



Published in final edited form as:

Cell. 2017 October 19; 171(3): 683–695.e18. doi:10.1016/j.cell.2017.09.017.

## EGFR ligands differentially stabilize receptor dimers to specify signaling kinetics

Daniel M. Freed<sup>1,2,7</sup>, Nicholas J. Bessman<sup>3,6,7</sup>, Anatoly Kiyatkin<sup>1,2</sup>, Emanuel Salazar-Cavazos<sup>4</sup>, Patrick O. Byrne<sup>5</sup>, Jason O. Moore<sup>3</sup>, Christopher C. Valley<sup>4</sup>, Kathryn M. Ferguson<sup>1,2</sup>, Daniel J. Leahy<sup>5</sup>, Diane S. Lidke<sup>4</sup>, and Mark A. Lemmon<sup>1,2,3,8</sup>

<sup>1</sup>Department of Pharmacology, Yale University School of Medicine, New Haven, CT 06520, USA

<sup>2</sup>Yale Cancer Biology Institute, Yale University, West Haven, CT 06516, USA

<sup>3</sup>Department of Biochemistry and Biophysics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA 19104-6059, USA

<sup>4</sup>Department of Pathology and UNM Comprehensive Cancer Center, University of New Mexico Health Science Center, Albuquerque, NM 87131, USA

<sup>5</sup>Department of Molecular Biosciences, University of Texas at Austin, Austin, TX 78712, USA

### Summary

Epidermal growth factor receptor (EGFR) regulates many crucial cellular programs, with seven different activating ligands shaping cell signaling in distinct ways. Using crystallography and other approaches, we show how the EGFR ligands epiregulin (EREG) and epigen (EPGN) stabilize different dimeric conformations of the EGFR extracellular region. As a consequence, EREG or EPGN induce less stable EGFR dimers than EGF – making them partial agonists of EGFR dimerization. Unexpectedly, this weakened dimerization elicits more sustained EGFR signaling than seen with EGF, provoking responses in breast cancer cells associated with differentiation rather than proliferation. Our results reveal how responses to different EGFR ligands are defined by receptor dimerization strength and signaling dynamics. These findings have broad implications for understanding receptor tyrosine kinase (RTK) signaling specificity. Our results also suggest parallels between partial and/or biased agonism in RTKs and G protein-coupled receptors, as well as new therapeutic opportunities for correcting RTK signaling output.

Correspondence to: Mark A. Lemmon.

<sup>6</sup>Present address: Department of Medicine, Weill Cornell Medicine, New York, NY 10021, USA

<sup>7</sup>These authors contributed equally

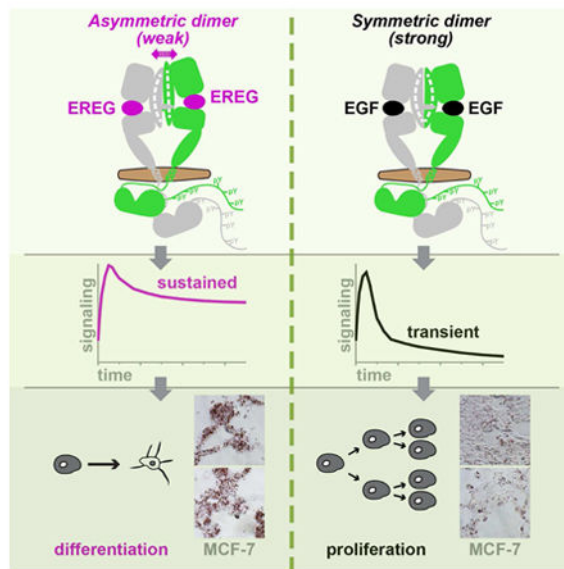
<sup>8</sup>Lead contact

Supplemental Information: Supplemental information including seven figures and two tables can be found in the online version of this article.

**Author Contributions:** D.M.F., N.J.B., and M.A.L. designed the overall project, analyzed structures, and wrote the manuscript. D.M.F. and N.J.B. generated all materials and performed crystallographic and solution biophysical studies (assisted by J.O.M., K.M.F., and M.A.L.). A.K., D.M.F., or N.J.B. performed cell-signaling studies. P.O.B. and D.J.L. performed FRET studies in membranes, and E.S.C. and D.S.L. carried out single particle tracking studies (assisted by C.C.V.) All authors contributed to analysis of results and editing of the manuscript.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Graphical abstract



## Keywords

receptor tyrosine kinase; dimerization; crystallography; kinetic proofreading; growth factor; signaling specificity

## Introduction

Structural understanding of receptor tyrosine kinase (RTK) activation has advanced greatly in recent years (Lemmon and Schlessinger, 2010), with the epidermal growth factor receptor (EGFR/ErbB1) being among the most intensely studied (Kovacs et al., 2015). A primary focus has been to understand how ligand binding drives RTKs from their inactive to active forms, assuming that a single ligand-induced dimeric species represents the active state. Although this assumption may be reasonable for RTKs with only one activating ligand, it breaks down when trying to explain how EGFR family RTKs signal differently in response to their multiple cognate ligands (Sweeney and Carraway, 2000; Wilson et al., 2009).

EGFR is activated by seven different growth factors (Harris et al., 2003), which fall into two groups based on receptor-binding affinity. The high-affinity ligands are EGF, transforming growth factor- $\alpha$  (TGF $\alpha$ ), betacellulin (BTC), and heparin binding EGF-like growth factor (HB-EGF), which bind cell surface EGFR with apparent  $K_{ds}$  of 0.1-1 nM. The low-affinity ligands are epiregulin (EREG), epigen (EPGN), and amphiregulin (AREG), which bind 10-100 fold more weakly. Numerous studies report distinct EGFR-dependent cellular responses to the different ligands (Wilson et al., 2009), with a given cell line responding differently to individual EGFR ligands in terms of cell proliferation (Wilson et al., 2012), differentiation (Kochupurakkal et al., 2005; Rizzi et al., 2013), and/or motility (Willmarth and Ethier, 2006). Individual EGFR ligands also induce qualitatively and quantitatively

different downstream signals (Knudsen et al., 2014; Ronan et al., 2016; Wilson et al., 2012), and are linked to unique phenotypes *in vivo* (Wilson et al., 2009).

It remains unclear from current mechanistic understanding how different ligands could promote distinct cellular signaling responses through the same RTK. Crystal structures have described how EGF or TGF $\alpha$  induce formation of activated EGFR dimers (Ferguson et al., 2003; Garrett et al., 2002; Kovacs et al., 2015; Ogiso et al., 2002). Our more recent work (Bessman et al., 2014), however, has argued that relationships between extracellular ligand binding and receptor dimerization are more complex than suggested by these models – allowing the possibility that different EGFR ligands induce distinct dimers (Wilson et al., 2009). Here, we describe crystallographic and cellular studies that reveal how two EGFR ligands, epiregulin and epigen, do indeed drive the EGFR extracellular region into dimers with different structures. We show that the resulting ligand-induced dimers are weaker and more short-lived than those induced by EGF, altering signaling kinetics in a way that profoundly influences cellular outcome. Our findings suggest unexpected parallels with the biased agonism observed for some G protein-coupled receptor (GPCR) ligands (Lane et al., 2017; Wacker et al., 2017), and further suggest that kinetic proofreading might play an important role in controlling ligand discrimination and RTK signaling specificity (Swain and Siggia, 2002).

## Results and Discussion

### Epiregulin stabilizes a unique EGFR extracellular dimer

To ask whether individual EGFR ligands can stabilize different dimeric configurations of the EGFR extracellular region (sEGFR), we first determined the 2.9 Å structure (Figure 1A) of a complex between epiregulin (EREG) and sEGFR501 (EREG/sEGFR501). The sEGFR501 construct was used for the earliest structural studies of EGFR (Garrett et al., 2002), and contains both  $\beta$ -helix/solenoid ligand-binding domains of the receptor (domains I and III) plus the complete cysteine-rich laminin-like domain II that includes the key dimerization arm. Most of domain IV (residues 482-618), which contributes very little to dimerization (Dawson et al., 2005), is absent from sEGFR501.

As in every ErbB receptor dimer (Ferguson, 2008), dimerization of the EREG/sEGFR501 complex is mediated by domain II, with the dimerization arm at the center of the interface. Remarkably, however, as seen by comparing Figures 1A and B, the relationship between the two sEGFR501 protomers (grey and green) in the EREG/sEGFR501 dimer is strikingly different from that seen for TGF $\alpha$ /sEGFR501 (Garrett et al., 2002). The EREG/sEGFR501 dimer (Figure 1A) displays a distinct asymmetry that contrasts with the two-fold symmetry seen in TGF $\alpha$ /sEGFR501 (Figure 1B) and EGF/sEGFR dimers (Garrett et al., 2002; Lu et al., 2010; Ogiso et al., 2002). As cartooned in the lower parts of Figures 1A and B, this asymmetry arises because domain II in the right-hand (green) epiregulin-bound sEGFR molecule has failed to undergo the ‘bend’ usually observed upon wedging of ligand between domains I and III. As a result, the ‘unbent’ domain II in the right-hand (green) receptor molecule of the EREG/sEGFR501 dimer projects beyond its (grey) neighbor at the top of the dimer interface by  $\sim 7$  Å (Figure 1C).

Importantly, the fact that the EREG/sEGFR501 dimer differs in structure from other sEGFR dimers confirms the proposal that individual EGFR ligands can stabilize distinct EGFR conformations (Wilson et al., 2009), which might underlie differential signaling. It is appealing to hypothesize that the distinct conformation of epiregulin-activated EGFR allows it to engage a unique subset of downstream effectors – which would make this ligand a biased agonist (Lane et al., 2017). Alternatively, epiregulin might be a partial agonist that simply signals less strongly (but through the same effectors) at a given degree of receptor occupancy – as suggested in studies of other EGFR ligands (Macdonald-Obermann and Pike, 2014).

### **EREG/sEGFR501 dimer asymmetry mimics that seen in *D. melanogaster* EGFR**

The key structural differences responsible for asymmetry of the EREG/sEGFR501 dimer are clearest in the upper part of the dimer interface, detailed in Figure 1C. The bend in domain II of the left-hand (grey) molecule allows its N-terminal region to contact its unbent counterpart in the right-hand (green) molecule. The result is an intimate interface between the two sEGFR501 molecules in the upper part of Figure 1C, burying 641 Å<sup>2</sup> of surface. The equivalent interface fails to form in the symmetric TGF $\alpha$ /sEGFR501 dimer (Figure 1D), with only 213 Å<sup>2</sup> buried (Garrett et al., 2002). Domain II residues engaged directly in the asymmetric epiregulin-induced sEGFR501 dimer interface (Q194, S196, P204, H209, P219, E221, and D238: bold in Figure 1C) do not participate in the interface of TGF $\beta$ - or EGF-induced dimers (Figure 1D). Intriguingly, however, these residues are conserved in *Drosophila* EGFR (dEGFR), where they make almost identical sets of interactions (Figure S1) in the asymmetric dimer of the dEGFR extracellular region induced by its ligand Spitz (Alvarado et al., 2010). Formation of the asymmetric dimer shown in Figures 1A and S1 is therefore an evolutionarily conserved property of EGFR.

### **Dimerization arm-mediated contacts are compromised in the EREG/sEGFR501 dimer**

The distinct sEGFR dimer structure induced by epiregulin led us to ask how such relatively small structural differences in the extracellular region might be propagated across the membrane to alter signaling. Studies using chemical biology tools have suggested that different EGFR ligands can stabilize distinct intracellular structures (Doerner et al., 2015; Scheck et al., 2012). Most other reports, by contrast, argue for loose or flexible linkage between extra- and intracellular regions of EGFR (Lu et al., 2010; Lu et al., 2012; Mi et al., 2011) – making straightforward conformational coupling mechanisms difficult to envision. We therefore also considered the possibility that EGFR dimers induced by different ligands might differ in their stability and/or lifetime. The increased buried surface in the N-terminal part of the domain II dimer interface (Figure 1C) initially suggested that EREG/sEGFR501 dimers might be stronger than TGF $\beta$ /sEGFR501 dimers. At the same time, however, key intermolecular contacts involving the grey dimerization arm (Figure 1C, far right and bottom panel) are substantially compromised by the asymmetry of the EREG/sEGFR501 dimer. As a result, the surface area buried in the central dimerization arm region of EREG/sEGFR501 dimers (1350 Å<sup>2</sup>) is ~30% lower than in TGF $\beta$ /sEGFR501 dimers (Garrett et al., 2002), and 40% lower than in EGF/sEGFR dimers (Lu et al., 2010). This reduction results primarily from loss of a key interaction between Y251 in the dimerization arm of the grey molecule and R285 in disulfide-bonded module m6 of its green dimerization partner (Figure 1C,D). In

TGF $\alpha$ /sEGFR501 and EGF/sEGFR dimers, Y251 and R285 make key intermolecular cation- $\pi$  interactions on both sides of the symmetric dimer (dotted lines in Figure 1D). Disrupting these by mutation abolishes ligand-induced receptor activation (Ogiso et al., 2002). The EREG/sEGFR501 dimer only retains this interaction on the left side of the dimer (Figure 1C). On the right side, by contrast, Y251 and R285 are separated by  $>12$  Å (Figure 1C, lower panel) because the grey dimerization arm is displaced – failing to dock into its canonical site (involving modules m4, m5, and m6) and instead making only limited interactions with modules m6 and m7. EREG/sEGFR501 dimer asymmetry also distorts the ‘buttressed’ module m6 interactions involving D279 and H280 (asterisk in Figure 1D), which are crucial for EGFR dimerization (Dawson et al., 2005). Thus, despite additional interactions in the N-terminal part of the EREG/sEGFR501 dimer interface, the distorted dimerization arm interactions suggest that the EREG-induced sEGFR dimer may actually be weaker than TGF $\alpha$ -or EGF-induced dimers. Reduced dimerization strength could be just as important for altered signaling as any specific structural differences, perhaps weakening signaling by the occupied receptor (epiregulin behaving as a partial agonist).

### Epigen-bound sEGFR501 crystallizes as a monomer

A 3.0 Å resolution crystal structure of epigen (EPGN) bound to sEGFR501 (EPGN/sEGFR501) provides further weight to the argument that certain EGFR ligands induce weakened receptor dimers. Unexpectedly, the crystallized EPGN/sEGFR501 complex was monomeric (Figure 2A). The most extensive receptor/receptor interface in the EPGN/sEGFR501 crystals involves only limited domain III contacts, forming a dimer that is sterically infeasible in the membrane. The absence of domain II-mediated dimers in EPGN/sEGFR501 crystals cannot reflect incomplete ligand saturation, since epigen binds sEGFR501 with  $K_d = 2.8$   $\mu$ M (Figure S2A) yet was present at 280  $\mu$ M during crystallization (and  $\sim 13$  mM in crystals). Moreover, clear electron density for epigen was seen in the ligand-binding site (Figure S3A). As in all other ligand-bound ErbB receptor extracellular regions, epigen bridges domains I and III, which adopt positions characteristic of the ‘untethered’ receptor conformation (Ferguson, 2008). Whereas all other ligand/sErbB complexes have crystallized as domain II-mediated dimers, the EPGN/sEGFR501 complex structure instead closely resembles the monomeric ErbB2 extracellular region (Figure 2B), which has no known ligand and does not homodimerize (Ferguson et al., 2000; Garrett et al., 2003).

In addition to resembling ErbB2, the epigen-bound sEGFR501 structure (red in Figures 2 and 3A) overlays remarkably well (C $\alpha$  rmsd of 1.1 Å) with the right-hand molecule in the EREG/sEGFR501 complex (green in Figure 3A). By contrast, EPGN/sEGFR501 overlays poorly (C $\alpha$  rmsd  $>3$  Å) with the left-hand (grey) molecule in the epiregulin-bound dimer (Figure 3B) because of significant domain rearrangements. As suggested by the similarity apparent in Figure 2, epigen-bound sEGFR501 also overlays well with unliganded extracellular regions from ErbB2 (blue in Figure 3C) and the *Drosophila* EGFR (orange/brown in Figure 3C), which both likewise homodimerize weakly (or not at all) in solution (Alvarado et al., 2009; Ferguson et al., 2000). These extracellular regions all have an unbent, or straight, domain II (denoted with a dashed line in Figure 3C) – common to all unactivated ErbB receptors (Ferguson, 2008) – that apparently precludes (or weakens) canonical

homodimerization. Indeed, the absence of a bend in domain II makes it impossible to model EPGN/sEGFR501 homodimers without either severe steric clashes at the interface or minimal contact surface (Figure S2B) – regardless of how hypothetical dimers are constructed. Strongly dimerizing ErbB receptor extracellular regions, by contrast, all have a characteristic bend in domain II, as depicted in Figure 3D for the TGF $\alpha$ /sEGFR501 and neuregulin-1 $\beta$ (NRG1 $\beta$ )/sErbB4 complexes as well as the left-hand (grey) molecule in the EREG/sEGFR501 complex. Structural details of ligand binding by sEGFR501 are summarized in Figures S3C and D, and share strong similarities with those seen in other liganded sEGFR structures.

### Epiregulin and epigen fail to induce strong sEGFR dimerization in solution

We next used small-angle X-ray scattering (SAXS) to compare strengths of ligand-induced sEGFR501 dimers in solution. As in previous studies (Lemmon et al., 1997), we monitored  $I(0)/c$  (normalized scattering intensity extrapolated to zero angle) to measure weight-averaged molecular mass and thus ligand-induced dimerization. Saturating sEGFR501 (at 70  $\mu$ M) with EGF elevates  $I(0)/c$  by two-fold (Figure 4A) through stoichiometric EGF-induced sEGFR501 dimerization (Lemmon et al., 1997). The same increase is seen for the other high-affinity EGFR ligands, TGF $\alpha$ , HB-EGF and BTC (Figures 4A and S4). By contrast,  $I(0)/c$  increases only  $\sim$ 1.3-fold when saturating epiregulin or epigen is added, which is the result expected when ligand binds without dimerization – as seen when EGF is added to ‘dimarm\*’ sEGFR501, a well-studied variant with six mutations that abolish dimerization (Dawson et al., 2005; Garrett et al., 2002; Valley et al., 2015). Failure to detect dimerization in this assay (with 70 mM sEGFR501) places a lower limit of  $\sim$ 200  $\mu$ M on  $K_d$ , compared with  $\sim$ 3  $\mu$ M for TGF $\alpha$ /sEGFR501 (Dawson et al., 2005) – arguing that sEGFR501 dimers induced by epiregulin or epigen are  $>$ 60-fold weaker than those induced by TGF $\alpha$ . Importantly, despite being a low-affinity EGFR ligand like epiregulin and epigen, amphiregulin promotes substantial sEGFR501 dimerization (red in Figures 4A and S4F) – demonstrating that the observed dimerization differences do not simply reflect reduced ligand/receptor affinity.

### Epiregulin and epigen promote activation of EGFR in cells

The weak homodimerization of EREG/sEGFR501 and EPGN/sEGFR501 complexes prompted us next to ask whether these complexes might preferentially heterodimerize with other ErbB receptors. Efforts to detect epiregulin- or epigen-induced sEGFR heterodimers with sErbB2 or sErbB3 using SAXS (with or without the ErbB3 ligand NRG1 $\beta$ ), however, gave the same negative results that we reported for EGF (Ferguson et al., 2000). Moreover, studies in SKBR3 and T47D breast cancer cells revealed no enhancement of ErbB2 activation by epiregulin or epigen compared with that seen for EGF.

With selective heterodimerization ruled out, it was important to determine whether epiregulin or epigen can induce sufficient homodimerization of intact EGFR in cell membranes to support signaling. We expressed human EGFR in *Drosophila* S2 cells as a null background in which other human ErbB receptors cannot contribute to activation (since they are absent). Both epiregulin and epigen activate EGFR robustly in these assays (Figure 4B), reaching EGFR autophosphorylation levels similar to those seen with saturating EGF.

Quantitating these data yields EC<sub>50</sub> values (Figure 4C) that reflect the ~100-fold lower affinities of epiregulin and epigen for EGFR (Figure S2A). Three important conclusions can be drawn from Figures 4B and C. First, both ligands can activate EGFR in cells without the need for other ErbB receptors. Second, dimerization arm (dimarm\*) mutations that abolish activation by EGF (Valley et al., 2015) also eliminate EGFR activation by epiregulin or epigen (Figure 4B), arguing that these ligands rely on the same EGFR dimerization interface. Third, despite the fact that both ligands bind EGFR ~100-fold more weakly than EGF, and induce substantially weaker sEGFR dimers, epiregulin and epigen appear to be full agonists of receptor phosphorylation – arguing that their distinct signaling properties do not simply reflect altered potency.

### Epiregulin and epigen are partial agonists of EGFR dimerization in cells

We next asked how reduced EREG/sEGFR501 and EPGN/sEGFR501 dimer strength is manifest in the dimerization properties of intact cell surface EGFR. We first used a FRET approach with the extracellular region (ECR) plus transmembrane (TM) domain of EGFR fused to intracellular fluorescent proteins (FPs). We measured FRET as a function of receptor density (Chen et al., 2010a) in individual membrane vesicles generated from CHO cells (Del Piccolo et al., 2012) that lack endogenous EGFR, with EGFR<sub>ECR-TM-eYFP</sub> as FRET donor and EGFR<sub>ECR-TM-mCherry</sub> as acceptor (Figures 5A and S5A,B). Adding EGF, epiregulin (Figure 5B), or epigen (Figure 5C) substantially increased FRET compared with that seen without ligand (open grey circles), confirming that all three ligands induce EGFR oligomerization in a membrane context. EGF-bound receptor was maximally dimeric at all receptor densities (black circles in Figure 5B and C). For epiregulin or epigen, by contrast, FRET was clearly dependent on receptor density – consistent with weaker dimerization. Fitting the unbinned data (Figure S5B) to equations describing monomer/dimer equilibria suggested that EREG/EGFR<sub>ECR-TM-FP</sub> complexes dimerize 10-50 fold more strongly than EPGN/EGFR<sub>ECR-TM-FP</sub> complexes (Table S2) – consistent with crystallization of the EREG/sEGFR501 complex (but not the EPGN/sEGFR501 complex) as a dimer. The receptor density range below which epiregulin or epigen induce less FRET than EGF (~500 receptors per  $\mu\text{m}^2$ ) in Figures 5B and C corresponds to a local EGFR concentration of ~80  $\mu\text{M}$ , or ~200,000 receptors/cell. Most responsive EGFR-expressing cells have substantially lower numbers of receptors than this (Shi et al., 2016), suggesting that the dimerization affinity differences we observe are physiologically relevant.

Cognizant of the fact that dimerization of intact EGFR also involves significant contributions from the intracellular kinase domain (Kovacs et al., 2015) and juxtamembrane domain (Red Brewer et al., 2009), it was important to extend these studies to full-length EGFR. We labeled N-terminally HA-tagged EGFR, stably expressed in CHO cells, with quantum dots (QD605 and QD655) for two-color single particle tracking (SPT) studies as previously described (Low-Nam et al., 2011; Valley et al., 2015). Analyzing trajectories from multiple single quantum dots (see Figure S5C for examples) gave the ensemble mean square displacement (MSD) plots shown in Figure 5D, from which diffusion coefficients ( $D$ ) were calculated (see inset for 95% confidence intervals). Adding saturating EGF reduced the mean  $D$  value by 2.5-fold (from 0.036 to 0.014  $\mu\text{m}^2\text{s}^{-1}$ ), consistent with previous reports (Chung et al., 2010; Valley et al., 2015). Saturating the receptor with epiregulin or epigen (at

20  $\mu\text{M}$ ) also reduced EGFR mobility, but to a smaller extent – reducing  $D$  by just under 1.9-fold. As shown in Figure 5E, the distribution of  $D$  values for individual CHO cells treated with epiregulin or epigen was significantly different from the distribution seen with EGF. Similar trends were also observed in HeLa cells. We interpret this ligand-induced reduction in EGFR mobility as a manifestation of both receptor dimerization and activation-dependent signaling complex assembly. Although the relationship between diffusion and dimerization is not straightforward (Low-Nam et al., 2011), it seems reasonable to argue that the smaller effects of epiregulin and epigen than EGF on EGFR mobility reflect weaker dimerization of the full-length receptor bound to these ligands. We also attempted to capture epiregulin- or epigen-induced dimers using two-color tracking to determine dimer lifetimes (off-rates) and compare with those reported for EGF-induced dimers (Low-Nam et al., 2011). Unfortunately, despite much effort, we were not able to observe a sufficiently large number of epiregulin- or epigen-induced dimers before the saturating ligand concentrations induced receptor endocytosis. Qualitatively, however, the data (and indeed the difficulty in detecting long-lived dimers) suggest shorter dimer lifetimes for EGFR activated by epiregulin or epigen than by EGF – as expected if dimerization contacts are impaired. It should be noted that no reduction in  $D$  can be detected in equivalent studies of intact dimarm\* EGFR when saturated with EGF (Valley et al., 2015).

Taken together, the data in Figure 5 argue that the impaired dimerization of EREG/sEGFR501 and EPGN/sEGFR501 complexes – in crystals and in solution – accurately reflects the behavior of the intact receptor when activated by these ligands in cells. Since the (receptor-mediated) dimerization induced by epiregulin or epigen is clearly below the maximal dimerization capability of EGFR – even when saturating the receptor – these ligands can therefore be said to behave as partial agonists with respect to EGFR dimerization, despite the fact that they appear to function as full agonists with respect to receptor phosphorylation.

### Reduced dimerization alters receptor activation kinetics

The fact that epiregulin and epigen appear to function as partial agonists of dimerization, yet as full agonists of receptor phosphorylation, led us to hypothesize that altered dimer strengths (rather than structural changes *per se*) might underlie ligand discrimination by EGFR – particularly if coupled to modified signaling kinetics. A long history of studies of Erk activation kinetics downstream of RTKs has shown how signaling outcome can be profoundly different for transient versus sustained signaling (Marshall, 1995). Recent studies have further suggested that receptor-level mechanisms play an important role in defining these kinetics (Sparta et al., 2015). Since epiregulin and epigen induce weaker EGFR dimers than EGF, we hypothesized that they might also promote more short-lived activation of the receptor than EGF. Remarkably, however, we found exactly the opposite (Figure 6). We performed these studies in MCF-7 breast adenocarcinoma cells, which express only a few thousand EGFR molecules per cell (Shi et al., 2016), so that signaling differences arising from reduced receptor dimerization strength would be most evident. As shown in Figure 6A, tyrosine autophosphorylation of EGFR was substantially more sustained following activation with epiregulin or epigen than with EGF. Whereas EGF-induced EGFR phosphorylation at Y1173, Y845, or Y1086 returns to baseline within 20-45 minutes of



initial stimulation, it remains elevated even 90-120 minutes after initial stimulation with epiregulin or epigen, and does not fall below 50% of the peak level after one hour (Figure 6B). We also observed the same behavior in T47D cells, which express significantly higher EGFR levels (~50,000 per cell) than MCF-7 cells (Figure S6A). Moreover, the sustained nature of EGFR activation was retained when epiregulin and epigen were added at concentrations (1  $\mu$ M or 100 nM) substantially below saturation (Figure S6B), arguing that it is not simply a consequence of high ligand concentrations. Further, analysis of the concentration dependence of EGFR phosphorylation at 5 minutes and 1 hour gave similar results at each time point for epiregulin and epigen (Figure S6C). The sustained nature of the autophosphorylation signal at any ligand concentration suggests that it is indeed a feature of the EREG/EGFR or EPGN/EGFR complex rather than a consequence of low ligand affinity or receptor occupancy. Importantly, we also showed that amphiregulin promotes transient EGFR autophosphorylation (Figure S6D) despite binding the receptor with a similar (or lower) affinity than epiregulin or epigen (Ronan et al., 2016; Wilson et al., 2012). This is consistent with the ability of amphiregulin to induce sEGFR501 dimerization (Figure 4A), and further argues that reduced dimerization and sustained receptor activation are not simply consequences of low ligand/receptor affinity.

### **Sustained EGFR activation by epiregulin and epigen alters cellular responses**

Sustained signaling of the sort seen in Figure 6 – if propagated to the level of Erk – is typically associated with cell differentiation, whereas transient responses tend to be linked with proliferation (Lemmon et al., 2016; Marshall, 1995). Consistent with this idea, several reports have described induction of differentiation by epiregulin or epigen in settings where EGF instead induces proliferation. Epiregulin induces differentiation of SK-N-BE neuroblastoma cells (Rizzi et al., 2013), for example, and has been reported to regulate differentiation of several tissue types (Riese and Cullum, 2014). Similarly, epigen promotes differentiation of breast cancer cells, prostate epithelial cells, PC12 cells, and endothelial cells in settings where EGF or TGF $\alpha$  do not (Kochupurakkal et al., 2005). Prompted by these observations we assessed the ability of epiregulin and epigen to induce differentiation of MCF-7 cells, monitoring intracellular lipid droplet accumulation by Oil Red O staining. As shown in Figure 7A, both epiregulin and epigen promote substantial intracellular cytoplasmic lipid droplet accumulation. This is true at saturating (10  $\mu$ M) or sub-saturating (100 nM) concentrations of these ligands, whereas saturating EGF induces no such response. In agreement with previous studies (Herrero et al., 2016; Nagashima et al., 2007), the ErbB3/ErbB4 ligand NRG1 $\beta$  also promotes MCF-7 cell differentiation (Figure 7A). The distinct biological response to epiregulin and epigen (compared to EGF) correlates with both sustained EGFR activation (Figure 6) and sustained Erk activation (Figure 7B). By contrast, EGF induces transient Erk activation (as expected) – as do all other EGFR ligands, including the low-affinity ligand amphiregulin (Figure 7C) – and fails to induce cell differentiation. Thus, EGFR ligands that promote weak receptor dimerization appear, somewhat counter-intuitively, to induce more sustained signaling responses at the level of both EGFR and Erk activation.

## Sustained signaling responses to epigen are mediated exclusively by EGFR

Before ascribing the sustained (differentiation-inducing) Erk signaling by epiregulin and epigen shown in Figure 7 solely to altered EGFR activation, it is important to ask whether these ligands can bind other ErbB receptors. Epigen has been reported to bind EGFR exclusively (Wilson et al., 2009), whereas epiregulin can activate ErbB4 as well as EGFR (Komurasaki et al., 1997). We confirmed this specificity using competition surface plasmon resonance (SPR) studies (Figure S7A-C) and signaling studies in Ba/F3 cells (Figure S7D), leading us to conclude that the sustained EGFR and Erk signaling seen for epigen (compared with EGF) is independent of any effects on other ErbB family receptors, and reflects the ability of EGFR alone to discriminate between epigen and EGF.

## Differences in EGFR-dependent responses to epiregulin and epigen

Since epiregulin (but not epigen) binds detectably to sErbB4 (Figure S7C) and activates ErbB4 expressed in Ba/F3 cells (Figure S7D), we next asked whether the sustained signaling responses to epiregulin in MCF-7 cells require ErbB4 – by knocking down ErbB4 using siRNA. As shown in Figure S7E, epiregulin-induced EGFR and Erk phosphorylation remain sustained following knock down of ErbB4 to levels barely detectable by Western blotting. This finding suggests – as with epigen – that sustained epiregulin signaling seen in Figures 6 and 7B reflects differential activation of EGFR rather than an ErbB4-dependent effect. Moreover, it is important to note that both BTC and HB-EGF activate ErbB4 as well as EGFR (Elenius et al., 1997; Riese et al., 1996), yet these ligands promote transient Erk activation (Figure 7C). As expected, ErbB4 knockdown had no effect on EGFR or Erk activation by epigen (Figure S7F), but we did notice in these experiments that epiregulin promotes more robust ErbB3 phosphorylation than epigen or EGF. Importantly, this difference is also propagated into differential Akt signaling for the three different ligands (Figure 7D). Epiregulin promotes sustained Akt activation (lasting >90 minutes), EGF promotes transient Akt activation, and the response to epigen has intermediate kinetics.

## Conclusions

Our studies reveal that EGFR dimers induced by epiregulin and epigen differ structurally from those induced by EGF or TGF $\alpha$ . If the observed conformational differences could be communicated to the intracellular region of the receptor as suggested (Doerner et al., 2015), this might contribute to biased agonism. In addition (or instead), however, our results suggest that altered structural dynamics of the activated receptor dimer help define signaling specificity. This possibility was recently explored for GPCRs (Manglik et al., 2015; Nygaard et al., 2013), and our results argue that the time evolution of receptor activation (rather than receptor structure per se) may likewise be one of the key parameters altered when EGFR is stimulated with epiregulin or epigen rather than EGF. The conformational ensemble explored by the activated receptor might be different for each ligand, which could in turn alter the way in which (and for how long) the receptor couples to downstream effectors, as recently suggested for biased agonism of GPCRs (Furness et al., 2016; Klein Herenbrink et al., 2016; Lane et al., 2017; Wacker et al., 2017).

In terms of EGFR signaling, it seems paradoxical at first thought that reduced EGFR dimer strength and lifetime seen with epiregulin and epigen causes more sustained (rather than more transient) activation of EGFR – leading to sustained Erk activation and induction of MCF-7 cell differentiation rather than proliferation. Thus, partial agonism at the level of receptor dimerization (but not receptor phosphorylation) is manifest as apparent biased agonism with respect to cell fate. How might more sustained EGFR autophosphorylation (and Erk activation) arise from weaker, shorter-lived, EGFR dimers? One possibility is a form of kinetic proofreading (McKeithan, 1995), as has been discussed extensively for T cell receptor signaling. RTKs such as EGFR require multiple phosphorylation events for complete activation and, as pointed out previously (Swain and Siggia, 2002), it is easy to see how short-lived dimers might not reach the end of a progressive multi-site phosphorylation ‘program’ – and might thus fail to elicit all possible signals. If the short-lived EGFR dimers induced by epiregulin or epigen do not complete the full complement of phosphorylation events and therefore fail to engage a key (late) negative feedback signal, this would result in loss of the transient nature of the signal seen with EGF (Lemmon et al., 2016). Loss of such a negative feedback event would promote the more sustained signaling seen in our steady state studies.

Putative negative feedback(s) normally responsible for the transient nature of EGF-induced EGFR activation, but lost or impaired with epiregulin and epigen, might include receptor internalization and receptor dephosphorylation. All EGFR agonists induce receptor internalization, but alterations in receptor ubiquitinylation, recycling, or degradation of endocytosed receptors may be important. Indeed, differences in all of these activities have been reported across EGFR ligands (Francavilla et al., 2016; Roepstorff et al., 2009). The issue is not simply one of receptor recycling versus degradation, however; EGFR activated by either TGF $\alpha$  or epiregulin is internalized and recycled (Roepstorff et al., 2009), yet TGF $\alpha$  resembles EGF rather than epiregulin in its signaling kinetics (Figure 7C). Reduced recruitment of one or more tyrosine phosphatases following receptor activation by epiregulin or epigen is another important possibility. Indeed, depleting PTP1B by siRNA, for example, has been shown to cause normally transient EGF-induced EGFR phosphorylation to become sustained (Eden et al., 2010). Along similar lines, feedback that modulates the ability of adaptor molecules such as Shc to protect phosphotyrosines in EGFR from dephosphorylation could determine the extent to which signals are transient or sustained (Kholodenko et al., 1999).

Other recent studies have also described functional selectivity or biased signaling through receptors with a single TM domain. In one example, biased signaling by c-Kit was achieved by engineered stem cell factor variants with altered dimerization (Ho et al., 2017). In another case, erythropoietin (EPO) harboring a pathogenic mutation was found to signal aberrantly because of altered binding kinetics rather than strength (Kim et al., 2017). Our results show how biased agonism or functional selectivity can occur with EGFR and its natural human ligands. We explain the structural basis of this, but our findings argue that it is not the structure of the receptor dimer itself that alters the signaling outcome. Instead, the strength and lifetime of the activated receptor dimer – which result from the structural alterations – appear to play an important role. This has important implications for understanding how different RTKs (and other receptors) can be harnessed to promote diverse signaling

outcomes, and for considering how to modulate them therapeutically. For example, our results suggest the possibility of using well-defined biologics to modulate or bias aberrant EGFR signaling rather than simply blocking activity.

## Contact for Reagent and Resource Sharing

Requests for further information or reagents may be directed to the Lead Contact, Mark A. Lemmon (mark.lemmon@yale.edu).

## Experimental Model and Subject Details

### Cell culture

**Insect cells**—*D. melanogaster* Schneider 2 (S2) cells were maintained at 27°C in Schneider's Insect Medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum. S2 cells were originally derived from a primary culture of late stage *D. melanogaster* embryos, and are male. For protein expression, stably transfected S2 cell pools were grown at 27°C in ESF 921 Insect Cell Culture Medium (Expression Systems) or EX-CELL 420 Serum Free Medium (Sigma-Aldrich) in the presence of 200 µg/ml hygromycin-B. *Spodoptera frugiperda* Sf9 cells were propagated at 27°C in ESF 921 Insect Cell Culture Medium (Expression Systems). Sf9 cells were originally established from immature ovaries of female *S. frugiperda* pupae.

**Mammalian cells**—Mammalian cells were all grown in a humidified incubator with 5% CO<sub>2</sub>. MCF-7 breast cancer cells (ATCC HTB-22, first isolated in 1970 from a 69 year old woman) were cultured in complete DMEM/F-12 medium (ThermoFisher Scientific) containing 10% fetal bovine serum (FBS) (Atlanta Biologicals #S11550), 0.01 mg/ml human recombinant insulin (Santa Cruz) as well as 100 U/ml penicillin and 100 µg/ml streptomycin (ThermoFisher Scientific). T47D breast cancer cells (ATCC HTB-133, first established in the late 1970s from a 54 year old woman) were cultured in complete RPMI 1640 medium containing L-glutamine and HEPES (ThermoFisher Scientific), supplemented with 10% fetal bovine serum (FBS) (Atlanta Biologicals #S11550), 0.02 mg/ml human recombinant insulin (Santa Cruz) and 100 U/ml penicillin and 100 µg/ml streptomycin (ThermoFisher Scientific). Ba/F3 cells (ATCC HB-283, an IL-3 dependent murine male pro B cell line, precise origin unclear) were maintained in RPMI 1640 medium containing L-glutamine and HEPES (ThermoFisher Scientific), supplemented with 10% FBS, 1 mM sodium pyruvate (ThermoFisher Scientific), 1 ng/ml interleukin-3 and 100 U/ml penicillin and 100 µg/ml streptomycin (ThermoFisher Scientific). Chinese hamster ovary (CHO) cells (ATCC CCL-61, derived from ovary of *Cricetulus griseus* female adult in 1957) were maintained in DMEM/F12 medium supplemented with 10% FBS and 2 mM glutamine.

## Method Details

### Protein expression and purification

#### EGFR family extracellular regions

**Expression constructs:** Recombinant baculoviruses directing expression of histidine-tagged sEGFR501 and its variants were generated as described (Dawson et al., 2005), with cDNA encoding amino acids 1-501 of the EGFR extracellular region (plus 6 appended C-terminal histidines) subcloned between the *Bam* *HI* and *Xba* *I* sites of the Bac-to-Bac pFastBac™1 vector (ThermoFisher Scientific) using the native EGFR signal sequence. Equivalent constructs for the ErbB3 and ErbB4 extracellular regions (sErbB3(1-500) and sErbB4(1-497)) were also generated. A dimerization-deficient ‘dimarm\*’ sEGFR501 variant was generated by site-directed mutagenesis with the following substitutions: Y246E, N247A, T249D, Y251E, Q252A and M253D (Dawson et al., 2005; Garrett et al., 2002).

**Protein expression:** Protein expression was induced by infection with appropriate recombinant baculovirus of 4-8 liters of Sf9 cells in ESF 921 medium at a density of  $\sim 2 \times 10^6$  cells/ml. Conditioned medium was harvested 3-4 days post-infection, concentrated  $\sim 4$ -fold, and diafiltered against 4 volumes of 25 mM MES (pH 6.0), 150 mM NaCl (buffer A). Histidine-tagged protein was captured on a 2 ml Ni-NTA column (Qiagen) at 4°C. After extensive washing in buffer A containing 20 mM imidazole, sEGFR501, sErbB3(1-500) or sErbB4(1-497) protein was eluted using an imidazole gradient (100 to 300 mM) at pH 6.0. Proteins were buffer exchanged into 25 mM MES (pH 6.0), 50 mM NaCl (buffer S), loaded onto an SO<sup>3-</sup> cation exchange column, and eluted during an isocratic step at 240 mM NaCl (or 24 mS/cm). Fractions containing sErbB protein were pooled, concentrated, and further purified by size exclusion chromatography (Superose 6, GE Healthcare) in 10 mM HEPES (pH 8.0), 150 mM NaCl (buffer B). Protein purity was assessed by overloaded Coomassie-stained SDS-PAGE.

#### Production of EGFR ligands

**Expression constructs:** cDNA encoding the EGF domains from human epiregulin and epigen was amplified by PCR and subcloned into a modified pMT/BiP/V5-His A (ThermoFisher Scientific) vector used for expressing EGF domains in S2 cells as Spitz fusion proteins (Alvarado et al., 2010; Klein et al., 2008). The expressed protein has an N-terminal BiP signal sequence that is followed by a hexahistidine tag, residues 44-76 from *D. melanogaster* Spitz (numbered as in UniProt Q01083), a Factor Xa cleavage site (I-E-G-R) and then the relevant EGF domain. The subcloned epiregulin and epigen EGF domain fragments correspond to residues D56-K116 and E49-Y109, respectively, of the complete chains (including signal sequence) of UniProt entries O14944 (epiregulin) and Q6UW88 (epigen). The EGF domains present in the crystal structures are numbered according to the mature chains as designated in UniProt, with residues S2-V48 for epiregulin (as in PDB entries 5E8D and 1K36) and C38-T79 for epigen. The EGF plus heparin-binding domains of human amphiregulin were subcloned into an analogous modified pFastbac™1 vector for expression in the same way, using the amphiregulin fragment corresponding to S101-S190 of the complete UniProt chain (P15514), or residues S82-S171 of the mature protein.

**Protein expression:** For epiregulin and epigen expression, stably transfected *D. melanogaster* S2 cell pools were established using standard manufacturer protocols, and were grown to a density of  $\sim 4\text{--}6 \times 10^6$  cells/ml before induction with  $500 \mu\text{M}$   $\text{CuSO}_4$  for 4-5 days at  $27^\circ\text{C}$ . For amphiregulin expression, expression from Sf9 cells directed by recombinant baculovirus was achieved as described above. For all three ligands, culture medium was concentrated  $\sim 4$ -fold following expression and diafiltered against 4 volumes of buffer A. Histidine-tagged Spitz fusion proteins were then purified by Ni-NTA and cation exchange chromatography as described above for sEGFR501. Ligands were eluted from the  $\text{SO}_3^-$  column using a gradient from 50 mM to 1 M NaCl in buffer S, and protein-containing fractions were buffer exchanged into 10 mM HEPES, 100 mM NaCl, 2 mM  $\text{CaCl}_2$ , pH 7.0 on a Superose 12 size exclusion column (GE Healthcare). Proteins were concentrated to  $>1$  mg/ml and cleaved overnight at  $4^\circ\text{C}$  with  $20 \mu\text{g}$  Factor Xa (New England BioLabs) to remove the N-terminal Spitz-derived sequence. Uncleaved fusion protein was removed by incubation with Ni-NTA beads in buffer B plus 20 mM imidazole (to prevent non-specific binding of the ligands). EGF domains were finally purified using a Superdex Peptide size exclusion column (GE Healthcare) equilibrated in buffer B.

Carrier-free human EGF (EMD Millipore), TGF $\alpha$  (EMD Millipore), betacellulin (PeproTech), and HB-EGF (R&D Systems) were purchased and dissolved in 10 mM HEPES, 150 mM NaCl, pH 8.0. For mammalian cell stimulation experiments, human EGF was purchased from R&D Systems and reconstituted in sterile PBS buffer containing 1 mg/ml bovine serum albumin, and human NRG1 $\beta$  (R&D Systems) was reconstituted in sterile PBS buffer.

### X-ray crystallography

**Crystallization**—Crystals of sEGFR501 bound to epiregulin or epigen were obtained using the hanging-drop vapor diffusion method, mixing equal volumes of protein complex with reservoir solution and equilibrating over this reservoir at  $21^\circ\text{C}$ . For EREG/sEGFR501 ( $190 \mu\text{M}$  sEGFR501/ $230 \mu\text{M}$  epiregulin), reservoir solution contained 30 mM citric acid, 70 mM Bis-Tris propane, 16% PEG3350, pH 7.6. Crystals of up to  $\sim 0.20 \times 0.20 \times 0.05$  mm were flash frozen following brief exposure to a cryoprotectant of 30 mM citric acid, 70 mM Bis-Tris propane, 20% PEG8000, 15% glycerol, pH 7.6. For EPGN/sEGFR501 ( $210 \mu\text{M}$  sEGFR501/ $280 \mu\text{M}$  epigen), reservoir solution contained 100 mM magnesium formate and 15% PEG3350. Crystals of  $\sim 0.15 \times 0.15 \times 0.10$  mm were cryoprotected in a solution of 20% PEG8000 and 15% glycerol and flash frozen in liquid nitrogen.

**Data collection and structure determination**—Crystals of EREG/sEGFR501 diffracted to  $2.94 \text{ \AA}$  resolution at APS beamline 23ID-D, and belonged to space group  $\text{P2}_1$  (Table S1). Diffraction data were anisotropic by  $0.4 \text{ \AA}$  along the  $c^*$  axis. The asymmetric unit contained 51% solvent and four EREG/sEGFR501 molecules (i.e., two 2:2 EREG/sEGFR501 dimers), with each 2:2 dimer related by a polypeptide backbone root-mean-square deviation (rmsd) of  $1.4 \text{ \AA}$  after refinement. Crystals of EPGN/sEGFR501 diffracted to  $3.0 \text{ \AA}$  resolution at APS beamline 23ID-B, and belonged to space group  $\text{P4}_322$  (Table S1). The asymmetric unit contained 63% solvent and two EPGN/sEGFR501 complexes, related by a polypeptide backbone rmsd of  $2.3 \text{ \AA}$  after refinement.

Data were processed with HKL3000. Diffraction patterns for the EPGN/sEGFR501 crystals contained contributions from three lattices, which were identified and integrated separately using the program PROTEUM2 (Bruker), and then scaled together using the program PHENIX (Adams et al., 2010). Due to significant anisotropy in the data, ellipsoidal truncation and anisotropic scaling were also performed using the diffraction anisotropy server (<https://services.mbi.ucla.edu/anisotrope/>). Structures were solved by the method of molecular replacement (MR) using Phaser (CCP4, 1994), with two fragments of EGFR from PDB ID 3NJP (Lu et al., 2010) as search models: domain I plus the N-terminal portion of domain II (residues 1-239) and domain III (residues 310-479). The resulting MR maps showed clear electron density for ligand in each binding site. Cycles of model building using the program Coot (Emsley and Cowtan, 2004) were alternated with rounds of refinement in Refmac (CCP4, 1994) or PHENIX (Adams et al., 2010), employing composite omit maps generated with PHENIX. TLS refinement was employed in later stages (as implemented in PHENIX), with anisotropic motion tensors refined for each of the receptor domains and ligand molecules. Final structures were refined using PHENIX and validated with the MolProbity and wwPDB servers.

**Structure representation**—Structural figures were made using PyMol. Structural overlays in Figure 3 employed the dimer formed between chains A and D in the EREG/sEGFR501 structure, and chain A in the EPGN/sEGFR501 structure. Overlays were similar when other molecules in the asymmetric units were used.

### Small-Angle X-ray Scattering (SAXS)

SAXS data were recorded at 4°C on a Rigaku PSAXS S-Max3000 pinhole camera system with a Rigaku 007HF rotating anode source and a Rigaku 300 mm wire grid ASM DTR 200 detector, with 20-80 min exposures. Protein concentration was 4 mg/ml (70 μM) in 10 mM HEPES, 150 mM NaCl, pH 8.0. Ligands were added at a 1.2-fold molar excess (84 μM), such that >90% ligand saturation of sEGFR501 was reached in each case. Data were reduced using the program SAXSGui v2.05.02 (Rigaku America & JJ X-Ray Systems ApS, Lyngby, Denmark) and matching buffers were subtracted using PRIMUS (Konarev et al., 2003) to yield the final scattering profile in which intensity ( $I$ ) is plotted as a function of  $q$  ( $4\pi\sin\theta/\lambda$ , where  $2\theta$  is the scattering angle). All samples were monodisperse as evidenced by linear Guinier regions (Figure S4). Values for the scattering intensity extrapolated to zero angle,  $I(0)$ , were calculated from the Guinier region where  $q^*R_g < 1.4$ , normalized by mass concentration of receptor protein, and then divided by the value of  $I(0)/c$  measured for unliganded (monomeric) sEGFR501 collected on the same day to give a value for fold-change in oligomeric state (Lemmon et al., 1997). Importantly, ligand and sEGFR501 concentrations used in all SAXS experiments are >25-fold above  $K_d$  for even the weakest ligands (Figure S2A), ensuring full saturation of all binding sites.

### Quantitative imaging FRET microscopy

**Constructs and protein expression**—The extracellular and transmembrane regions of human EGFR (residues 1-647 of the mature protein) were fused at the C-terminus to either mCherry or EYFP, connected via a GSGGSGGS flexible linker in pcDNA3.1+ (ThermoFisher Scientific) for mammalian cell expression of EGFR<sub>ECR</sub>-TM-mCherry and

EGFR<sub>ECCR-TM-eYFP</sub> fusions. The sequence for EYFP was modified using site directed mutagenesis to generate the EYFP-A206K dimerization-defective variant. For each imaging experiment, CHO cells were seeded in 35 mm dishes at a density of  $2 \times 10^4$  cells per well. After 24 hours, cells were transiently transfected with EGFR<sub>ECCR-TM-mCherry</sub> and EGFR<sub>ECCR-TM-eYFP</sub> expression constructs using Lipofectamine 3000 (ThermoFisher Scientific). 18-24 hours post-transfection, cells were washed twice with 30% PBS diluted in water (1 minute each), and then incubated with vesiculation buffer (200 mM NaCl, 100 mM bicine pH 8.5, 5 mM KCl, 0.75 mM CaCl<sub>2</sub>, 0.5 mM MgCl<sub>2</sub>). Vesiculation was allowed to proceed for 5-13 hours (Del Piccolo et al, 2012). The solution was then transferred to Nunc Lab-Tek II Chamber Slides (ThermoFisher Scientific), which were mounted onto a Nikon C1 laser scanning confocal microscope for imaging. All ligands added to the vesicles were allowed to equilibrate for at least one hour prior to imaging.

**FRET analysis**—Images were processed using a MATLAB (MathWorks) program developed in the laboratory of Kalina Hristova, which finds the boundary of each vesicle, verifies that the vesicle is present in donor, acceptor, and FRET scans. It then fits the intensity profile across the membrane to a Gaussian function. The approach used for calculating FRET efficiency and calibration were developed in previous studies (Chen et al., 2010a; Chen et al, 2010b; Kavran et al, 2014). Absolute protein concentration in the membrane is first calculated for donor and acceptor (Figure S5A) by comparing the fluorescence intensity in vesicles with the intensities measured from a dilution series of purified fluorescent protein standards (EYFP and mCherry). Bleed-through coefficients are calculated for each experiment (typically  $\sim 0.3$  and  $\sim 0.2$  for EYFP and mCherry, respectively). FRET data ( $E_{app}$ ), as shown in Figure S5B, are then corrected for contributions resulting from stochastic interactions that occur between proteins in the membrane (referred to as ‘proximity FRET’ or  $E_{proximity}$ ) using the model described (Wolber and Hudson, 1979), so that  $E_{dimerization} = E_{app} - E_{proximity}$ . The processed data (95-274 points) are subsequently fit to a monomer-dimer equilibrium model using GraphPad Prism (Table S2). Two parameters are fit:  $E$  (FRET efficiency within a dimer), where:

$$f_{dim} \times E = \frac{E_{app} - E_{prox}}{f_a}$$

and  $K_d$  (dissociation constant), where:

$$X_{dim} = \frac{1}{8} (K_d + 4X_{tot} - \sqrt{K_d(K_d + 8X_{tot})})$$

$X_{dim}$  is the concentration of dimer molecules,  $X_{tot}$  is total concentration of receptor, and  $f_a$  is the fraction of acceptor molecules in each vesicle.

**Data binning, error analysis, and fitting**—In order to obtain experimental averages and standard errors for the large number of data points in Figure S5B once processed, we need to account for the fact that the probability of having exactly the same EGFR<sub>ECCR-TM-FP</sub>



concentration in several vesicles is very low. Means and errors are therefore calculated using data with similar EGFR<sub>ECR-TM-FP</sub> concentrations that are binned in the  $x$  and  $y$  axes. For Figures 5B and 5C, the bin centers were spaced at intervals of 100 receptors per square micron, with the first bin centered at 150 receptors per square micron. When necessary, the range of the last bin was increased to ensure that the bin included at least 3 data points. Fits for each dataset described in Table S2 were calculated using the unbinned data. As described for other systems using this approach (Chen et al, 2010a), the data binning/averaging step does not affect data fitting, but facilitates visual comparison of the data and the best fit.

### Single-particle tracking studies

The approaches used for single-particle tracking of HA-tagged EGFR and diffusion analyses were established in previous studies (Low-Nam et al, 2011; Valley et al, 2015). The method involves conjugating anti-HA Fab fragments to quantum dots (QDs), by incubating an equimolar mixture of anti-HA-biotin (#12158167001, Roche) and QD605- or QD655-streptavidin (Q10101MP or Q10121MP, ThermoFisher Scientific) in PBS/1% BSA at 4°C for 2 h with agitation before imaging. CHO cell lines stably expressing HA-tagged EGFR (at approximately  $2.4 \times 10^5$  receptors per cell) were plated in Nunc Lab-Tek 8-well chambered coverglasses (#155411, ThermoFisher Scientific) and allowed to adhere overnight. Cells were imaged in Tyrode's imaging buffer (135 mM NaCl, 10 mM KCl, 0.4 mM MgCl<sub>2</sub>, 1 mM CaCl<sub>2</sub>, 10 mM HEPES, 20 mM glucose, 0.1% BSA, pH 7.2). Cells were incubated with 400 pM anti-HA-QD605/655 for 10 min at 37°C and unbound HA-QD was removed by extensive washing. Cells were imaged in the absence of ligand and for 1-8 min after adding saturating concentrations of the relevant ligand (50 nM EGF, 20 μM epiregulin, 20 μM epigen). Wide-field imaging was performed using an Olympus IX71 inverted microscope (60×, 1.2 NA water objective) with 1.6× extra magnification, and an objective heater (Bioptechs) to maintain cells at 34-36°C. A mercury lamp with a 436/10 nm band-pass excitation filter and a 50/50 neutral-density filter was used to excite the QDs. An OptoSplit image splitter (Cairn Research) with a 625 nm dichroic filter and the appropriate band-pass filters (655/40 nm and 605/20 nm, Chroma) was used to separate QD emission before detection by an EMCCD camera (Andor iXon 887), with a single pixel equivalent to 166.67 nm. Individual QD-labeled receptors were localized and tracked using methods developed previously (Valley et al., 2015). All image processing was performed using MATLAB together with the MATLAB toolbox for image-processing DIPImage (Delft University of Technology), and diffusion analysis of trajectories was conducted by fitting the first five points of the Mean Square Displacement curve;  $MSD = offset + 4D_{1-5} t$ .

### Cell signaling studies

**Signaling of human EGFR in S2 cells**—For analysis of ligand-induced EGFR activation in insect cells, *Drosophila* S2 cell pools stably expressing full-length wild-type or mutated EGFR were grown to mid-log phase and serum-starved overnight. Cells ( $4 \times 10^6$ ) were washed with ice-cold binding buffer (10 mM HEPES, 150 mM NaCl, 1% bovine serum albumin, pH 8.0, and stimulated with ligand (or left unstimulated) in this buffer for 10 min on ice. Cells were quickly lysed in binding buffer containing 1% NP-40, 1 mM phenylmethylsulfonyl fluoride, 1 μg/ml aprotinin, 1 μg/ml leupeptin, 5 mM sodium orthovanadate, plus Halt phosphatase inhibitor cocktail (ThermoFisher Scientific), and lysis

supernatants were subjected to immunoblotting with 3 µg/ml anti-EGFR ab-10 (Lab Vision) and 1:500 diluted anti-phosphotyrosine pY20 (Santa Cruz Biotechnology), with detection using LI-COR.

**Signaling by endogenous EGFR in breast cancer cells**—For mammalian cell studies, cells were starved for 6 to 16 hours, and then left unstimulated or stimulated with ligand for noted times at 37°C, using growth factor ligands present in starvation medium. After stimulation, medium was removed and cells were placed on ice and immediately lysed with scraping in ice-cold cell lysis buffer (Cell Signaling Technology) supplemented with PhosSTOP phosphatase inhibitor (Roche) and Complete protease inhibitor (Roche). Cell lysis supernatants were subjected to immunoblotting using the Xcell Surelock Electrophoresis system (ThermoFisher Scientific) and NuPAGE Novex 4-12% Bis-Tris Protein Gels (ThermoFisher Scientific) and the multistrip western blotting procedure (Aksamitiene et al., 2015). With the exception of Grb2 (C-23) from Santa Cruz Biotechnology, all antibodies were from Cell Signaling Technology, and were used at 1:1000 unless otherwise noted: phospho-p44/42 MAPK (Erk1/2) pT202/pY204 (E10), Akt pS473 (D9E), EGFR pY1173 (53A5 at 1:800), EGFR pY845 (2231), EGFR pY1086 (2220 at 1:500), EGFR pY1068 (2234 at 1:500), ErbB2 pY1221/pY1222 (6B12 at 1:700), ErbB3 pY1289 (D1B5 at 1:700), ErbB4 pY1284 (21A9 at 1:700), EGFR (D38B1), ErbB2 (D8F12 at 1:700), ErbB3 (D22C5 at 1:700), and ErbB4 (111B2 at 1:700). Secondary antibodies were either horse anti-mouse IgG horseradish peroxidase-linked antibody (Cell Signaling Technology) or goat anti-rabbit IgG (H+L) cross-adsorbed secondary antibody horseradish peroxidase conjugate (ThermoFisher Scientific), used at dilutions of 1:10,000 and 1:40,000 respectively. In cases where phosphorylation of EGFR Y1173 and Y845 were assessed using the same samples, pY845 blots were stripped with Restore Western Blot Stripping Buffer (ThermoFisher Scientific) and reprobed with pY1173 antibodies (in Figures 6A and S6D). In all cases, detection was by enhanced chemiluminescence using SuperSignal West Pico Chemiluminescent Substrate (ThermoFisher Scientific) and a Kodak Image Station 440CF (Kodak Scientific).

### Oil Red O staining

MCF-7 cells were seeded into 24-well plates containing poly-lysine coated glass coverslips at  $0.2 \times 10^6$  cells per well, and were allowed to grow for 24 hours. Cells were serum-starved for an additional 24 hours before stimulation with the indicated doses of EGFR ligands. Following incubation at 37°C in 5% CO<sub>2</sub> for 6 days, cells were fixed with 3.7% formaldehyde in PBS for 30 min. Cells were then washed with water, incubated in 60% isopropanol for 5 min, and stained with Oil Red O solution for 5 min. Oil Red O solution contained Oil Red O powder (Alfa Aesar) dissolved in isopropanol (300 mg per 100 ml), of which 3 parts were mixed with 2 parts of water and filtered. After staining, cells were rinsed with water and visualized and photographed using a ZEISS Axio Observer.A1 inverted microscope.

### Transient cell transfection with siRNA

Cells were harvested 30 min before transfection and resuspended in antibiotic-free complete medium. For each experiment,  $2 \times 10^6$  cells were aliquoted into Eppendorf tubes and

centrifuged at  $100 \times g$  for 10 min at room temperature. Supernatant was removed, and the cell pellet was resuspended in 100  $\mu$ l of Ingenio transfection solution (Mirus) containing 100 nM ErbB4 siRNA (OriGene). Control cells were transfected with 100 nM AllStars Negative Control siRNA (Qiagen). Cell suspensions containing siRNA were electroporated using the P-020 program for MCF-7 cells on a Nucleofector 2b device (Lonza). Immediately after electroporation, 0.5 ml of antibiotic-free complete medium, pre-equilibrated at 37°C in a CO<sub>2</sub> incubator, was added to the cuvette. The cell suspension was gently transferred into 60  $\times$  15mm CellBind surface dishes (Corning), and the final volume adjusted to 2 ml by addition of antibiotic-free complete medium. Cells were allowed to attach for 6 hours before addition of penicillin/streptomycin solution. The culture medium was aspirated the next day and replaced with 5 ml of fresh complete medium. Cell stimulation experiments were performed 72 hours post-transfection.

### Isothermal Titration Calorimetry (ITC)

All isothermal titration calorimetry (ITC) experiments were conducted on a MicroCal ITC200 instrument at 4°C. Purified receptor and ligand proteins were exchanged into 20 mM HEPES, pH 8.0, 150 mM NaCl and 3.4 mM EDTA by dialysis. [sEGFR501] in the calorimeter cell ranged from 6 to 26  $\mu$ M, and [ligand] in the injection syringe ranged from 150 to 315  $\mu$ M. Data were fit to a single-site binding model using the Origin software package (OriginLab) to derive  $H$ ,  $S$ , and  $K_D$  values. These values are reported as the mean  $\pm$  SD derived from at least three independent experiments. ITC plots shown in Figure S2 are representative of three independent experiments.

### Surface plasmon resonance (SPR)

SPR experiments to analyze binding of ErbB ligands (Ferguson et al, 2000) were performed on a Biacore 3000 instrument. EGF or NRG1 $\beta$  were immobilized on a CM5 sensorchip using amine coupling. Purified sEGFR501, sErbB3(1-500), or sErbB4(1-497) was injected at 5  $\mu$ l/min for 8 min (sufficient for binding to reach steady state) in degassed 10 mM HEPES (pH 8.0), 150 mM NaCl, 3 mM EDTA and 0.005% Surfactant P-20 at room temperature – either alone or pre-incubated with competitor ligand. Between injections, the sensorchip surface was regenerated using a 20  $\mu$ l injection of 10 mM sodium acetate (pH 5.0) containing 1 M NaCl. The final steady-state signal was background-corrected by subtracting the signal obtained with a control surface. For initial determination of receptor/ligand affinities, SPR signal values were plotted against [sErbB] and fit to a simple single-site saturation-binding model. Subsequent competition binding experiments in Figure S7 were set up with sErbB protein present at the measured  $K_d$  values for EGF or NRG1 $\beta$  binding. sErbB protein and competitor ligand (at the indicated concentrations) were pre-incubated for at least 30 minutes at room temperature prior to injection over the sensorchip.

## Quantification and Statistical Analysis

### Ligand-induced dimerization analysis by SAXS

To quantify ligand-induced dimerization, the mean  $I(0)/c$  value measured for unliganded, monomeric, sEGFR501 or sEGFR501-dimarm\* sample (collected on the same day) was set to a relative value of 1.0 (no dimerization), and all other  $I(0)/c$  values were normalized to

this value (to give fold dimerization). Mean  $\pm$  standard deviation of  $I(0)/c$  for each ligand was plotted in GraphPad Prism for 3-6 repeats of each experiment (8 for the EGF control), using protein from two different protein preparations (for independent biological replicates) for TGF $\alpha$ , HB-EGF, BTC and AREG, and three different ligand preparations for EREG and EPGN.

### **Analysis of EGFR activation in S2 cells**

Quantification of band intensities for phosphorylated and total EGFR was performed with LI-COR Image Studio software. The ratio of these signal intensities (calculated as phosphorylated EGFR divided by total EGFR) at each ligand concentration was determined, and the background value from the unstimulated sample subtracted. Data for at least three experimental repeats (using at least three different ligand preparations) were plotted as log[ligand] versus response using GraphPad Prism, from which the maximum response for each experiment was determined. Results for each concentration were then normalized by the maximum response for the relevant experiment, and values of mean response  $\pm$  SD were plotted in Figure 4C as log[ligand] versus response to determine mean EC<sub>50</sub> values.

### **Ligand-induced dimerization analysis by quantitative FRET microscopy**

Error bars in Figures 5B and 5C show standard error of the mean in the y-axis, and standard deviation of the mean in the x-axis, for data binned as described above. For EPGN and EREG, plots in Figure 5 and S5 represent data merged from four quantitative FRET experiments; for each ligand, two biological repeats were each measured twice. Data plotted for EGF represent an aggregate of three independent experiments from the same stock of ligand. Data were plotted using GraphPad Prism.

### **Mobility analysis by Single-Particle Tracking**

Data were analyzed from three independent experiments, each employing a different ligand preparation. Diffusion coefficients were obtained by a weighted linear fit of time intervals 1 to 5 of the mean square displacement (MSD) plot. Differences in diffusion distributions were evaluated for statistical significance (P value < 0.05) using Welch's *t*-test. Data were plotted and evaluated using GraphPad Prism.

### **Quantification of receptor phosphorylation in MCF-7 cells**

Signal intensities of a given protein were normalized by the corresponding peak value at 5 min stimulation, and expressed as percentages of the peak value. Independent ligand preparations were used for each experiment repetition. Results were plotted using GraphPad Prism.

### **Western blot image manipulation**

Raw images from the LI-COR Odyssey Fc (Figure 4B) or Kodak Image Station (all other gel figures) were imported into Adobe Photoshop, and linear contrast stretching was manually applied using the 'Levels' function – so that the darkest points of the image are black, and background is brought into the visible grey scale so that all features are registered. For multiple parallel experiments that employed the same camera exposure time, this procedure

was standardized (selecting the same upper and lower bounds). Cropped images of gel slices are individually boxed in the figures shown.

### Structure determination and analysis

The statistical analysis of the structural models is provided in Table S1. Analysis of molecular contacts, rmsd values and buried surface areas were calculated using the CCP4 software package (CCP4, 1994).

### Data and software availability

The coordinates and structure factors for the EREG/sEGFR501 and EPGN/sEGFR501 have been deposited in the protein data bank (PDB) with accession numbers of 5WB7 and 5WB8 respectively.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

We thank members of the Lemmon and Ferguson laboratories for helpful discussions and comments, and Kalina Hristova for access to microscopes. Crystallographic data were collected at the Advanced Photon Source, Argonne National Laboratory using GM/CA beamline 23-ID, with assistance at the 8<sup>th</sup> CCP4/APS School. GM/CA has been funded by NCI (ACB-12002) and NIGMS (AGM-12006). This work was supported by NIH grants F32-GM109688 (D.M.F.), T32-GM008275 (N.J.B.), R01-CA112552 (K.M.F.), R01-GM099092 (D.J.L.), P50-GM085273 (D.S.L.), R01-CA198164 (M.A.L. and K.M.F.), and U54-CA193417 (M.A.L.).

### References

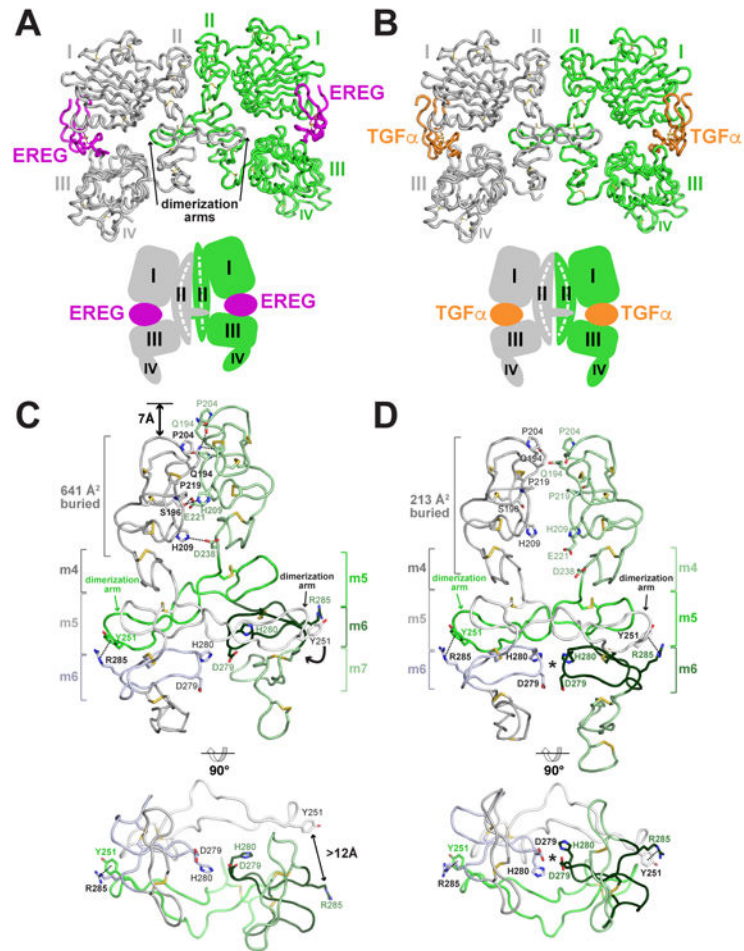
- Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 2010; 66:213–221. [PubMed: 20124702]
- Aksamitiene E, Hoek JB, Kiyatkin A. Multistrip Western blotting: a tool for comparative quantitative analysis of multiple proteins. *Methods Mol Biol.* 2015; 1312:197–226. [PubMed: 26044004]
- Alvarado D, Klein DE, Lemmon MA. ErbB2 resembles an autoinhibited invertebrate epidermal growth factor receptor. *Nature.* 2009; 461:287–291. [PubMed: 19718021]
- Alvarado D, Klein DE, Lemmon MA. Structural basis for negative cooperativity in growth factor binding to an EGF receptor. *Cell.* 2010; 142:568–579. [PubMed: 20723758]
- Bessman NJ, Bagchi A, Ferguson KM, Lemmon MA. Complex relationship between ligand binding and dimerization in the epidermal growth factor receptor. *Cell Rep.* 2014; 9:1306–1317. [PubMed: 25453753]
- CCP4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr.* 1994; 50:760–763. [PubMed: 15299374]
- Chen L, Novicky L, Merzlyakov M, Hristov T, Hristova K. Measuring the energetics of membrane protein dimerization in mammalian membranes. *J Am Chem Soc.* 2010a; 132:3628–3635. [PubMed: 20158179]
- Chen L, Placone J, Novicky L, Hristova K. The extracellular domain of fibroblast growth factor receptor 3 inhibits ligand-independent dimerization. *Sci Signal.* 2010b; 3:ra86. [PubMed: 21119106]
- Chung I, Akita R, Vandlen R, Toomre D, Schlessinger J, Mellman I. Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature.* 2010; 464:783–787. [PubMed: 20208517]

- Dawson JP, Berger MB, Lin CC, Schlessinger J, Lemmon MA, Ferguson KM. Epidermal growth factor receptor dimerization and activation require ligand-induced conformational changes in the dimer interface. *Mol Cell Biol.* 2005; 25:7734–7742. [PubMed: 16107719]
- Del Piccolo N, Placone J, He L, Agudelo SC, Hristova K. Production of plasma membrane vesicles with chloride salts and their utility as a cell membrane mimetic for biophysical characterization of membrane protein interactions. *Anal Chem.* 2012; 84:8650–8655. [PubMed: 22985263]
- Doerner A, Scheck R, Schepartz A. Growth factor identity is encoded by discrete coiled-coil rotamers in the EGFR juxtamembrane region. *Chem Biol.* 2015; 22:776–784. [PubMed: 26091170]
- Eden ER, White IJ, Tsapara A, Futter CE. Membrane contacts between endosomes and ER provide sites for PTP1B-epidermal growth factor receptor interaction. *Nat Cell Biol.* 2010; 12:267–272. [PubMed: 20118922]
- Elenius K, Paul S, Allison G, Sun J, Klagsbrun M. Activation of HER4 by heparin-binding EGF-like growth factor stimulates chemotaxis but not proliferation. *EMBO J.* 1997; 16:1268–1278. [PubMed: 9135143]
- Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 2004; 60:2126–2132. [PubMed: 15572765]
- Ferguson KM. Structure-based view of epidermal growth factor receptor regulation. *Annu Rev Biophys.* 2008; 37:353–373. [PubMed: 18573086]
- Ferguson KM, Berger MB, Mendrola JM, Cho HS, Leahy DJ, Lemmon MA. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol Cell.* 2003; 11:507–517. [PubMed: 12620237]
- Ferguson KM, Darling PJ, Mohan MJ, Macatee TL, Lemmon MA. Extracellular domains drive homo- but not hetero-dimerization of erbB receptors. *EMBO J.* 2000; 19:4632–4643. [PubMed: 10970856]
- Francavilla C, Papetti M, Rigbolt KT, Pedersen AK, Sigurdsson JO, Cazzamali G, Karemora G, Blagoev B, Olsen JV. Multilayered proteomics reveals molecular switches dictating ligand-dependent EGFR trafficking. *Nat Struct Mol Biol.* 2016; 23:608–618. [PubMed: 27136326]
- Furness SG, Liang YL, Nowell CJ, Halls ML, Wookey PJ, Dal Maso E, Inoue A, Christopoulos A, Wootten D, Sexton PM. Ligand-dependent modulation of G protein conformation alters drug efficacy. *Cell.* 2016; 167:739–749. [PubMed: 27720449]
- Garrett TPJ, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Kofler M, Jorissen RN, Nice EC, Burgess AW, et al. The crystal structure of a truncated ErbB2 ectodomain reveals an active conformation, poised to interact with other ErbB receptors. *Mol Cell.* 2003; 11:495–505. [PubMed: 12620236]
- Garrett TPJ, McKern NM, Lou M, Elleman TC, Adams TE, Lovrecz GO, Zhu HJ, Walker F, Frenkel MJ, Hoyne PA, et al. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell.* 2002; 110:763–773. [PubMed: 12297049]
- Harris RC, Chung E, Coffey RJ. EGF receptor ligands. *Exp Cell Res.* 2003; 284:2–13. [PubMed: 12648462]
- Herrero A, Casar B, Colón-Bolea P, Agudo-Ibáñez L, Crespo P. Defined spatiotemporal features of RAS-ERK signals dictate cell fate in MCF-7 mammary epithelial cells. *Mol Biol Cell.* 2016; 27:1958–1968. [PubMed: 27099370]
- Ho CC, Chhabra A, Starkl P, Schnorr PJ, Wilmes S, Moraga I, Kwon HS, Gaudenzio N, Sibilano R, Wehrman TS, et al. Decoupling the functional pleiotropy of stem cell factor by tuning c-Kit signaling. *Cell.* 2017; 168:1041–1052. [PubMed: 28283060]
- Kavran JM, McCabe JM, Byrne PO, Connacher MK, Wang Z, Ramek A, Sarabipour S, Shan Y, Shaw DE, Hristova K, et al. How IGF-1 activates its receptor. *Elife.* 2014; 3:e03772.
- Kholodenko BN, Demin OV, Moehren G, Hoek JB. Quantification of short term signaling by the epidermal growth factor receptor. *J Biol Chem.* 1999; 274:30169–30181. [PubMed: 10514507]
- Kim AR, Ulirsch JC, Wilmes S, Unal E, Moraga I, Karakukcu M, Yuan D, Kazerounian S, Abdulhay NJ, King DS, et al. Functional selectivity in cytokine signaling revealed through a pathogenic EPO mutation. *Cell.* 2017; 168:1053–1064. [PubMed: 28283061]

- Klein DE, Stayrook SE, Shi F, Narayan K, Lemmon MA. Structural basis for EGFR ligand sequestration by Argos. *Nature*. 2008; 453:1271–1275. [PubMed: 18500331]
- Klein Herenbrink C, Sykes DA, Donthamsetti P, Canals M, Coudrat T, Shonberg J, Scammells PJ, Capuano B, Sexton PM, Charlton SJ, et al. The role of kinetic context in apparent biased agonism at GPCRs. *Nat Commun*. 2016; 7:10842. [PubMed: 26905976]
- Knudsen SL, Mac AS, Henriksen L, van Deurs B, Grøvdal LM. EGFR signaling patterns are regulated by its different ligands. *Growth Factors*. 2014; 32:155–163. [PubMed: 25257250]
- Kochupurakkal BS, Harari D, Di-Segni A, Maik-Rachline G, Lyass L, Gur G, Kerber G, Citri A, Lavi S, Eilam R, et al. Epigen, the last ligand of ErbB receptors, reveals intricate relationships between affinity and mitogenicity. *J Biol Chem*. 2005; 280:8503–8512. [PubMed: 15611079]
- Komurasaki T, Toyoda H, Uchida D, Morimoto S. Epiregulin binds to epidermal growth factor receptor and ErbB-4 and induces tyrosine phosphorylation of epidermal growth factor receptor, ErbB-2, ErbB-3 and ErbB-4. *Oncogene*. 1997; 15:2841–2848. [PubMed: 9419975]
- Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Cryst*. 2003; 36:1277–1282.
- Kovacs E, Zorn JA, Huang Y, Barros T, Kuriyan J. A structural perspective on the regulation of the epidermal growth factor receptor. *Annu Rev Biochem*. 2015; 84:739–764. [PubMed: 25621509]
- Lane JR, May LT, Parton RG, Sexton PM, Christopoulos A. A kinetic view of GPCR allosterity and biased agonism. *Nat Chem Biol*. 2017; 13:929–937. [PubMed: 28820879]
- Lemmon MA, Bu Z, Ladbury JE, Zhou M, Pinchasi D, Lax I, Engelman DM, Schlessinger J. Two EGF molecules contribute additively to stabilization of the EGFR dimer. *EMBO J*. 1997; 16:281–294. [PubMed: 9029149]
- Lemmon MA, Freed DM, Schlessinger J, Kiyatkin A. The dark side of cell signaling: Positive roles for negative regulators. *Cell*. 2016; 164:1172–1184. [PubMed: 26967284]
- Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. *Cell*. 2010; 141:1117–1134. [PubMed: 20602996]
- Liu P, Cleveland TE 4th, Bouyain S, Byrne PO, Longo PA, Leahy DJ. A single ligand is sufficient to activate EGFR dimers. *Proc Natl Acad Sci U S A*. 2012; 109:10861–10866. [PubMed: 22699492]
- Low-Nam ST, Lidke KA, Cutler PJ, Roovers RC, van Bergen en Henegouwen PM, Wilson BS, Lidke DS. ErbB1 dimerization is promoted by domain co-confinement and stabilized by ligand binding. *Nat Struct Mol Biol*. 2011; 18:1244–1249. [PubMed: 22020299]
- Lu C, Mi LZ, Grey MJ, Zhu J, Graef E, Yokoyama S, Springer TA. Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. *Mol Cell Biol*. 2010; 30:5432–5443. [PubMed: 20837704]
- Lu C, Mi LZ, Schürpf T, Walz T, Springer TA. Mechanisms for kinase-mediated dimerization of the epidermal growth factor receptor. *J Biol Chem*. 2012; 287:38244–38253. [PubMed: 22988250]
- Macdonald-Obermann JL, Pike LJ. Different epidermal growth factor (EGF) receptor ligands show distinct kinetics and biased or partial agonism for homodimer and heterodimer formation. *J Biol Chem*. 2014; 289:26178–26188. [PubMed: 25086039]
- Manglik A, Kim TH, Masureel M, Altenbach C, Yang Z, Hilger D, Lerch MT, Kobilka TS, Thian FS, Hubbell WL, et al. Structural insights into the dynamic process of  $\beta$ 2-adrenergic receptor signaling. *Cell*. 2015; 161:1101–1111. [PubMed: 25981665]
- Marshall CJ. Specificity of receptor tyrosine kinase signaling: transient versus sustained extracellular signal-regulated kinase activation. *Cell*. 1995; 80:179–185. [PubMed: 7834738]
- McKeithan TW. Kinetic proofreading in T-cell receptor signal transduction. *Proc Natl Acad Sci U S A*. 1995; 92:5042–5046. [PubMed: 7761445]
- Mi LZ, Lu C, Li Z, Nishida N, Walz T, Springer TA. Simultaneous visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor. *Nat Struct Mol Biol*. 2011; 18:984–989. [PubMed: 21822280]
- Nagashima T, Shimodaira H, Ide K, Nakakuki T, Tani Y, Takahashi K, Yumoto N, Hatakeyama M. Quantitative transcriptional control of ErbB receptor signaling undergoes graded to biphasic response for cell differentiation. *J Biol Chem*. 2007; 282:4045–4056. [PubMed: 17142811]

- Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, Pan AC, Liu CW, Fung JJ, Bokoch MP, et al. The dynamic process of  $\beta(2)$ -adrenergic receptor activation. *Cell*. 2013; 152:532–542. [PubMed: 23374348]
- Ogiso H, Ishitani R, Nureki O, Fukai S, Yamanaka M, Kim JH, Saito K, Sakamoto A, Inoue M, Shirouzu M, et al. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell*. 2002; 110:775–787. [PubMed: 12297050]
- Red Brewer M, Choi SH, Alvarado D, Moravcevic K, Pozzi A, Lemmon MA, Carpenter G. The juxtamembrane region of the EGF receptor functions as an activation domain. *Mol Cell*. 2009; 34:641–651. [PubMed: 19560417]
- Riese DJ 2nd, Bermingham Y, van Raaij TM, Buckley S, Plowman GD, Stern DF. Betacellulin activates the epidermal growth factor receptor and erbB-4, and induces cellular response patterns distinct from those stimulated by epidermal growth factor or neuregulin-beta. *Oncogene*. 1996; 12:345–353. [PubMed: 8570211]
- Riese DJ 2nd, Cullum RL. Epiregulin: roles in normal physiology and cancer. *Semin Cell Dev Biol*. 2014; 28:49–56. [PubMed: 24631357]
- Rizzi M, Pittarella P, Sabbatini M, Renò F. Epiregulin induces human SK-N-BE cell differentiation through ERK1/2 signaling pathway. *Growth Factors*. 2013; 31:90–97. [PubMed: 23734838]
- Roepstorff K, Grandal MV, Henriksen L, Knudsen SL, Lerdrup M, Grovdal L, Willumsen BM, van Deurs B. Differential effects of EGF ligands on endocytic sorting of the receptor. *Traffic*. 2009; 10:1115–1127. [PubMed: 19531065]
- Ronan T, Macdonald-Obermann JL, Huelsmann L, Bessman NJ, Naegle KM, Pike LJ. Different epidermal growth factor receptor (EGFR) agonists produce unique signatures for the recruitment of downstream signaling proteins. *J Biol Chem*. 2016; 291:5528–5540. [PubMed: 26786109]
- Scheck RA, Lowder MA, Appelbaum JS, Schepartz A. Bipartite tetracysteine display reveals allosteric control of ligand-specific EGFR activation. *ACS Chem Biol*. 2012; 7:1367–1376. [PubMed: 22667988]
- Shi T, Niepel M, McDermott JE, Gao Y, Nicora CD, Chrisler WB, Markillie LM, Petyuk VA, Smith RD, Rodland KD, et al. Conservation of protein abundance patterns reveals the regulatory architecture of the EGFR-MAPK pathway. *Sci Signal*. 2016; 9:rs6. [PubMed: 27405981]
- Sparta B, Pargett M, Minguet M, Distor K, Bell G, Albeck JG. Receptor level mechanisms are required for epidermal growth factor (EGF)-stimulated extracellular signal-regulated kinase (ERK) activity pulses. *J Biol Chem*. 2015; 290:24784–24792. [PubMed: 26304118]
- Swain PS, Siggia ED. The role of proofreading in signal transduction specificity. *Biophys J*. 2002; 82:2928–2933. [PubMed: 12023215]
- Sweeney C, Carraway KL 3rd. Ligand discrimination by ErbB receptors: differential signaling through differential phosphorylation site usage. *Oncogene*. 2000; 19:5568–5573. [PubMed: 11114736]
- Valley CC, Arndt-Jovin DJ, Karedla N, Steinkamp MP, Chizhik AI, Hlavacek WS, Wilson BS, Lidke KA, Lidke DS. Enhanced dimerization drives ligand-independent activity of mutant epidermal growth factor receptor in lung cancer. *Mol Biol Cell*. 2015; 26:4087–4099. [PubMed: 26337388]
- Wacker D, Stevens RC, Roth BL. How ligands illuminate GPCR molecular pharmacology. *Cell*. 2017; 170:414–427. [PubMed: 28753422]
- Willmarth NE, Ethier SP. Autocrine and juxtacrine effects of amphiregulin on the proliferative, invasive, and migratory properties of normal and neoplastic human mammary epithelial cells. *J Biol Chem*. 2006; 281:37728–37737. [PubMed: 17035230]
- Wilson KJ, Gilmore JL, Foley J, Lemmon MA, Riese DJ 2nd. Functional selectivity of EGF family peptide growth factors: implications for cancer. *Pharmacol Ther*. 2009; 122:1–8. [PubMed: 19135477]
- Wilson KJ, Mill C, Lambert S, Buchman J, Wilson TR, Hernandez-Gordillo V, Gallo RM, Ades LM, Settleman J, Riese DJ 2nd. EGFR ligands exhibit functional differences in models of paracrine and autocrine signaling. *Growth Factors*. 2012; 30:107–116. [PubMed: 22260327]
- Wolber PK, Hudson BS. An analytic solution to the Förster energy transfer problem in two dimensions. *Biophys J*. 1979; 28:197–210. [PubMed: 262548]

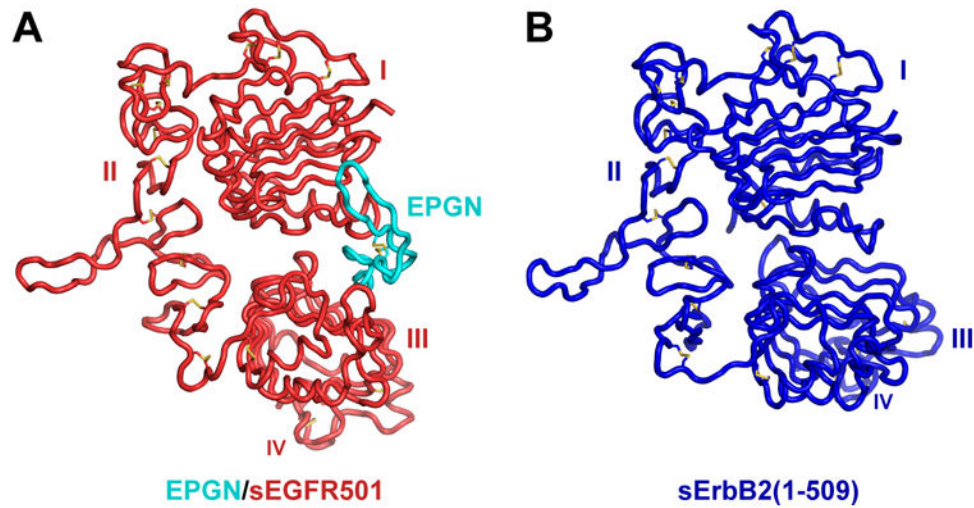




### Figure 1. Epiregulin induces asymmetric sEGFR dimers

Distinct sEGFR501 dimer structures induced by (A) EREG and (B) TGFα (PDB ID 1MOX), aligned using the left (grey) protomer. Asymmetry of the EREG/sEGFR501 complex is emphasized in the lower cartoons, depicting ‘bent’ and ‘straight’ domain II configurations with white dashed lines.

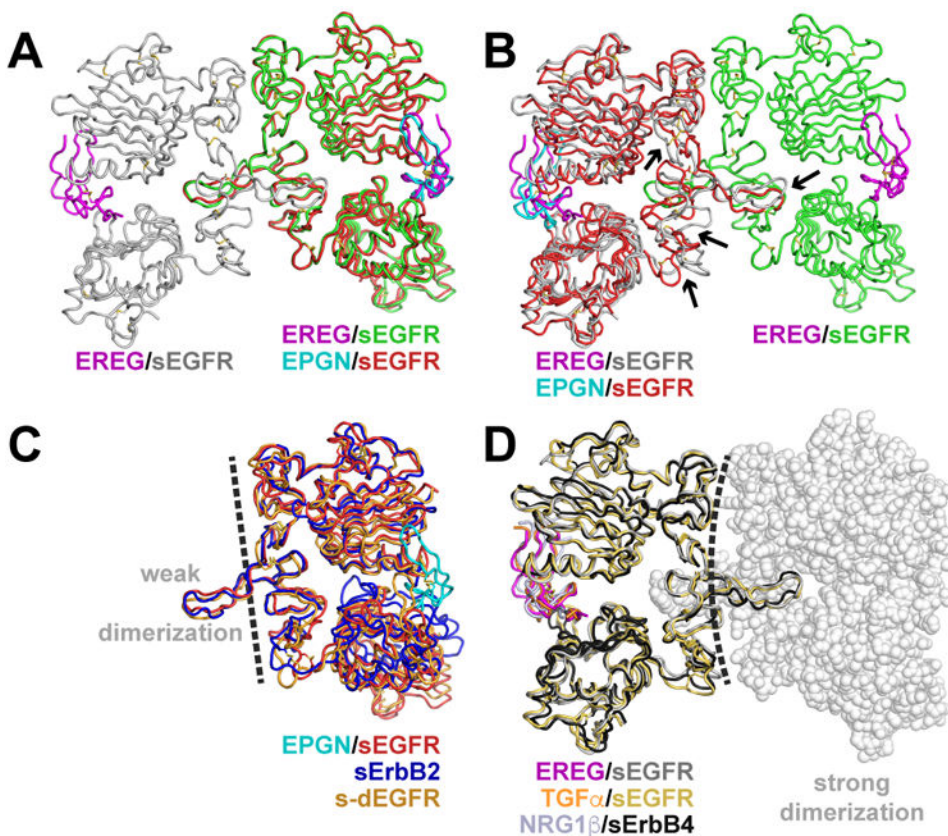
Close-up of domain II dimer interfaces induced by (C) EREG and (D) TGFα viewed from the side (upper) and bottom (lower). In (C), the side view shows a 7 Å upward shift of the green receptor (right) relative to the grey receptor (left) in the EREG/sEGFR501 complex. C-terminal disulfide-bonded modules in each domain II are colored different shades of grey or green, with selected interface residues labeled – in bold when involved directly in interactions. The arrow on the right of (C) denotes an outward shift of the grey dimerization arm (including Y251) that prevents Y251/R285 contacts. Asterisk in (D) marks key ‘buttressed’ intermolecular contacts involving D279 and H280. See also Figure S1 and Table S1.



**Figure 2. Epigen-bound sEGFR is monomeric**

(A) Ribbon structure of EPGN-bound sEGFR501, with sEGFR501 colored red and EPGN cyan.

(B) Structure of sErbB2 (residues 1-509 – analogous to sEGFR501) in the same orientation as in (A), from PDB ID 2A91. See also Figures S2 and S3, and Table S1.

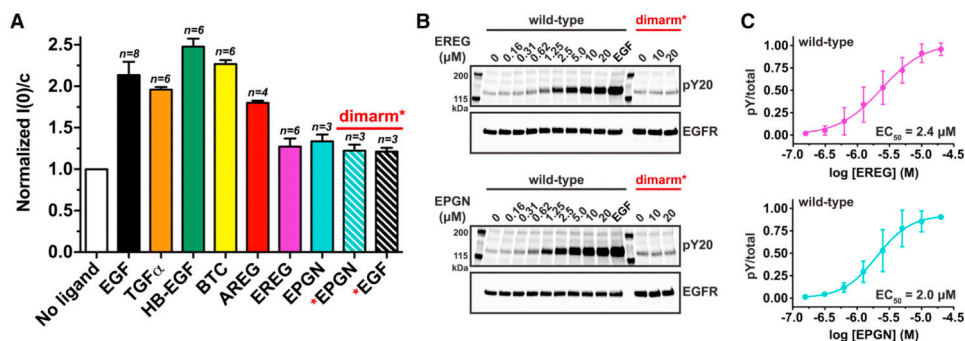


**Figure 3. Distinct sEGFR domain II conformations for EREG and EPGN**

, (A) EPGN-bound sEGFR501 (red) overlays well with the right-hand (green) molecule of the EREG/sEGFR501 dimer, but deviates significantly (B) when overlaid on the left-hand (grey) molecule. Areas of significant divergence are highlighted with black arrows.

(C) EPGN-bound sEGFR501 (red) overlays well with the unliganded *Drosophila* EGFR extracellular region (s-dEGFR), shown in orange/brown, and sErbB2(1-509) shown in dark blue. The right-hand (green) molecule of the EREG/sEGFR501 dimer also falls into this category. Domain II is unbent in each of these structures, as depicted by the straight black dashed line.

(D) The left-hand (grey) molecule of the asymmetric EREG/sEGFR501 dimer overlays well with both TGF $\alpha$ -bound sEGFR501, shown in gold and the NRG1 $\beta$ -bound ErbB4 extracellular region (sErbB4) from PDB ID 3U7U, shown in black (Liu et al., 2012). The domain II dimerization interface is distinctly bent in each of these dimerization-competent structures, as depicted by the curved black dashed line and interactions with the space-filling sEGFR501 model (from the TGF $\alpha$ /sEGFR501 dimer) shown at right. See also Figure S3.

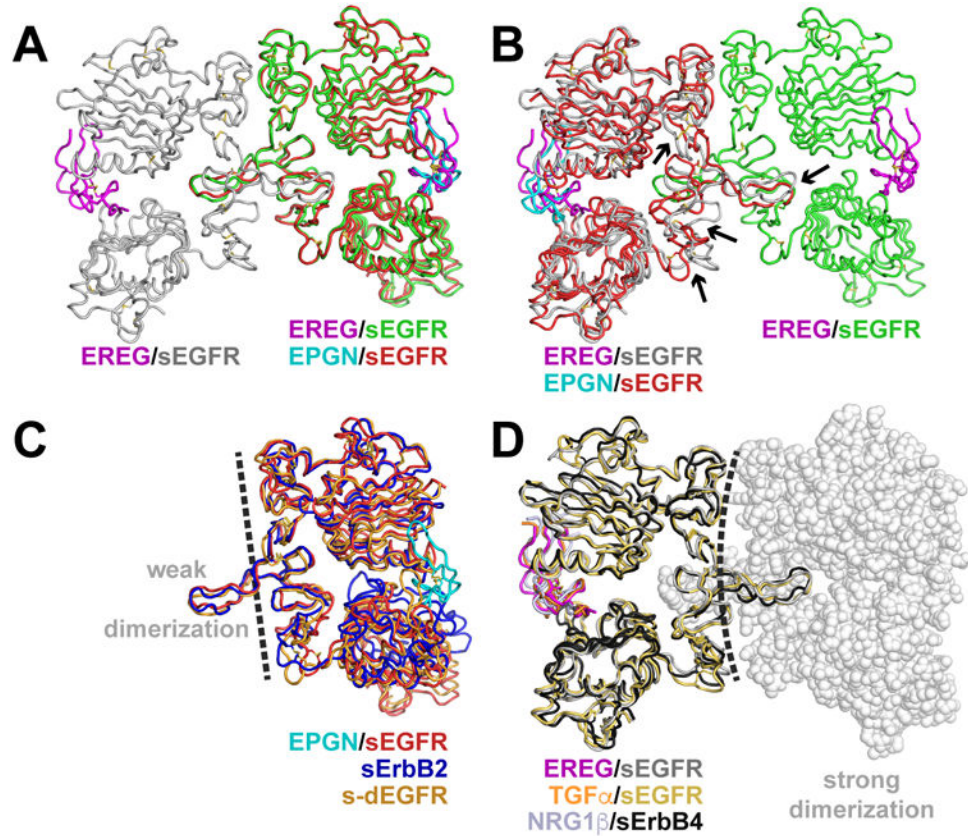


**Figure 4. EGFR dimerization and activation by different ligands**

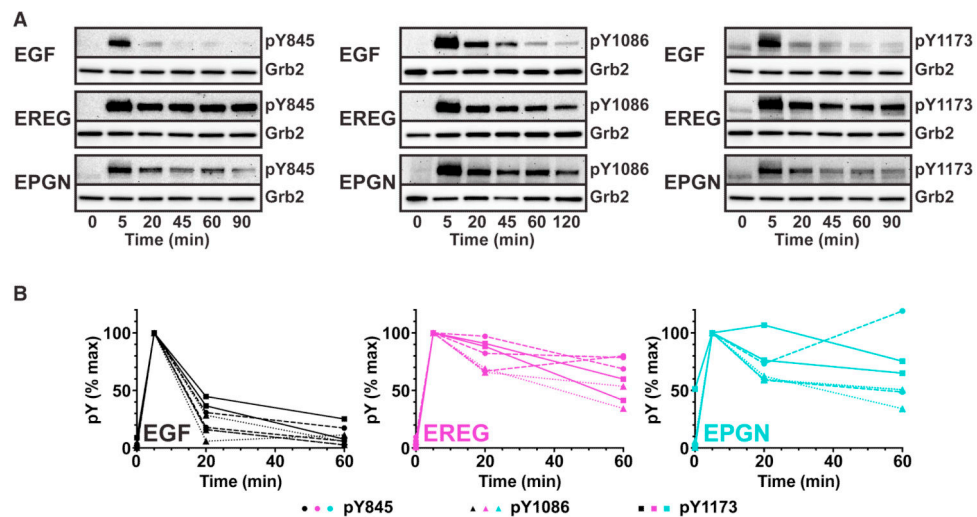
(A) SAXS-derived normalized  $I(0)/c$  values for 70  $\mu\text{M}$  sEGFR501 without ligand (white bar) or with saturating concentrations (84  $\mu\text{M}$ ) of EGFR ligands (colored bars).  $I(0)/c$  values represent fold increases over that seen for sEGFR501 monomers. Hatched bars represent data for dimerization arm-mutated (dimarm\*) sEGFR501. Values of the mean ( $\pm$  SD) and  $n$  are presented. Representative Guinier regions for each biological replicate are plotted in Figure S4.

(B) Activation of human EGFR in stable *Drosophila* S2 cell lines expressing wild-type human EGFR or the dimarm\* variant, stimulated with EREG (upper), EPGN (lower), or 100 nM EGF as positive control. Each experiment represents three biological repeats.

(C) LI-COR quantitation for data from three biological repeats ( $\pm$  SD) of the experiment shown in (B). See also Figure S4.



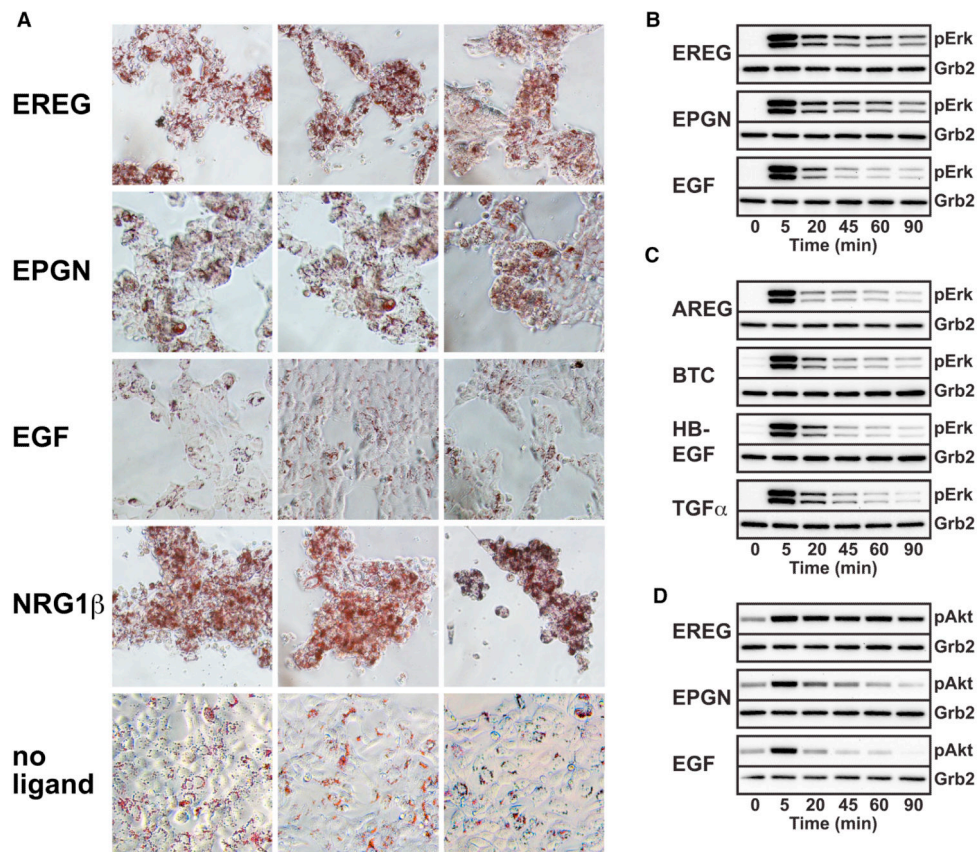
**Figure 5. Epiregulin and epigen induce weaker and shorter-lived EGFR dimers than EGF**  
**(A)** Cartoon of quantitative FRET experiments.  
**(B, C)** Quantitative  $\text{EGFR}_{\text{ECR-TM-eYFP}}/\text{EGFR}_{\text{ECR-TM-mCherry}}$  FRET data (see Methods) for EREG **(B)** and EPGN **(C)** at 1  $\mu\text{M}$  (open circles) or 20  $\mu\text{M}$  (closed circles), plotted as a function of receptor density. FRET with no ligand (grey open circles) or with saturating (100 nM) EGF (black circles) is plotted for comparison. Standard error is plotted in the  $y$ -axis and standard deviation in the  $x$ -axis, for binned data. Best-fit curves of unbinned data to the dimerization model described in Methods are plotted (see Table S2).  
**(D, E)** Single particle tracking of HA-tagged full-length EGFR in CHO cells. Diffusion of quantum dot-labeled EGFRs was monitored without ligand (grey), or with saturating EGF (50 nM), EREG (20  $\mu\text{M}$ ) or EPGN (20  $\mu\text{M}$ ). Ensemble mean square displacement (MSD) is plotted for  $N > 1834$  trajectories per condition **(D)**, and diffusion coefficient distribution across cells is plotted for  $N > 61$  cells per condition **(E)**. The inset in **(D)** shows MSD at small displacements with shaded areas representing 95% confidence intervals from fits. Distributions in **(E)** are compared using Welch's t-test (\* $P = 0.048$ , \*\* $P = 0.006$ ). See also Figure S5 and Table S2.



**Figure 6. EGFR activation by epiregulin or epigen is sustained**

(A) Representative time-courses of EGFR phosphorylation at Y1173, Y845, and Y1086 in MCF-7 cells induced by saturating levels of EGF (16 nM), EREG (20  $\mu$ M) or EPGN (20  $\mu$ M). Anti-Grb2 is used as loading control. Data for pY1173 were generated by stripping and reprobing pY845 blots, so use the same loading controls.

(B) Quantitation of EGFR phosphorylation time courses, normalized by signal at 5 minutes. Data are plotted on the same graph for multiple independent experiments quantitating phosphorylation at Y1173 (squares), Y845 (circles) and Y1086 (triangles). See also Figure S6.



**Figure 7. Sustained signaling by epiregulin and epigen promotes MCF-7 cell differentiation**

(A) Oil Red O staining of MCF-7 cells stimulated with saturating EREG (10  $\mu$ M), EPGN (10  $\mu$ M), EGF (16 nM), NRG1 $\beta$  (25 nM), or with no ligand.

(B, C) Representative time-courses of Erk phosphorylation (at T202 and Y204) in MCF-7 cells induced by saturating levels of each EGFR ligand. (B) EREG and EPGN were added at 20  $\mu$ M, and EGF at 16 nM. (C) AREG was added at 20  $\mu$ M, BTC and HB-EGF at 15 nM, and TGF $\alpha$  at 25 nM.

(D) Representative time courses of Akt S473 phosphorylation in MCF-7 cells after treatment with EREG (20  $\mu$ M), EPGN (20  $\mu$ M), or EGF (16 nM).

See also Figure S7.