# Supplemental Materials

***Molecular Biology of the Cell***

Valley et al.

**SUPPORTING INFORMATION**

**Table S1.** Diffusion coefficients and dimer off-rates for HA-EGFR-WT, HA-EGFR-L858R, and HA-EGFR-ΔL747-P753insS.

| | | D ($\mu m^2 s^{-1}$) | s.e.m. ($\pm\mu m^2 s^{-1}$) | $K_{off,dimer}$ ($s^{-1}$) | $K_{off,dimer}$ Error ($s^{-1}$) | # of dimers |
|---|---|---|---|---|---|---|
| EGFR WT | Unliganded (0-EGF:2-EGFR) | 0.0442 | 0.003 | 0.31 | 0.026 | 262 |
| | EGF-bound (2-EGF:2-EGFR) | 0.0138 | 0.002 | 0.12 | 0.016 | 175 |
| | Singly liganded (1-EGF:2EGFR) | - | - | 0.22 | 0.036 | 67 |
| EGFR L858R | Unliganded (0-EGF:2-EGFR) | 0.0215 | 0.002 | 0.16 | 0.028 | 260 |
| | EGF-bound (2-EGF:2-EGFR) | 0.015 | 0.005 | 0.16 | 0.031 | 97 |
| | Singly liganded (1-EGF:2EGFR) | - | - | 0.14 | 0.032 | 35 |
| EGFR ΔL747-P753insS | Unliganded (0-EGF:2-EGFR) | 0.0226 | 0.002 | 0.16 | 0.036 | 260 |
| | EGF-bound (2-EGF:2-EGFR) | 0.0125 | 0.001 | 0.11 | 0.026 | 61 |
| | Singly liganded (1-EGF:2EGFR) | - | - | 0.16 | 0.029 | 63 |

Diffusion coefficients (D) represent the mobile component from two-component fitting of the distribution of squared displacements. Shown are the mean D and s.e.m. for at least three independent experiments.

Dimer off-rates ($K_{off,dimer}$) and error estimates ($K_{off,dimer}$ Error) are calculated using Hidden Markov Model analysis (Low-Nam *et al.*, 2011).

Diffusion results are shown in Figure 2 F-H, and dimerization results are shown in Figure 2 A-E.

**Table S2.** Fluorescence lifetimes for ACP-EGFR-WT and ACP-EGFR-L858R.

| | Donor | | | Donor + Acceptor | | | Donor + Acceptor + EGF | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\tau_2$ (ns) | $\tau_2$ error | n | $\tau_2$ (ns) | $\tau_2$ error | n | $\tau_2$ (ns) | $\tau_2$ error | n |
| EGFR-WT | 3.93 | 0.04 | 223 | 3.16 | 0.06 | 177 | 3.41 | 0.05 | 75 |
| EGFR-L858R | 3.83 | 0.03 | 99 | 3.34 | 0.03 | 135 | 3.55 | 0.03 | 47 |

Shown are the donor (Oregon Green 488) fluorescence lifetime values, $\tau_2$, for EGFR-WT and EGFR-L858R alone (Donor), with NR12 acceptor (Donor + Acceptor), or in the presence of 30 nM EGF (Donor + Acceptor + EGF).

The fluorescence lifetime error ($\tau_2$ error) represents the s.e.m. for the total number of cells (n) acquired for several independent experiments over multiple days.

See Figure 4.

**Figure S1**

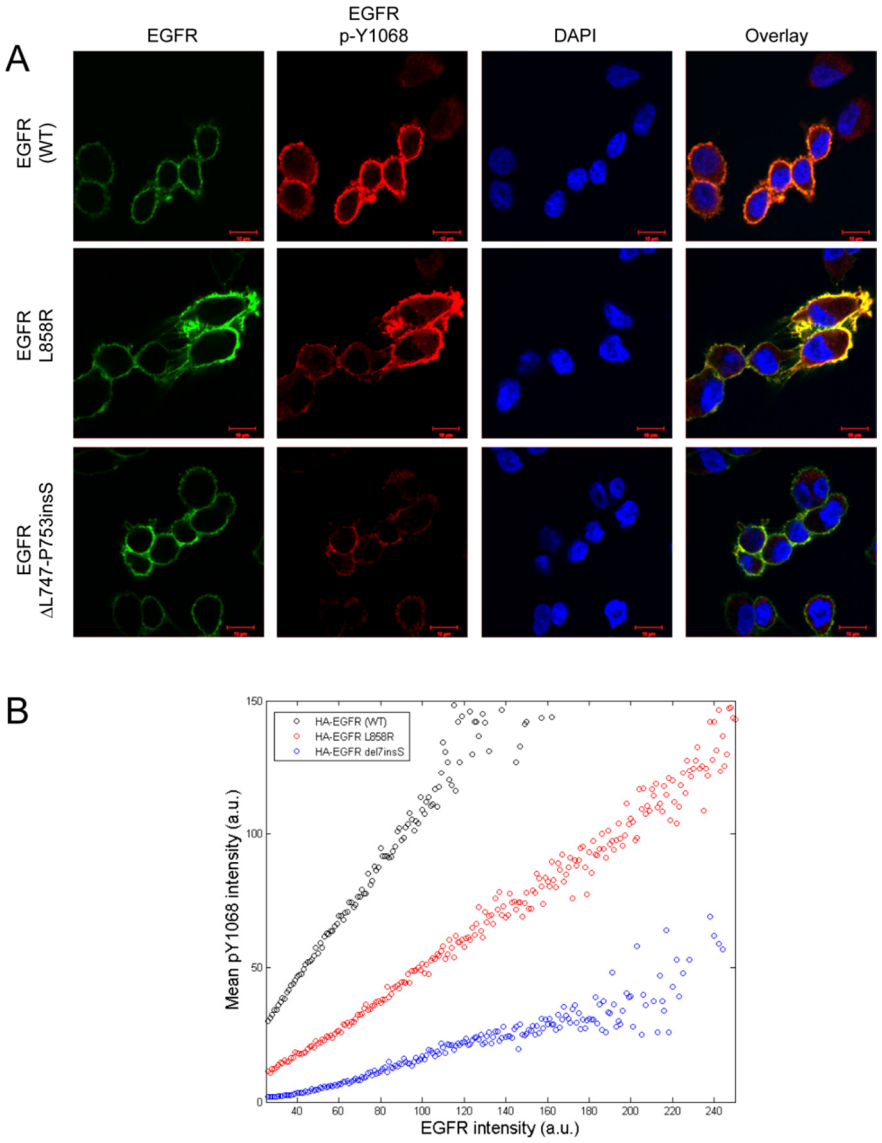



**Figure S1.** CHO cells expressing HA-EGFR-WT, HA-EGFR-L858R, and HA-EGFR-ΔL747-P753insS were treated with dark EGF and labeled for EGFR, p-Y1068 EGFR, and DAPI. **(A)** Fluorescence images show EGF induces phosphorylation of EGFR in the plasma membrane. Note that EGF induces phosphorylation of EGFR in cells expressing low levels of receptor. **(B)** Quantification of EGFR expression and phosphorylation as in Figure 1C. In the presence of saturating concentration of EGF, receptor phosphorylation is independent of expression level, shown as a linear increase in phosphorylation v. receptor expression. This is consistent with previous reports of wild type EGFR (Endres *et al.*, 2013). Note that EGF-induced phosphorylation is lower in EGFR mutants.

**Figure S2**

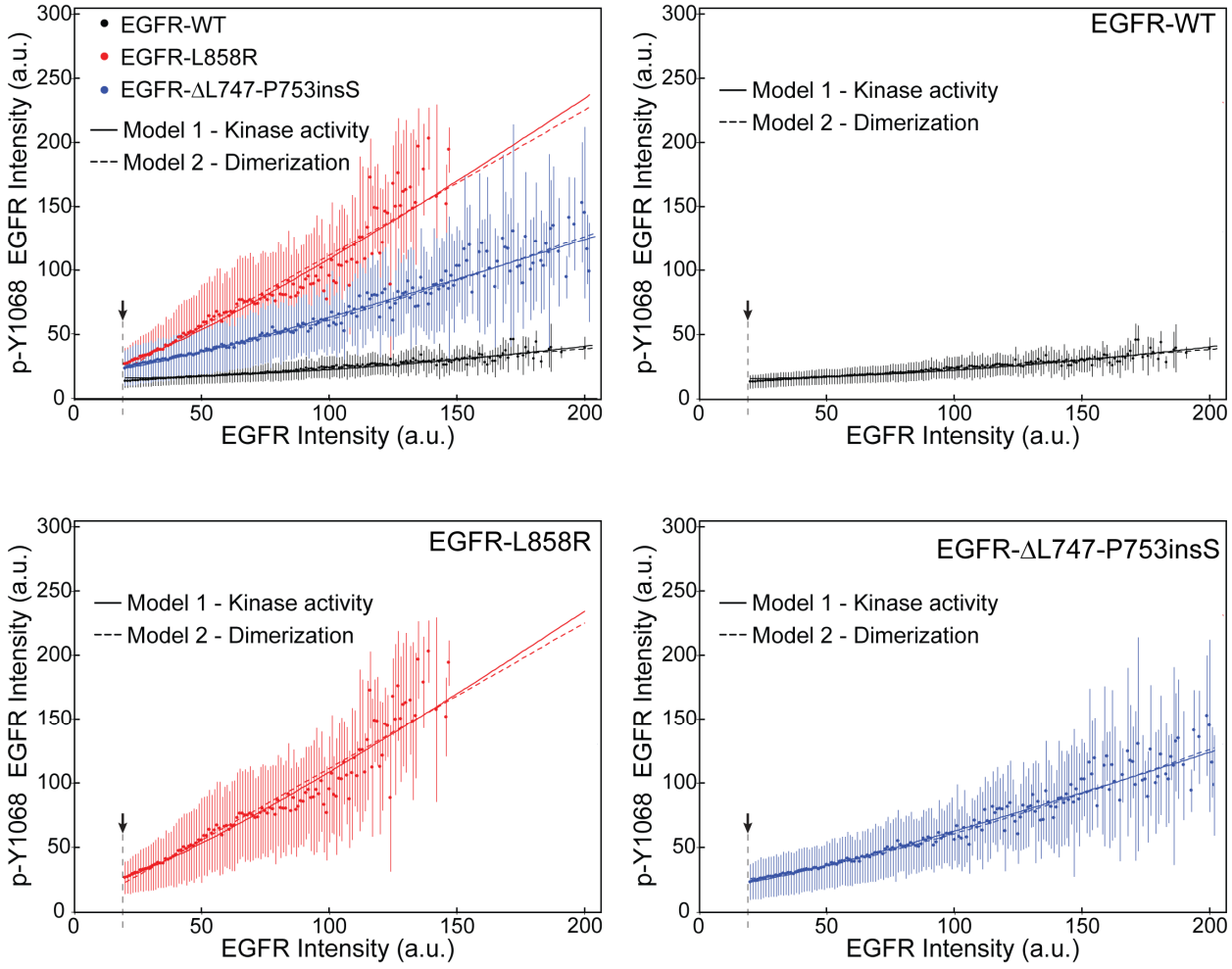



**Figure S2.** Correlation of EGFR phosphorylation (p-Y1068) and total EGFR expression in CHO cells expressing HA-tagged EGFR-WT, -L858R, or - ΔL747-P753insS. Shown are the raw data (mean ± standard deviation) and the equilibrium model fits for changes in kinase activity (solid line, Model 1) or changes in dimerization kinetics (dashed line, model 2). See Figure 1C.

**Figure S3**

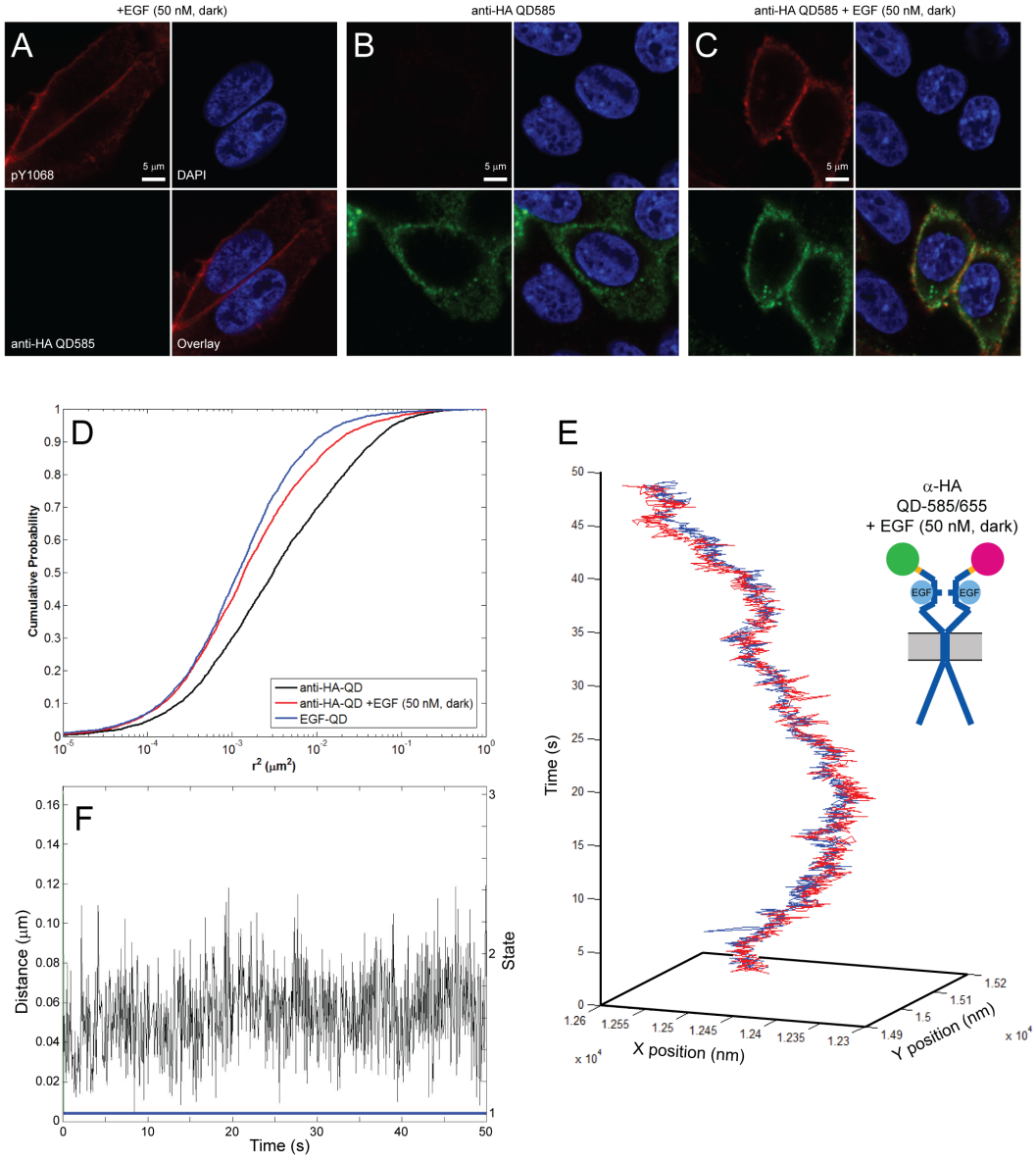



**Figure S3.** QD labeling of HA-EGFR via the N-terminal HA-tag does not prevent receptor dimerization or activation. **(A-C)** CHO cells expressing HA-EGFR-WT were pre-labeled with saturating anti-HA-QD585 and treated with saturating dark EGF (50 nM) as indicated. Cells were fixed and labeled for phospho-EGFR (pY1068) and DAPI. Note that saturated labeling with anti-HA QD does not activate EGFR in the absence of ligand **(B)** nor does it prevent ligand-induced phosphorylation **(C)**. **(D-F)** CHO cells were labeled with anti-HA QD585/QD655, and two-color single particle tracking was used to calculate mobility and visualize receptor dimerization. **(D)** The reduction in mobility upon activation of EGFR tracked with anti-HA-QD in the presence of saturating dark EGF is similar to that observed tracking with EGF-QD. **(E)** EGFR tracked with anti-HA-QD forms stable dimers in the presence of saturating EGF. These events are rare relative to tracking with EGF-QD, due to the high number of dark ligand-occupied receptors, but are classified as dimers using hidden Markov Model analysis **(F)**.

**Figure S4**

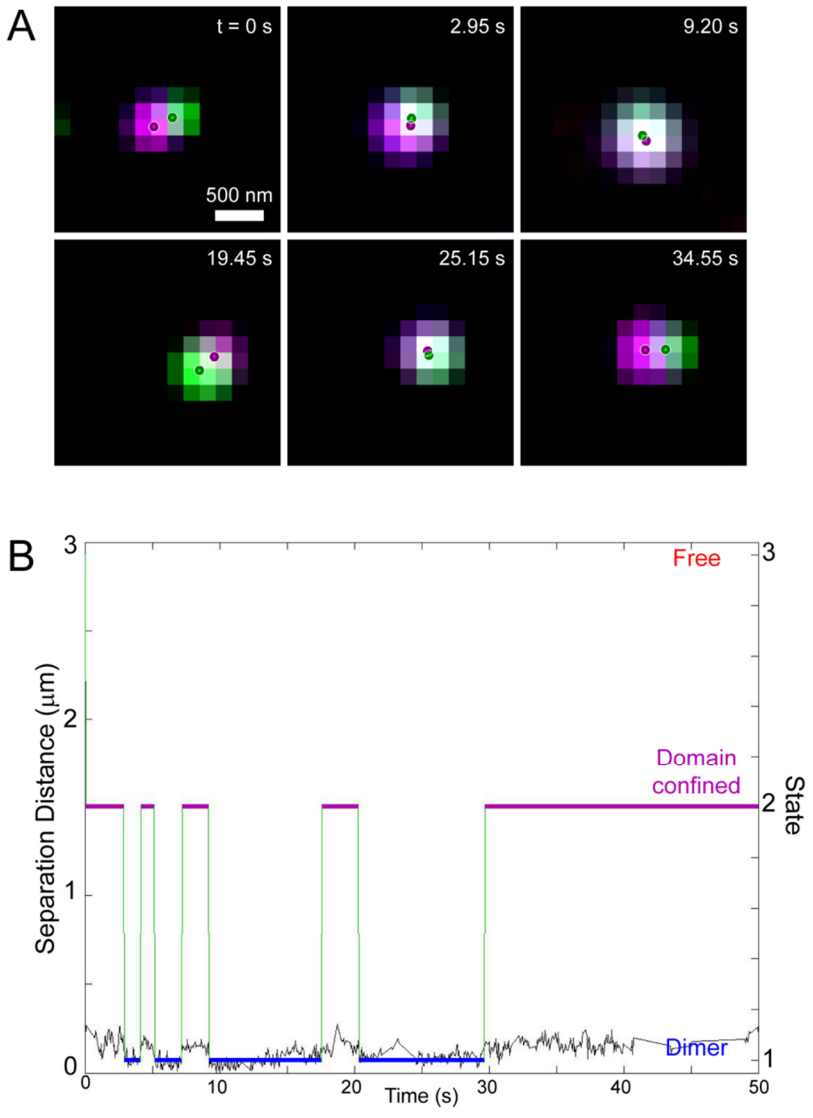
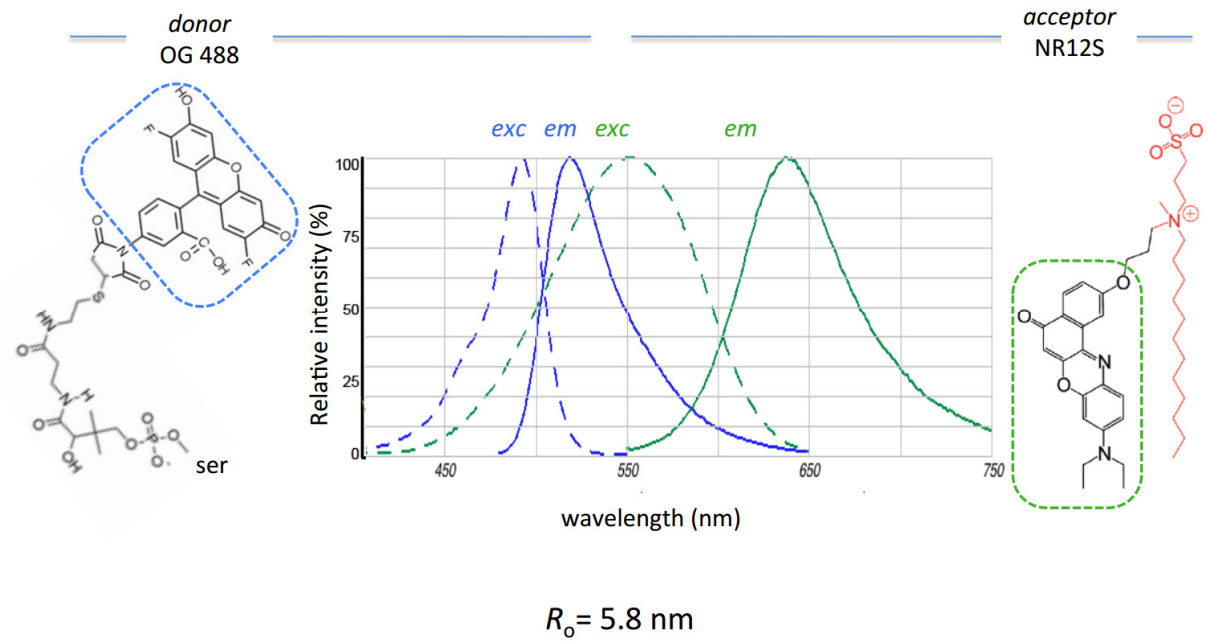



**Figure S4. (A)** Raw data (pixelated, Gaussian-filtered green/magenta image) and corresponding localizations (green/magenta circles) for two color single particle tracking of EGFR-ΔL747-P753inS in the absence of ligand illustrating stable dimerization. **(B)** The resulting Viterbi plot from Hidden Markov Model (HMM) analysis illustrating state changes over time for the dimer pairs shown above.

**Figure S5**




$R_o$= 5.8 nm

**Figure S5.** Chemical structures and normalized fluorescence spectra for Oregon Green 488 P-Pant (donor: excitation, dashed blue line; emission, solid blue line) and NR12S (acceptor: excitation, dashed green line; emission, solid green line).

**Figure S6**

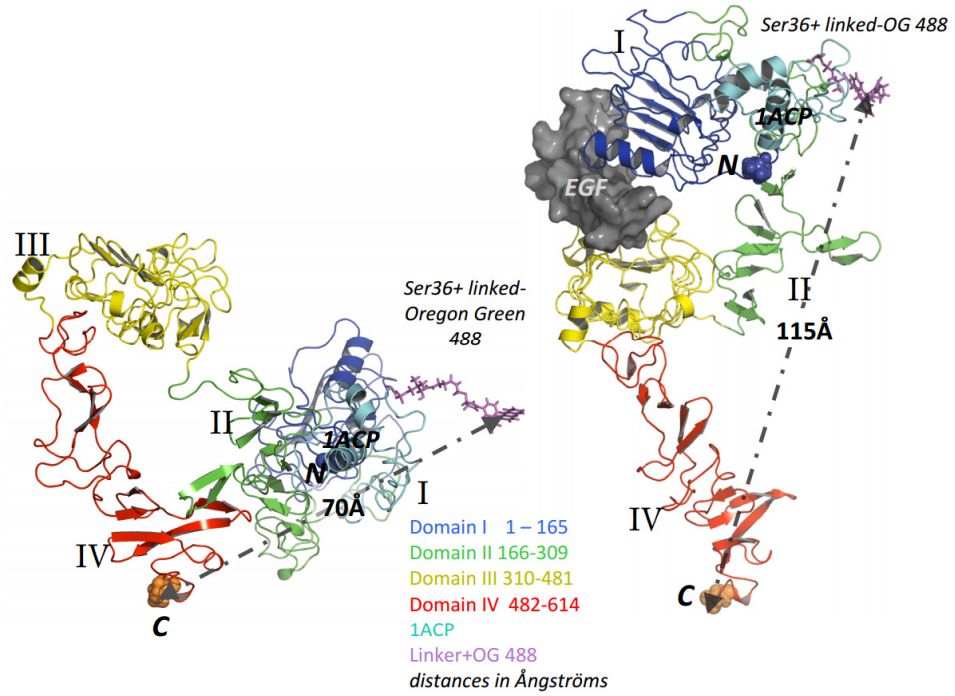



**Figure S6.** Snapshots of 10 ns MD molecular docking runs for reorientation of the ACP and the external domain structures of EGFR showing possible distances of the EGFR-ACP-tagged OG 488 to the plasma membrane. Models constructed in PyMOL with molecular dynamics calculations and minimizations made using the amber99sb-ildn force field and GROMACS 4.55. Starting structures: unliganded EGFR 1NQL + 1ACP + OG 488 (tethered conformation, left); ligand bound EGFR structures 3NJP + 1ACP + OG 488 (extended conformation, right). Distances between the N-terminal OG488 dye and the membrane-proximal receptor C-terminus are given in Ångströms.

**Figure S7**

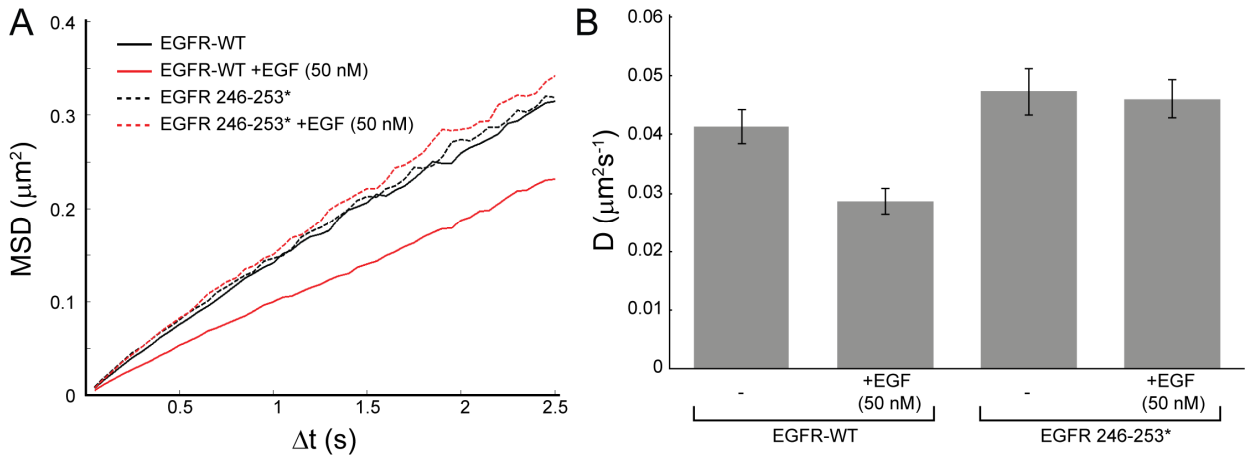



**Figure S7.** HeLa cells transiently expressing HA-EGFR-WT or HA-EGFR-246-253* having a mutated dimerization arm were tracked using anti-HA-QD in the absence or presence of saturating dark EGF (50 nM). **(A)** Mean squared displacement (MSD) plot showing changes in mobility of wild type EGFR in the absence (solid black line) and presence of saturating dark EGF (solid red line) consistent with activity, and no measurable change in mobility of the EGFR-246-253* mutant with the addition of EGF (compare dashed black and red lines). **(B)** Diffusion coefficients for EGFR-WT and EGFR-246-253* in the absence and presence of EGF. Shown is the estimated diffusion coefficient from linear regression of the MSD plot shown in (A) and the 95% confidence interval.

**Figure S8**

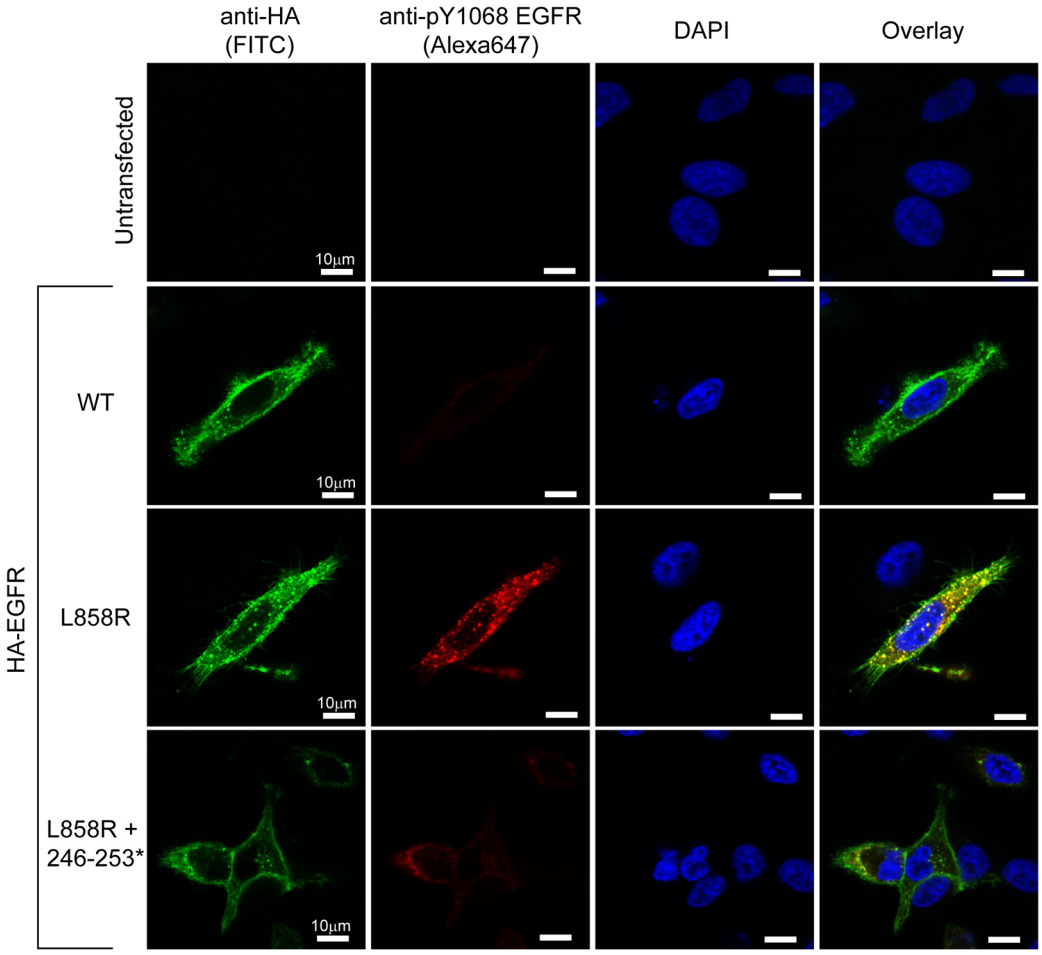



**Figure S8.** HeLa cells transiently expressing HA-EGFR-WT, HA-EGFR-L858R, or HA-EGFR-L858R+246-253* were labeled for HA-EGFR expression and phosphorylated EGFR (Y1068) and analyzed by immunofluorescence. Untransfected cells show no labeling with anti-HA (FITC) or pY1068 (Alexa647). Cells transfected with HA-EGFR-WT show robust receptor expression in the membrane but minimal levels of EGFR phosphorylation. HA-EGFR-L858R exhibits robust phosphorylation, however HA-EGFR-L858R+246-253* has diminished phosphorylation relative to HA-EGFR-L858R.

**Figure S9**

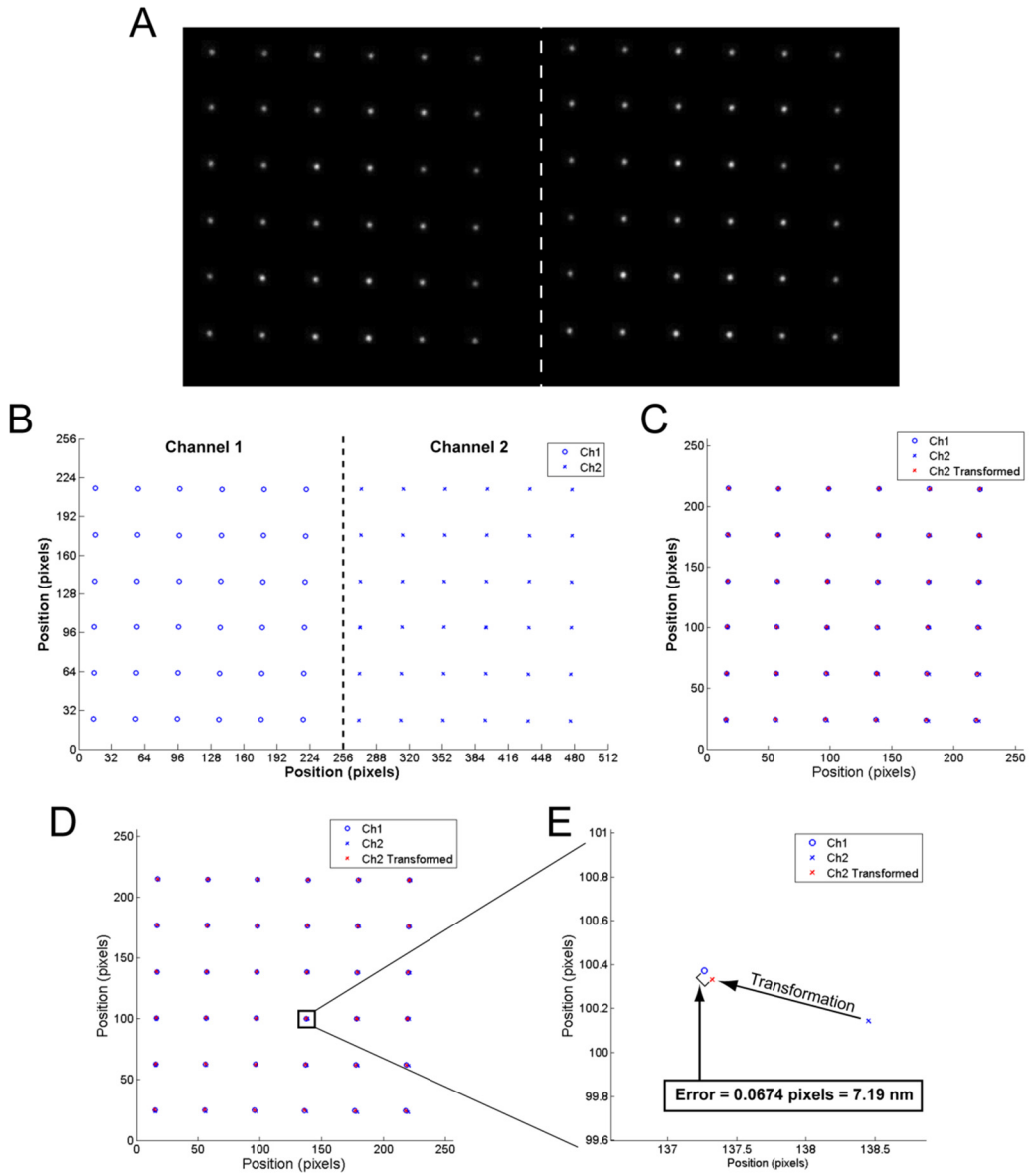



**Figure S9.** Overlay of two-color super-resolution data using fiducial channel registration and affine transformation. **(A)** Fiducial bead calibration series is collected using excitation of Tetraspeck beads in both channels, shown as left (pixels 1-256) and right (pixels 257-512) halves of an EMCCD. The white dashed line indicates approximation of the channel separation on the EMCCD. **(B)** Localization of Tetraspeck fiducial beads in a 6×6 array for channel 1 (left) and channel 2 (right). The dashed line indicates the separation between channels. **(C)** After a linear shift of channel 2 localizations by 256 pixels, an affine transformation matrix is generated to overlay channel 2 onto channel 1 with < 10 nm RMS error. **(D)** Applying the transformation matrix to a separate datasets yields good overlay of independent fiducial beads taken hours later, demonstrating the robustness of the transformation and stability of the optical setup. **(E)** Enlarged region from **(D)** showing transformation and calculation of the error for a single fiducial bead. The error for this bead is 7.2 nm. Repeating over all combinations of channel registration datasets, we estimate the overlay registration error to be 10.4 ± 2.74 nm.

**Supplementary Video Legends**

**Movie S1 – Unliganded, resting HA-EGFR-WT visualized via SPT using two-color QD-anti-HA.** CHO cells expressing HA-EGFR-WT were labeled with anti-HA-QD585 and anti-HA-QD655. Cells were imaged at 20 frames per second. The two channel data (green/magenta pixels) were localized and tracked over time (green/magenta circles and lines). The highlighted localizations and tracks correspond to the distance trace and Viterbi plot shown in Figure 2B. Note that interactions between resting receptors are transient, and that the receptors are highly mobile in the plasma membrane. Playback is 20 frames per second. Scale bar represents 0.5 μm.

**Movie S2 – Activated HA-EGFR-WT visualized via SPT using two-color QD-EGF.** CHO cells expressing HA-EGFR-WT were labeled with EGF-conjugated QD585 and QD655. Cells were imaged at 20 frames per second. The two channel data (green/magenta pixels) were localized and tracked over time (green/magenta circles and lines). The highlighted localizations and tracks correspond to the distance trace and Viterbi plot shown in Figure 2C. Notice that interactions are long lived and mobility is reduced compared to the resting receptor (resting EGFR-WT shown in Movie S1). Playback is 20 frames per second. Scale bar represents 0.5 μm.

**Movie S3 – Unliganded HA-EGFR-L858R visualized via SPT using two color QD-anti-HA.** CHO cells expressing HA-EGFR-L858R were labeled with anti-HA-QD585 and anti-HA-QD655. Cells were imaged at 20 frames per second. The two channel data (green/magenta pixels) were loalized and tracked over time (green/magenta circles and lines). The highlighted localizations and tracks correspond to the distance trace and Viterbi plot shown in Figure 2D. Notice that, relative to the unliganded EGFR-WT (Movie S1), interactions are long lived and mobility is reduced, albeit in the absence of ligand. Playback is 20 frames per second. Scale bar represents 0.5 μm.

**Movie S4 – Unliganded HA-EGFR-ΔL747-P753insS visualized via SPT using two color QD-anti-HA.** CHO cells expressing HA-EGFR-ΔL747-P753insS were labeled with anti-HA-QD585 and anti-HA-QD655. Cells were imaged at 20 frames per second. The two channel data (green/magenta pixels) were localized and tracked over time (green/magenta circles and lines). The highlighted localizations and tracks correspond to the distance trace and Viterbi plot shown in Figure S4. Again, notice that, relative to the unliganded EGFR-WT (Movie S1), interactions are long lived and mobility is reduced in the absence of ligand. Playback is 20 frames per second. Scale bar represents 0.5 μm.

# Supplementary Note - Description and discussion of the model

The data of Figure 1C reveal differences between EGFR-WT and mutant forms in their dependence of phospho-EGFR level on total EGFR level. We considered two possible explanations for these differences: (1) differences in intrinsic kinase activity and (2) differences in receptor dimerization affinity. Mathematical models were developed that incorporate these explanatory mechanisms (Models 1 and 2) and these models (and generalizations) were used to analyze the data. As discussed in the main text, both models are consistent with the data (Fig. 1C). According to Model 1, where dimerization affinity is assumed to be the same for each species, EGFR-L858R has higher kinase activity than EGFR-ΔL747-P753insS, which in turn has higher kinase activity than EGFR-WT. In contrast, according to Model 2, where intrinsic kinase activity is assumed the same for each species, EGFR-L858R dimerizes more readily than EGFR-ΔL747-P753insS, which in turn dimerizes more readily than EGFR-WT.

## The simplest model for monomer-dimer equilibrium

Models 1 and 2 are special cases of a more general model, which can be derived for diverse receptor dimerization mechanisms. The simplest mechanism is $M_X + M_X = D_X$, where $M_X$ is an EGFR monomer of form X (X = WT, ΔL747-P753insS or L858R) and $D_X$ is a homodimer of EGFR-X (i.e. dimeric EGFR-WT, EGFR-ΔL747-P753insS or EGFR-L858R). For this mechanism, at equilibrium (or more generally at steady state where detailed balance holds (Gorban and Yablonsky, 2011)), $K_{D,X}[D_X]=[M_X]^2$, where $[M_X]$ is the equilibrium abundance of monomeric EGFR-X, $[D_X]$ is the equilibrium abundance of dimeric EGFR-X, and $K_{D,X}$ is the equilibrium dissociation constant for EGFR-X dimerization. When other more complicated

reaction schemes are considered, such as a scheme wherein EGFR is taken to be bivalent (i.e., to dimerize via two interfaces, one in the ectodomain and one in the endodomain), the same relationship between monomer and dimer abundances can be found, and in these cases, $K_{D,X}$ is an *apparent* (rather than true) equilibrium dissociation constant, meaning that it is a function of multiple/other (constant) binding parameters. For a brief comparison of these more complicated models see below.

If we assume EGFR-X is conserved, such that $[R_{T,X}] = [M_X] + 2[D_X]$, where $[R_{T,X}]$ is the total abundance of EGFR-X, we find that $[M_X]$ and $[D_X]$ are related to $[R_{T,X}]$ as follows:

$$[M_X] = (K_{D,X}/4) \times ((1 + 8[R_{T,X}]/K_{D,X})^{1/2} - 1) \quad (1)$$

$$[D_X] = (K_{D,X}/16) \times ((1 + 8[R_{T,X}]/K_{D,X})^{1/2} - 1)^2 \quad (2)$$

The general model, from which Models 1 and 2 derive, is based not only on the above considerations but also on a measurement model, i.e., transformations (functions) that relate measured fluorescence intensities, which are recorded in arbitrary units (a.u.) that depend on instrument settings, to the variables in Eqs. (1) and (2), meaning the abundances of monomers, dimers, and total receptor.

The measurement model is obtained by assuming that fluorescence intensities measured in immunofluorescence assays of phosphorylated EGFR-X and total EGFR-X levels, denoted $F_{P,X}$ and $F_{R,X}$, are linearly related to $[M_X]$, $[D_X]$, and $[R_{T,X}]$ as follows:

$$F_{P,X} = a_{P,X}([D_X] + \mu_X[M_X]) + b_{P,X} \quad (3)$$

$$F_{R,X} = a_R[R_{T,X}] + b_{R,X} \quad (4)$$

where $b_{P,X}$ and $b_{R,X}$ are background fluorescence levels (a.u.) matched to receptor form X (because the behavior of each receptor form was characterized by an independent series of immunofluorescence assays, all performed on the same day with the same instrument settings);

$a_{P,X}$ is an intensive parameter reflecting the average contribution of a dimer of receptors of form X to the fluorescence intensity $F_{P,X}$; $\mu_X$ is a dimensionless ratio of intensive parameters that reflects the average contribution of a receptor monomer of form X to the fluorescence intensity $F_{P,X}$ relative to that of a receptor dimer; and $a_R$ is an intensive parameter, taken to be independent of receptor form X, reflecting the contribution of each receptor, whether in a monomeric or dimeric state, to the fluorescence intensity $F_{R,X}$. We consider $a_{P,X}$ (and $\mu_X$) to depend on receptor form because the different receptor forms potentially have different intrinsic kinase activities. Indeed, differences in kinase activities have been reported for EGFR-L858R (Zhang *et al.*, 2006). Recall that the mutations being considered affect the kinase domain of EGFR. Because EGFR autophosphorylation depends on receptor dimerization and basal phosphatase activity is high (Kleiman *et al.*, 2011), we expect $\mu_X$ to be small. We allow for a non-zero value because phosphorylated receptors leaving dimers can be expected to remain phosphorylated for some finite amount of time (Verveer *et al.*, 2000; Sawano *et al.*, 2002; Reynolds *et al.*, 2003). We assume that $a_R$ is independent of receptor form because the labeling antibody used in immunofluorescence assays is expected to recognize all receptor forms under consideration without distinction. Recall that $b_{R,X}$ corresponds to the value of $F_{R,X}$ for a cell expressing no EGFR-X (Eq. (4)), assuming even moderate levels of average EGFR expression we expect $b_{R,X}$ to be negligible.

From Eqs. (1)–(4), we obtain the following equations, which characterize the equilibrium or steady-state dependence of $F_{P,X}$ on $F_{R,X}$ for each receptor form X:

$$F_{P,X} = b_{P,X} + (a_{P,X}/a_R)(a_R \times K_{D,X})(\mu_X \times N_X/4 + N_X^2/16) \qquad (5)$$

where

$$N_X = (1 + 8(F_{R,X} - b_{R,X})/(a_R \times K_{D,X}))^{1/2} - 1 \qquad (6)$$

We used these equations for each X (WT, ΔL747-P753insS and L858R) to analyze the data of Figure 1C, focusing on two simplifications: Models 1 and 2.

Model 1 is derived from Eqs. (5) and (6) by setting $\mu_X = 0$ and $b_{R,X} = 0$ for each X (as a simplification) and by requiring that the (apparent) equilibrium dissociation constants for the different receptor forms be the same: $K_{D,WT} = K_{D,\Delta L747\text{-}P753insS} = K_{D,L858R} = K_D$. In Model 1, the ratios $a_{P,WT}/a_R$, $a_{P,\Delta L747\text{-}P753insS}/a_R$ and $a_{P,L858R}/a_R$, which reflect the intrinsic kinase activities of the different receptor forms, are allowed to have different values. Model 2 is derived from Eqs. (5) and (6) by setting $\mu_X = 0$ and $b_{R,X} = 0$ for each X (as a simplification) and by requiring that the kinase activities of the different receptor forms be the same, such that $a_{P,WT}/a_R = a_{P,\Delta L747\text{-}P753insS}/a_R = a_{P,L858R}/a_R = a_P/a_R$.

In Model 2, the (apparent) equilibrium dissociation constants are allowed differ, allowing the quantities $a_R \times K_{D,WT}$, $a_R \times K_{D,\Delta L747\text{-}P753insS}$ and $a_R \times K_{D,L858R}$ to have different values. These values are expressed in the same arbitrary units of the fluorescence intensity measurements. The free (and potentially identifiable) parameters of Model 1 can be taken as $K_D$, $a_{P,WT}/a_R$, $a_{P,\Delta L747\text{-}P753insS}/a_R$, $a_{P,L858R}/a_R$, $b_{P,WT}$, $b_{P,\Delta L747\text{-}P753insS}$ and $b_{P,L858R}/a_R$. The free (and potentially identifiable) parameters of Model 2 can be taken as $a_P/a_R$, $K_{D,WT}/a_R$, $K_{D,\Delta L747\text{-}P753insS}/a_R$, $K_{D,L858R}/a_R$, $b_{P,WT}$, $b_{P,\Delta L747\text{-}P753insS}$ and $b_{P,L858R}/a_R$. We note that three parameters in each model, the background fluorescence intensities ($b_{P,X}$ for each receptor form X), can be readily estimated by linear extrapolation from the data at low $F_{R,X}$ values in Fig. 1C to the *y*-axis, where $F_{R,X} = 0$, because $b_{P,X}$ is defined as the value of $F_{P,X}$ where $[R_{T,X}] = 0$ (Eq. (3)). Under our assumption that $b_{R,X} = 0$, $F_{R,X} = 0$ where $[R_{T,X}] = 0$. Evidently, $b_{P,X}$ is non-zero for each X (Fig. 1C). We assume that the non-zero background reflects a combination of non-specific antibody binding and/or microscope acquisition settings.

To estimate the values of the free parameters of Models 1 and 2, we used chi-square fitting, which is also known as weighted nonlinear least squares fitting. Minimal values of the chi-square function, which is discussed below, were found via the modified Levenberg-Marquardt algorithm implemented in LMDIF (Moré *et al.*, 1980) with the following input settings: `FTOL` = `XTOL` = `GTOL` = $10^{-15}$, `MAXFEV` = 1000, `EPSFCN` = 0.1, `MODE` = 1, `FACTOR` = 1, and `NPRINT` = 0. In fitting, we considered all of the data shown in Fig. 1C together. In other words, our best-fit parameter estimates represent the results of a global fit. The data in Figure 1C summarize more than 500 pairs of fluorescence intensity measurements from three series of receptor form X-matched immunofluorescence assays. For analysis, pairs of intensities were divided into bins centered on integer values of the labeling anti-EGFR fluorescence intensity, from 20 to 211 (a.u.). Accordingly, in Fig. 1C, the *x*-value of a data point reports the bin (which corresponds to the nearest integer value for measured anti-EGFR fluorescence intensities in the bin), and the *y*-value of a data point reports the empirical mean of the binned labeling anti-phosphotyrosine fluorescence intensities (a.u.). In the chi-square function, the calculated values are the values of $F_{P,X}$ obtained from Eqs. (5) and (6) for $F_{R,X}$ values from 20 to 211 (where observations were made) for each receptor form X (EGFR-WT, EGFR-ΔL747-P753insS and EGFR-L858R), the observed values are the empirical means for bins 20 to 211, and the weights are the inverses of either empirical standard deviations (for bins 20 through 80) or estimates of standard deviations (for bins 81 through 211). Estimation of dispersion, which increases with increasing anti-EGFR intensity (based on inspection of the data), was necessary because observations where the anti-EGFR fluorescence intensity is above 80 are limited. A simple function was deemed suitable for estimating standard deviations: $\log_{10} \sigma_X^2 = m \times F_{R,X} + b_X$, where $\sigma_X$ is the standard deviation expected at a given anti-EGFR fluorescence intensity, >80 (a.u.).

This function has one receptor form X-independent parameter (m) and one receptor form X-specific parameter ($b_X$), which were set at the following values: m = 0.0042, $b_{WT}$ = 1.35, $b_{\Delta L747\text{-}P753insS}$ = 2.37, and $b_{L858R}$ = 2.775.

We found 68% confidence limits on parameter estimates through bootstrapping, which involved repeated fitting to samples of the original data (Efron and Tibshirani, 1994; Press *et al.*, 2007). Sampling (with replacement) has the effect of randomly assigning weights to the data points. For each determination of a confidence limit, we considered 10,000 bootstrap samples of the data. Fitting was performed as described above. Each fitting run began with randomly selected starting values for the free parameters, i.e., at a randomly selected point in parameter space. Fitting runs sometimes failed to converge (i.e., LMDIF terminated before satisfaction of a convergence test). In these cases, fitting was simply restarted from a new randomly selected point in parameter space. At the end of a fitting run, quality of fit was checked by comparing the final chi-square value ($\chi_B^2$) to the final chi-square value obtained in fitting to the original (unweighted) data ($\chi^2$). In cases where $\chi_B^2 > 2\chi^2$ at convergence, the results were rejected (under the assumption that the fitting routine converged to a local minimum far from the global minimum) and fitting was restarted from a new randomly selected point in parameter space. In the vast majority of cases, the final accepted $\chi_B$ was close to $\chi$.

Although Models 1 and 2 both reasonably fit the data (Fig. 1C), the two models make distinct predictions about the extent of dimerization as a function of total receptor abundance, as can be seen in Figure S10. We note that Models 1 and 2 provide explanations of the data at opposite extremes. Phosphorylation differences are explained by kinase activity differences alone in the case of Model 1 and dimerization affinity differences alone in the case of Model 2. Because both models explain the data of Figure 1C, it is conceivable that the altered

phosphorylation of EGFR mutants and the observed differences in dependence of phospho-EGFR level on total EGFR expression may arise from a combination of both kinase activity and dimerization differences.

As discussed above, Models 1 and 2 are based on certain assumptions that constrain the values of the parameters in Eqs. (5) and (6). One assumption that is common to both models is that monomeric receptors contribute negligibly to the observed phosphorylation signal. To assess this assumption, we allowed non-zero values for $\mu_{WT}$, $\mu_{\Delta L747\text{-}P753insS}$ and $\mu_{L858R}$, obtaining an extension of Model 1 (Extended Model 1), and a non-zero value for $\mu$ (= $\mu_{WT}$ = $\mu_{\Delta L747\text{-}P753insS}$ = $\mu_{L858R}$), obtaining an extension of Model 2 (Extended Model 2). We then estimated parameters and confidence limits as described above for Extended Models 1 and 2. The results, which are given in Table S4, indicate that monomers, as expected, contribute much less to observed receptor phosphorylation than dimers. Quality of fit is not dramatically improved by accounting for monomer phosphorylation. The most improvement in quality of fit is obtained for the EGFR-WT data when considering Extended Model 2.

**Table S3.** Parameter estimates for Models 1 and 2.

| | **Model 1** ($\chi^2$ = 105) | | | **Model 2** ($\chi^2$ = 113) | | |
|---|---|---|---|---|---|---|
| Parameter | Best-fit value | Lower bound of 68% confidence limit | Upper bound of 68% confidence limit | Best-fit value | Lower bound of 68% confidence limit | Upper bound of 68% confidence limit |
| $a_R \times K_{D,WT}$ | 74.6 | 51.5 | 103 | 3440 | 2970 | 3970 |
| $a_{P,WT}/a_R$ | 0.394 | 0.357 | 0.432 | 2.61 | 2.33 | 2.90 |
| $\mu_{WT}$ | 0* | — | — | 0* | — | — |
| $b_{P,WT}$ | 12.7 | 12.4 | 13.0 | 15.4 | 15.1 | 15.7 |
| $a_R \times K_{D,\Delta L747\text{-}P753insS}$ | Same as WT | — | — | 379 | 290 | 476 |
| $a_{P,\Delta L747\text{-}P753insS}/a_R$ | 1.62 | 1.49 | 1.75 | Same as WT | — | — |
| $\mu_{\Delta L747\text{-}P753insS}$ | 0* | — | — | Same as WT | — | — |
| $b_{P,\Delta L747\text{-}P753insS}$ | 18.7 | 17.7 | 19.6 | 24.4 | 23.4 | 25.2 |
| $a_R \times K_{D,L858R}$ | Same as WT | — | — | 12.6 | 0.938 | 32.6 |
| $a_{P,L858R}/a_R$ | 3.33 | 3.09 | 3.61 | Same as WT | — | — |
| $\mu_{L858R}$ | 0* | — | — | Same as WT | — | — |
| $b_{P,L858R}$ | 17.4 | 15.4 | 19.0 | 9.11 | 3.80 | 13.2 |
| $b_R$ | 0* | — | — | 0* | — | — |

*Fixed value (i.e., the value of this parameter was not allowed to vary in fitting)

Parameter values given in this table are dimensionless or have arbitrary units.

**Table S4.** Parameter estimates for Extended Models 1 and 2.

| | **Extended Model 1** ($\chi^2$ = 105) | | | **Extended Model 2** ($\chi^2$ = 107) | | |
|---|---|---|---|---|---|---|
| Parameter | Best-fit value | Lower bound of 68% confidence limit | Upper bound of 68% confidence limit | Best-fit value | Lower bound of 68% confidence limit | Upper bound of 68% confidence limit |
| $a_R \times K_{D,WT}$ | 82.8 | 71.7 | 168 | 60200 | 22200 | Large |
| $a_{P,WT}/a_R$ | 0.349 | 0.263 | 0.438 | 2.98 | 2.68 | 3.28 |
| $\mu_{WT}$ | 0.161 | 0.00163 | 0.511 | 0.0429 | 0.0371 | 0.0489 |
| $b_{P,WT}$ | 11.8 | 10.5 | 12.6 | 10.5 | 10.1 | 11.0 |
| $a_R \times K_{D,\Delta L747\text{-}P753insS}$ | Same as WT | — | — | 707 | 564 | 855 |
| $a_{P,\Delta L747\text{-}P753insS}/a_R$ | 1.65 | 1.46 | 1.92 | Same as WT | — | — |
| $\mu_{\Delta L747\text{-}P753insS}$ | Small (≈0) | Small (≈0) | 0.132 | Same as WT | — | — |
| $b_{P,\Delta L747\text{-}P753insS}$ | 19.0 | 16.2 | 20.4 | 22.3 | 21.5 | 23.0 |
| $a_R \times K_{D,L858R}$ | Same as WT | — | — | 48.9 | 21.8 | 80.8 |
| $a_{P,L858R}/a_R$ | 3.40 | 3.13 | 3.77 | Same as WT | — | — |
| $\mu_{L858R}$ | 0.00404 | Small (≈0) | 0.124 | Same as WT | — | — |
| $b_{P,L858R}$ | 17.7 | 13.0 | 18.9 | 13.4 | 10.6 | 15.6 |
| $b_R$ | 0* | — | — | 0* | — | — |

*Fixed value (i.e., the value of this parameter was not allowed to vary in fitting)

Parameter values given in this table are dimensionless or have arbitrary units.

**Figure S10**

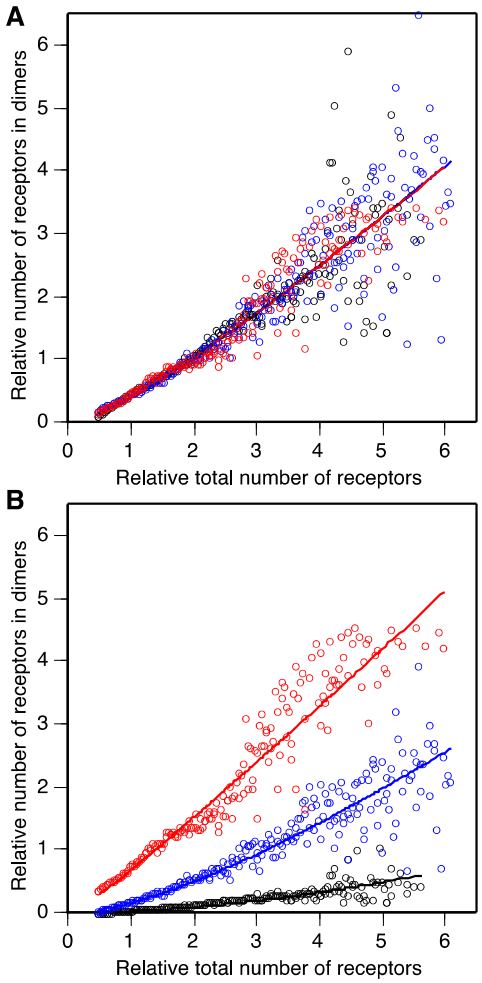



**Figure S10.** This figure displays the same information as Fig. 1C, except that fluorescence intensity measurements and calculated values for $F_{R,X}$ and $F_{P,X}$ (Eqs. (5) and (6)) have been transformed using Eqs. (3) and (4) and the values of $a_{P,X}$ and $b_{P,X}$ determined for **(A)** Model 1 and **(B)** Model 2. The horizontal axis reports $F_{R,X}/\langle F_{R,X}\rangle = [R_{T,X}]/[R_{T,X}]_0$, where $\langle F_{R,X}\rangle$=40.77 (a.u.) is the mean labeling intensity of EGFR-X and $[R_{T,X}]_0$ is the corresponding (unknown) EGFR-X copy number, for EGFR-WT data (black circles), EGFR-ΔL747-P753insS (blue circles), and EGFR-L858R data (red circles). The vertical axis reports $2(F_{P,X} - b_{P,X})/((a_{P,X}/a_R)\langle F_{R,X}\rangle) = 2[D_X]/[R_{T,X}]_0$ (i.e., the relative number of receptors in dimers). The curves shown in the top and bottom panels are obtained from Eqs. (5) and (6) and the parameter values given in Table S3.

# The equivalence of various equilibrium models for ligand-independent dimerization of EGFR

## 1 The simplest model for monomer-dimer equilibrium

Consider the following reaction scheme:

$$M + M \rightleftharpoons D \tag{7}$$

where $M$ represents a receptor monomer and $D$ represents a receptor dimer. At equilibrium,

$$D = KM^2 \tag{8}$$

where $K$ is the equilibrium association constant and $M$ and $D$ now represent abundances. Assuming conservation of receptor, we have

$$R_T = M + 2D = M + 2KM^2 \tag{9}$$

where $R_T$ is the total abundance of receptor. Solving the quadratic equation $2KM^2 + M - R_T = 0$, we find

$$M = \frac{-1 + \sqrt{1 + 8KR_T}}{4K} \tag{10}$$

From Eqs. (8) and (10), we find

$$D = \frac{(-1 + \sqrt{1 + 8KR_T})^2}{16K} \tag{11}$$

Let us assume linear relationships between abundances and fluorescence intensities, such that

$$F_P = a_P D + a'_P M + b_P \tag{12}$$

and

$$F_R = a_R R_T + b_R \tag{13}$$

where $F_P$ is the fluorescence intensity of anti-pY, $F_R$ is the fluorescence intensity of anti-EGFR, $b_P$ and $b_R$ are background intensities, and $a_P$, $a'_P$, and $a_R$ are scaling factors. In Eq. (12), the terms $a_P D$ and $a'_P M$ reflect the contributions of dimers and monomers, respectively, to the anti-pY fluorescence intensity. Rearranging, we find

$$F_P - b_P = (a_P/a_R)(a_R K_D)(\mu N/4 + N^2/16) \tag{14}$$

where $K_D \equiv 1/K$ is the equilibrium dissociation constant, $\mu \equiv a'_P/a_P$, and

$$N \equiv (1 + 8(F_R - b_R)/(a_R K_D))^{1/2} - 1 \tag{15}$$

In the above equations, there are only five identifiable parameters that relate $F_P$ and $F_R$, namely, $b_P$, $b_R$, $a_P/a_R$, $\mu$, and $a_R K_D$.

# 2 A model that accounts for two receptor dimerization interfaces and EGFR oligomers up to cyclic receptor dimers

Consider the following reaction scheme:

$$\begin{array}{ll} M + M \rightleftharpoons D_x & D_x \rightleftharpoons D_{xy} \\ M + M \rightleftharpoons D_y & D_y \rightleftharpoons D_{xy} \end{array} \tag{16}$$

where $M$ represents a receptor monomer, $D_x$ represents a receptor dimer connected via an ectodomain interface, $D_y$ represents a receptor dimer connected via an endodomain interface, and $D_{xy}$ represents a (cyclic) receptor dimer connected via both ecto- and endodomain interfaces. We neglect higher-order EGFR oligomers. At equilibrium,

$$D_x = K_x M^2 \quad D_y = K_y M^2 \quad D_{xy} = J_y D_x = J_x D_y \tag{17}$$

where $K_x$ is an equilibrium association constant that characterizes the ectodomain interface; $K_y$ is an equilibrium association constant that characterizes the endodomain interface; $J_x$ and $J_y$ are equilibrium constants that characterize cycle formation; and $M$, $D_x$, $D_y$, and $D_{xy}$ now represent abundances. For consistency with the principle of detailed balance, the following constraint must be satisfied:

$$K_x J_y = K_y J_x \tag{18}$$

Let us introduce the following definitions:

$$\begin{array}{lcl} \phi & \equiv & K_x J_y = K_y J_x \\ K & \equiv & K_x + K_y + \phi = K_x + K_y + K_x J_y = K_x + K_y + K_y J_x \\ D & \equiv & D_x + D_y + D_{xy} \end{array} \tag{19}$$

From Eqs. (17) and (19), it follows that $D = (K_x + K_y + \phi)M^2 = KM^2$. For conservation of receptor, we have

$$R_T = M + 2(D_x + D_y + D_{xy}) = M + 2D = M + 2KM^2 \tag{20}$$

where $R_T$ is the total abundance of receptor. As before, we take $F_P$ to be linearly related to the abundances of EGFR monomers and dimers and $F_R$ to be linearly related to the total abundance of EGFR (Eqs. (12) and (13)). Furthermore, as in the simplest model (Section 1), we have the following relations: $R_T = M + 2D$ and $D = KM^2$. Thus, under the assumption of Eq. (16) instead of Eq. (7), Eqs. (14) and (15) given earlier can be used to relate $F_P$ and $F_R$ with reinterpretation of the parameters. For example, $K_D$ in Eqs. (14) and (15) is now taken to be a function of the equilibrium constants that parameterize the reaction scheme of Eq. (16):

$$K_D \equiv 1/K = 1/(K_x + K_y + \phi) \tag{21}$$

# 3 A model that accounts for tethered and extended conformations of monomeric EGFR and one dimerization interface

Consider the following reaction scheme:

$$M_e \rightleftharpoons M_t \quad M_e + M_e \rightleftharpoons D \tag{22}$$

where $M_e$ represents a receptor monomer in an extended conformation, $M_t$ represents a receptor monomer in tethered conformation, and $D$ represents a receptor dimer. We assume that only extended receptors are competent for dimerization. At equilibrium,

$$M_t = K_0 M_e \quad D = K_1 M_e^2 \tag{23}$$

where $K_0$ is an equilibrium constant that characterizes the equilibrium distribution of receptor monomers into tethered and extended forms and $K_1$ is an equilibrium association constant that characterizes dimerization of monomers in the extended conformation. For conservation of receptor, we have

$$R_T = M_e + M_t + 2D = M + 2D = (1 + K_0)M_e + 2K_1 M_e^2 \tag{24}$$

where $R_T$ is the total abundance of receptor and $M \equiv M_e + M_t$ is the abundance of receptor monomer (in either conformation). Let us introduce the following definitions:

$$\begin{array}{rcl} f_e & \equiv & M_e/M = 1/(1 + K_0) \\ f_t & \equiv & M_t/M = 1 - f_e = K_0/(1 + K_0) \\ K & \equiv & K_1/(1 + K_0)^2 \end{array} \tag{25}$$

where $f_e$ ($f_t$) is the fraction of receptor monomers in the extended (tethered) conformation at equilibrium. With these definitions, it follows that $R_T = M + 2D$ and $D = KM^2$. These relationships are consistent with those derived from the reaction scheme of Eq. (7) (Section 1). Thus, we can relate $F_P$ and $F_R$ using Eqs. (14) and (15) with reinterpretation of the parameters. For example, we now interpret $K_D$ in these equations as a function of the equilibrium constants that parameterize the reaction scheme of Eq. (22):

$$K_D \equiv 1/K = (1 + K_0)^2/K_1 \tag{26}$$

# 4 A model that accounts for tethered and extended conformations of monomeric EGFR, two receptor dimerization interfaces, and EGFR oligomers up to cyclic receptor dimers

Let us consider the following reaction scheme:

$$\begin{array}{ll} M_e \rightleftharpoons M_t & \\ M_e + M_e \rightleftharpoons D_x & D_x \rightleftharpoons D_{xy} \\ M_e + M_e \rightleftharpoons D_y & D_y \rightleftharpoons D_{xy} \end{array} \tag{27}$$

The nomenclature is the same as that introduced earlier (see Sections 2 and 3). We are assuming that only receptor monomers in the extended conformation are competent for dimerization and we are neglecting oligomers of EGFR larger than dimers. At equilibrium,

$$M_t = K_0 M_e \quad D_x = K_x M_e^2 \quad D_y = K_y M_e^2 \quad D_{xy} = J_y D_x = J_x D_y \tag{28}$$

where $K_x J_y = K_y J_x$ from the principle of detailed balance. We note that $f_e \equiv M_e/M = 1/(1+K_0)$ and $f_t \equiv M_t/M = 1 - f_e = K_0/(1+K_0)$ are the fractions of extended and tethered receptor monomers, respectively. For conservation of receptor, we have

$$R_T = M_e + M_t + 2(D_x + D_y + D_{xy}) = M + 2D = (1+K_0)M_e + (K_x + K_y + \phi)M_e^2 \tag{29}$$

where $R_T$ is the overall abundance of EGFR, $M \equiv M_e + M_t = (1+K_0)M_e$, $D \equiv D_x + D_y + D_{xy}$, and $\phi \equiv K_x J_y = K_y J_x$. Let us introduce the following definition:

$$K \equiv \frac{K_x + K_y + \phi}{(1+K_0)^2} \tag{30}$$

With the definitions and relations introduced above, we can write $R_T = M + 2D$ and $D = KM^2$. These relationships are consistent with those derived from the reaction scheme of Eq. (7) (Section 1). Thus, we can relate $F_P$ and $F_R$ using Eqs. (14) and (15) with reinterpretation of the parameters. For example, we now interpret $K_D$ in these equations as a function of the equilibrium constants that parameterize the reaction scheme of Eq. (27):

$$K_D \equiv 1/K = (1+K_0)^2/(K_x + K_y + \phi) \tag{31}$$

.

# 5 A model wherein dimerization occurs only between EGFR monomers in distinct conformations

Consider the following reaction scheme:

$$\begin{aligned} M_0 &\rightleftharpoons M_a \\ M_0 &\rightleftharpoons M_b \\ M_a + M_b &\rightleftharpoons D \end{aligned} \tag{32}$$

where $M_0$ represents a monomer that is incompetent for dimerization, $M_a$ and $M_b$ represent monomers in conformations competent for dimerization, and $D$ represents a dimer. At equilibrium,

$$M_a = K_a M_0 \quad M_b = K_b M_0 \quad D = K' M_a M_b \tag{33}$$

where $K_a$ and $K_b$ are equilibrium constants that characterize the distribution of receptor monomers into the three possible conformations (0, $a$ and $b$) and $K'$ characterizes dimerization of monomers in the complementary $a$ and $b$ conformations. For conservation of receptor, we have

$$R_T = M_0 + M_a + M_b + 2D = M + 2D \tag{34}$$

where $R_T$ is the overall abundance of EGFR and $M \equiv M_0 + M_a + M_b$ is the total abundance of receptor monomers (i.e., the abundance of all monomers, regardless of conformation). It can be shown that $D = KM^2$, where

$$K \equiv \frac{K_a K_b K'}{(1 + K_a + K_b)^2} \tag{35}$$

The relationships $R_T = M + 2D$ and $D = KM^2$ are consistent with those derived from the reaction scheme of Eq. (7) (Section 1). Thus, we can relate $F_P$ and $F_R$ using Eqs. (14) and (15) with reinterpretation of the parameters. For example, we now interpret $K_D$ in these equations as a function of the equilibrium constants that parameterize the reaction scheme of Eq. (32):

$$K_D \equiv 1/K = (1 + K_a + K_b)^2/(K_a K_b K') \tag{36}$$

# 6 Relating fluorescence intensities to model variables

Throughout the discussion above, we have assumed that fluorescence intensities are related to model variables in accordance with Eqs. (12) and (13), which were derived within the context of the simplest model for monomer-dimer equilibrium (Section 1). Here, we will generalize the first of these equations, considering the possibility that monomeric and dimeric receptors in distinct states contribute differentially to the anti-pY fluorescence intensity, $F_P$. Let us focus on the model of Section 4, which considers two types of monomeric receptors (i.e., monomers in extended and tethered conformations) and three types of dimeric receptors (i.e., dimers connected via an ectodomain interface only, dimers connected via an endodomain interface only, and dimers connected via both ectodomain and endodomain interfaces). We will assume that a linear relationship exists between the abundance of a molecular species and its contribution to the total fluorescence intensity. Thus, we have

$$F_P = a_{P,x}D_x + a_{P,y}D_y + a_{P,xy}D_{xy} + a'_{P,e}M_e + a'_{P,t}M_t + b_P \tag{37}$$

where $b_P$ is the background fluorescence intensity and $a_{P,x}$, $a_{P,y}$, $a_{P,xy}$, $a'_{P,e}$, and $a'_{P,t}$ are scaling factors. Each factor determines how the abundance of a molecular species contributes to the total fluorescence intensity $F_P$. We will show that Eq. (37) reduces to a form identical to that of Eq. (12). In this demonstration, we will use the following relations, which follow from the results presented in Section 4:

$$\begin{array}{rcl} M_e/M & = & 1/(1+K_0) \\ M_t/M & = & K_0/(1+K_0) \\ D_x/D & = & K_x/(K_x+K_y+\phi) \\ D_y/D & = & K_y/(K_x+K_y+\phi) \\ D_{xy}/D & = & \phi/(K_x+K_y+\phi) \end{array} \tag{38}$$

The above relations allow us to rewrite Eq. (37) as follows:

$$F_P = [(a_{P,x}K_x+a_{P,y}K_y+a_{P,xy}\phi)/(K_x+K_y+\phi)]D+[(a'_{P,e}+a'_{P,t}K_0)/(1+K_0)]M+b_P \tag{39}$$

Simplifying, we obtain an expression that has the same form as Eq. (12):

$$F_P = \hat{a}_P D + \hat{a}'_P M + b_P \tag{40}$$

where

$$\hat{a}_P = \frac{a_{P,x}K_x + a_{P,y}K_y + a_{P,xy}\phi}{K_x + K_y + \phi} \tag{41}$$

and

$$\hat{a}'_P = \frac{a'_{P,e} + a'_{P,t}K_0}{1 + K_0} \tag{42}$$

As can be seen from Eqs. (41) and (42), the parameters $\hat{a}_P$ and $\hat{a}'_P$ in Eq. (40) are weighted sums of scaling factors that appear in Eq. (37). Each weight is a function of one or more equilibrium constants. Importantly, the values of $\hat{a}_P$ and $\hat{a}'_P$ are not determined by the value $K_D$ (Eq. (31)), i.e., these parameters can be estimated independently.

# 7 Summary

Measured fluorescence intensities of anti-pY and anti-EGFR can be related using Eqs. (14) and (15) for a variety of ligand-independent receptor dimerization mechanisms. However, the interpretation of the parameter $K_D$ in these equations depends on the reaction scheme being considered. In the simplest scheme, that of Section 1, $K_D$ is simply the equilibrium dissociation constant for receptor dimerization. In a more complicated scheme (Sections 2–5), $K_D$ is an apparent equilibrium constant, which is a function of multiple equilibrium constants. A limitation of the analysis summarized here is that we have neglected oligomers of EGFR larger than dimers. Reaction schemes that describe higher-order oligomer formation would require modification of Eqs. (14) and (15). We are essentially assuming that in the absence of ligand the predominant EGFR oligomer is a dimer. This assumption in not necessarily incompatible with ligand-induced formation of higher-order oligomers.

## Supplementary References


Efron, B., and Tibshirani, R. J. (1994). An Introduction to the Bootstrap, Taylor & Francis.

Endres, N. F. *et al.* (2013). Conformational Coupling across the Plasma Membrane in Activation of the EGF Receptor. Cell *152*, 543–556.

Gorban, A. N., and Yablonsky, G. S. (2011). Extended detailed balance for systems with irreversible reactions. Chem. Eng. Sci. *66*, 5388–5399.

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

Low-Nam, S. T., Lidke, K. A., Cutler, P. J., Roovers, R. C., van Bergen en Henegouwen, P. M. P., Wilson, B. S., and Lidke, D. S. (2011). ErbB1 dimerization is promoted by domain co-confinement and stabilized by ligand binding. Nat. Struct. Mol. Biol. *18*, 1244–1249.

Moré, J. J., Garbow, B. S., and Hillstrom, K. E. (1980). User Guide for MINPACK-1, Argonne National Laboratory.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). Numerical Recipes 3rd Edition: The Art of Scientific Computing, New York, NY, USA: Cambridge University Press.

Reynolds, A. R., Tischer, C., Verveer, P. J., Rocks, O., and Bastiaens, P. I. H. (2003). EGFR activation coupled to inhibition of tyrosine phosphatases causes lateral signal propagation. Nat. Cell Biol. *5*, 447–453.

Sawano, A., Takayama, S., Matsuda, M., and Miyawaki, A. (2002). Lateral propagation of EGF signaling after local stimulation is dependent on receptor density. Dev. Cell *3*, 245–257.

Verveer, P. J., Wouters, F. S., Reynolds, a R., and Bastiaens, P. I. (2000). Quantitative imaging of lateral ErbB1 receptor signal propagation in the plasma membrane. Science *290*, 1567–1570.

Zhang, X., Gureasko, J., Shen, K., Cole, P. A., and Kuriyan, J. (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. Cell *125*, 1137–1149.