. Ye, R. Kuruvilla, L. S. Zweifel, D. D. Ginty, Evidence in support of signaling endosome-based retrograde survival of sympathetic neurons. *Neuron*. **39**, 57–68 (2003).

21. A. Fortian, A. Sorkin, Live-cell fluorescence imaging reveals high stoichiometry of Grb2 binding to the EGF receptor sustained during endocytosis. *J Cell Sci*. **127**, 432–444 (2014).

22. C. M. Furcht, J. M. Buonato, N. Skuli, L. K. Mathew, A. R. Munoz Rojas, M. C. Simon, M. J. Lazzara, Multivariate signaling regulation by SHP2 differentially controls proliferation and therapeutic response in glioma cells. *J Cell Sci*. **127**, 3555–3567 (2014).

23. C. M. Furcht, A. R. Munoz Rojas, D. Nihalani, M. J. Lazzara, Diminished functional role and altered localization of SHP2 in non-small cell lung cancer cells with EGFR-activating mutations. *Oncogene*. **32**, 2346–55, 2355 e1–10 (2013).

24. L. K. Bieniasz, The von Neumann stability of finite-difference algorithms for the electrochemical kinetic simulation of diffusion coupled with homogeneous reactions. *Journal of Electroanalytical Chemistry*. **345**, 13–25 (1993).

25. D. C. Resasco, F. Gao, F. Morgan, I. L. Novak, J. C. Schaff, B. M. Slepchenko, *Virtual Cell: Computational tools for modeling in cell biology* (NIH Public Access, 2012; http://www.ncbi.nlm.nih.gov/pubmed/22139996), vol. 4.

26. M. L. Blinov, J. C. Schaff, D. Vasilescu, I. I. Moraru, J. E. Bloom, L. M. Loew, Compartmental and Spatial Rule-Based Modeling with Virtual Cell. *Biophysical Journal*. **113**, 1365–1372 (2017).

27. B. N. Kholodenko, MAP kinase cascade signaling and endocytic trafficking: A marriage of convenience? *Trends in Cell Biology*. **12** (2002), pp. 173–177.

28. C. M. Waters, K. C. Oberg, G. Carpenter, K. A. Overholser, Rate Constants for Binding, Dissociation, and Internalization of EGF: Effect of Receptor Occupancy and Ligand Concentration. *Biochemistry*. **29**, 3563–3569 (1990).

29. A. R. French, D. K. Tadaki, S. K. Niyogi, D. A. Lauffenburger, Intracellular trafficking of epidermal growth factor family ligands is directly influenced by the pH sensitivity of the receptor/ligand interaction. *Journal of Biological Chemistry*. **270**, 4334–4340 (1995).

30. Y. Shan, M. P. Eastwood, X. Zhang, E. T. Kim, A. Arkhipov, R. O. Dror, J. Jumper, J. Kuriyan, D. E. Shaw, Oncogenic mutations counteract intrinsic disorder in the EGFR kinase and promote receptor dimerization. *Cell*. **149**, 860–870 (2012).

31. R. Pepperkok, M. H. Bre, J. Davoust, T. E. Kreis, Microtubules are stabilized in confluent epithelial cells but not in fibroblasts. *Journal of Cell Biology* (1990), doi:10.1083/jcb.111.6.3003.

32. H. P. Erickson, Size and Shape of Protein Molecules at the Nanometer Level Determined by Sedimentation, Gel Filtration, and Electron Microscopy. *Biological Procedures Online*. **11**, 32–51 (2009).

33.   S. Marino, I. B. Hogue, C. J. Ray, D. E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*. **254**, 178–196 (2008).

34.   S. Surve, S. C. Watkins, A. Sorkin, EGFR-RAS-MAPK signaling is confined to the plasma membrane and associated endorecycling protrusions. *Journal of Cell Biology*. **220** (2021), doi:10.1083/JCB.202107103.

35.   R. J. Reddy, A. S. Gajadhar, E. J. Swenson, D. A. Rothenberg, T. G. Curran, F. M. White, Early signaling dynamics of the epidermal growth factor receptor. *Proceedings of the National Academy of Sciences*. **113**, 3114–3119 (2016).

36.   I. Pinilla-Macua, S. C. Watkins, A. Sorkin, Endocytosis separates EGF receptors from endogenous fluorescently labeled HRas and diminishes receptor signaling to MAP kinases in endosomes. *Proceedings of the National Academy of Sciences*. **113**, 2122–2127 (2016).

37.   K. Roepstorff, M. V. Grandal, L. Henriksen, S. L. J. Knudsen, M. Lerdrup, L. Grøvdal, B. M. Willumsen, B. van Deurs, Differential Effects of EGFR Ligands on Endocytic Sorting of the Receptor. *Traffic*. **10**, 1115–1127 (2009).

38.   F. Huang, D. Kirkpatrick, X. Jiang, S. Gygi, A. Sorkin, Differential Regulation of EGF Receptor Internalization and Degradation by Multiubiquitination within the Kinase Domain. *Molecular Cell*. **21**, 737–748 (2006).

39.   D. Rotin, B. Margolis, M. Mohammadi, R. J. Daly, G. Daum, N. Li, E. H. Fischer, W. H. Burgess, A. Ullrich, J. Schlessinger, SH2 domains prevent tyrosine dephosphorylation of the EGF receptor: identification of Tyr992 as the high-affinity binding site for SH2 domains of phospholipase C gamma. *EMBO J*. **11**, 559–567 (1992).

40.   K. Dittmann, C. Mayer, R. Kehlbach, H. P. Rodemann, Radiation-induced caveolin-1 associated EGFR internalization is linked with nuclear EGFR transport and activation of DNA-PK. *Mol Cancer*. **7**, 69 (2008).

41.   E. Sandilands, V. G. Brunton, M. C. Frame, The membrane targeting and spatial activation of Src, Yes and Fyn is influenced by palmitoylation and distinct RhoB/RhoD endosome requirements. *J Cell Sci*. **120**, 2555–2564 (2007).

42.   N. A. Kulak, G. Pichler, I. Paron, N. Nagaraj, M. Mann, Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nature Methods*. **11**, 319–324 (2014).

43.   R. Tsutsumi, J. Harizanova, R. Stockert, K. Schröder, P. I. H. Bastiaens, B. G. Neel, Assay to visualize specific protein oxidation reveals spatio-temporal regulation of SHP2. *Nature Communications*. **8**, 1–14 (2017).

44.   M. J. Lazzara, K. Lane, R. Chan, P. J. Jasper, M. B. Yaffe, P. K. Sorger, T. Jacks, B. G. Neel, D. A. Lauffenburger, Impaired SHP2-mediated extracellular signal-regulated kinase activation contributes to gefitinib sensitivity of lung cancer cells with epidermal growth factor receptor-activating mutations. *Cancer Res*. **70**, 3843–3850 (2010).

45.   M. Donepudi, M. D. Resh, c-Src trafficking and co-localization with the EGF receptor promotes EGF ligand-independent EGF receptor activation and signaling. *Cell Signal*. **20**, 1359–1367 (2008).

46.   M. C. Good, J. G. Zalatan, W. A. Lim, Scaffold proteins: hubs for controlling the flow of cellular information. *Science*. **332**, 680–686 (2011).

47.   S. v. Surve, P. J. Myers, S. A. Clayton, S. C. Watkins, M. J. Lazzara, A. Sorkina, Localization dynamics of endogenous fluorescently labeled RAF1 in EGF-stimulated cells. *Molecular Biology of the Cell*. **30**, 506–523 (2019).

48.   A. Montagner, A. Yart, M. Dance, B. Perret, J. P. Salles, P. Raynal, A novel role for Gab1 and SHP2 in epidermal growth factor-induced Ras activation. *J Biol Chem*. **280**, 5350–5360 (2005).

49.   T. A. Ahmed, C. Adamopoulos, Z. Karoulia, X. Wu, R. Sachidanandam, S. A. Aaronson, P. I. Poulikakos, SHP2 Drives Adaptive Resistance to ERK Signaling Inhibition in Molecularly Defined Subsets of ERK-Dependent Tumors. *Cell Reports*. **26**, 65-78.e5 (2019).

50.   S. Q. Zhang, W. Yang, M. I. Kontaridis, T. G. Bivona, G. Wen, T. Araki, J. Luo, J. A. Thompson, B. L. Schraven, M. R. Philips, B. G. Neel, Shp2 Regulates Src Family Kinase Activity and Ras/Erk Activation by Controlling Csk Recruitment. *Molecular Cell*. **13**, 341–355 (2004).

51.   R. J. Nichols, F. Haderk, C. Stahlhut, C. J. Schulze, G. Hemmati, D. Wildes, C. Tzitzilonis, K. Mordec, A. Marquez, J. Romero, T. Hsieh, A. Zaman, V. Olivas, C. McCoach, C. M. Blakely, Z. Wang, G. Kiss, E. S. Koltun, A. L. Gill, M. Singh, M. A. Goldsmith, J. A. M. Smith, T. G. Bivona, RAS nucleotide cycling underlies the SHP2 phosphatase dependence of mutant BRAF-, NF1- and RAS-driven cancers. *Nature Cell Biology*. **20**, 1064–1073 (2018).

52.   T. S. Batth, M. Papetti, A. Pfeiffer, M. A. X. Tollenaere, C. Francavilla, J. V. Olsen, Large-Scale Phosphoproteomics Reveals Shp-2 Phosphatase-Dependent Regulators of Pdgf Receptor Signaling. *Cell Reports*. **22**, 2784–2796 (2018).

53.   M. Dance, A. Montagner, J.-P. Salles, A. Yart, P. Raynal, The molecular functions of Shp2 in the Ras/Mitogen-activated protein kinase (ERK1/2) pathway. *Cellular Signalling*. **20**, 453–459 (2008).

54.   K. S. Grossmann, M. Rosário, C. Birchmeier, W. Birchmeier, The tyrosine phosphatase Shp2 in development and cancer. *Advances in cancer research*. **106** (2010), pp. 53–89.

55.   Y. Ren, Z. Chen, L. Chen, B. Fang, H. Win-Piazza, E. Haura, J. M. Koomen, J. Wu, Critical Role of Shp2 in Tumor Growth Involving Regulation of c-Myc. *Genes & Cancer*. **1**, 994 (2010).

56.   Y. Ren, S. Meng, L. Mei, Z. J. Zhao, R. Jove, J. Wu, Roles of Gab1 and SHP2 in Paxillin Tyrosine Dephosphorylation and Src Activation in Response to Epidermal Growth Factor*. *Journal of Biological Chemistry*. **279**, 8497–8505 (2004).

57.   J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nature Methods*. **9** (2012), pp. 676–682.

58.   U. Horzum, B. Ozdil, D. Pesen-Okvur, Step-by-step quantitative analysis of focal adhesions. *MethodsX*. **1**, 56–9 (2014).

59.   B. Moser, B. Hochreiter, R. Herbst, J. A. Schmid, Fluorescence colocalization microscopy analysis can be improved by combining object-recognition with pixel-intensity-correlation. *Biotechnology Journal*. **12**, 1600332 (2017).

60.   B. S. Hendriks, L. K. Opresko, H. S. Wiley, D. Lauffenburger, Quantitative analysis of HER2-mediated effects on HER2 and epidermal growth factor receptor endocytosis: distribution of homo- and heterodimers depends on relative HER2 levels. *J Biol Chem*. **278**, 23343–23351 (2003).

61.   Y. X. Fan, L. Wong, T. B. Deb, G. R. Johnson, Ligand regulates epidermal growth factor receptor kinase specificity: activation increases preference for GAB1 and SHC versus autophosphorylation sites. *J Biol Chem*. **279**, 38143–38150 (2004).

62.   A. Kiyatkin, E. Aksamitiene, N. I. Markevich, N. M. Borisov, J. B. Hoek, B. N. Kholodenko, Scaffolding protein Grb2-associated binder 1 sustains epidermal growth factor-induced mitogenic and survival signaling by multiple positive feedback loops. *Journal of Biological Chemistry*. **281**, 19925–19938 (2006).

63.   D. Barua, J. R. Faeder, J. M. Haugh, Structure-based kinetic models of modular signaling protein function: focus on Shp2. *Biophys J*. **92**, 2290–2300 (2007).