Supplemental Data

Figure S1, related to Figure 2. Analysis of molecular mass of EGFR-Fc and sEGFR-Zip.

(**A**) Sedimentation equilibrium AUC data, showing single-species fits for data obtained with sEGFR-Fc alone at 10 μM at a speed of 10000 r.p.m. Data were fit using Sedphat (Schuck, 2003) and yielded an excellent fit – with small random residuals – to a single species of 223 kDa. This compares well with the expected mass of an sEGFR-Fc dimer (190 kDa plus ~20% w/w carbohydrate). (**B**) An equivalent experiment to that shown in **A** for sEGFR-Fc, but with the addition of a 1.2-fold molar excess of TGFα. The best-fit molecular mass is unchanged (at 225 kDa), showing that the dimeric sEGFR-Fc is not further oligomerized upon ligand binding. (**C**) Size exclusion chromatography analysis of sEGFR[wild-type] (black curve), sEGFR-Zip (blue curve), and sEGFR-Fc (orange curve), injected onto a Superose 6 column at a concentration of 10 μM. The receptor elution volumes reveal that sEGFR-Fc and sEGFR-Zip are both dimers, while sEGFR[wild-type] is monomeric as expected.

Table S1, related to Table 1. ITC data for TGFα binding to sEGFR variants.

sEGFR variant	K_D (nM)[a]	ΔH (kcal/mol)[a]	ΔG (kcal/mol)[b]	$T\Delta S$ (kcal/mol)[c]
sEGFR[wild-type]	80 ± 27	+8.2 ± 1.6	-9.7	17.9
sEGFR[Y251A/R285S]	82 ± 17	+10.1 ± 2.1	-9.7	19.8
sEGFR-Fc[wild-type]	nd	+11.0 ± 0.9	Nd	nd

[a]Values are the mean ± standard deviation of at least three independent experiments.
[b]ΔG values are calculated from the mean K_D.
[c]$T\Delta S$ values are obtained by subtracting ΔG from the mean value for ΔH.

A EGF binding to sEGFR dom III

B Dissociation of EGF/sEGFR dimer
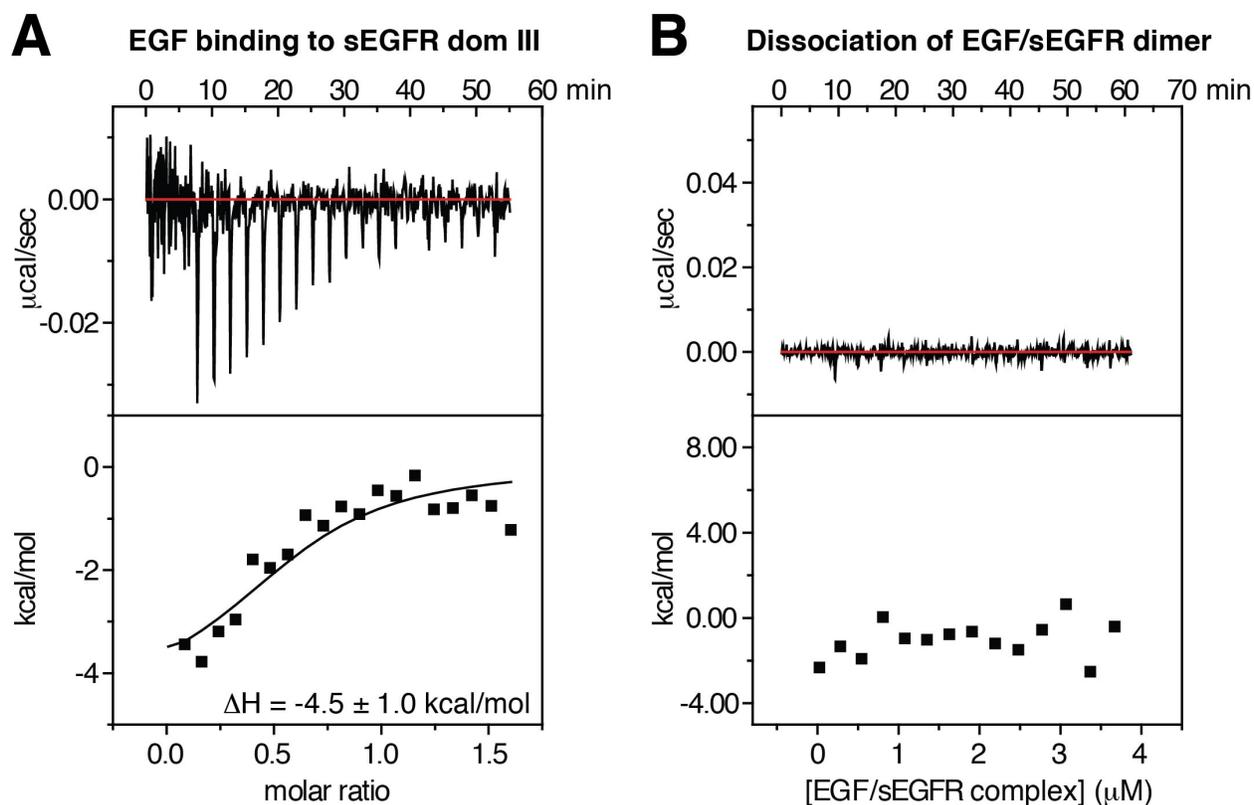
Figure S2, related to Figure 2. ITC of EGF binding to domain III, and EGF/sEGFR dimer dissociation.

(**A**) EGF was titrated into 10 μM sEGFR domain III purified as described (Lemmon et al., 1997). Fitting to a single binding-site model yielded a K_D value of 3.1 ± 2.6 μM, and a ΔH for EGF binding of -4.5 ± 1.0 kcal/mol. (**B**) To assess the enthalpy associated with dimerization of liganded sEGFR, a dilution experiment was performed in which 17.5 μM sEGFR[wild-type] to which a 1.3-fold molar excess of EGF had been added was diluted by injection into an ITC cell containing only buffer. 13 injections of 3 μl each were made, allowing measurement of dissociation heats in the ITC cell over an equilibrium concentration range of ~0.2 – 2.4 μM for the EGF/sEGFR[wild-type] complex. Importantly, EGF dissociation from sEGFR[wild-type] will be negligible over this concentration range. Because the heat of injection does not change systematically from the first injection to the last, and because the integrated heat of each injection is so low (and can not be distinguished from instrumental noise), we conclude that the ΔH for dimerization of EGF-bound sEGFR[wild-type] must be <<|2| kcal/mol.

The titrations in Figures 2B and 2C fit very well to a single entropy-driven (positive ΔH) ligand binding event. In our 1997 studies (Lemmon et al., 1997), which pre-dated structures of sEGFR, we inferred that this entropy-driven event reflects sEGFR dimerization, and that ligand binding has a small negative ΔH (based on studies of EGF binding to isolated domain III, repeated in **A**). Our new finding that the major entropy-driven event is unaffected by mutations that abolish sEGFR dimerization proves this wrong. Moreover, based on calorimetric dissociation experiments (**B**) we find that dimerization of ligand-bound sEGFR[wild-type] has a negligible ΔH. Thus, EGF binding is entropy driven, consistent with another recent report (Alvarenga et al., 2012).

3
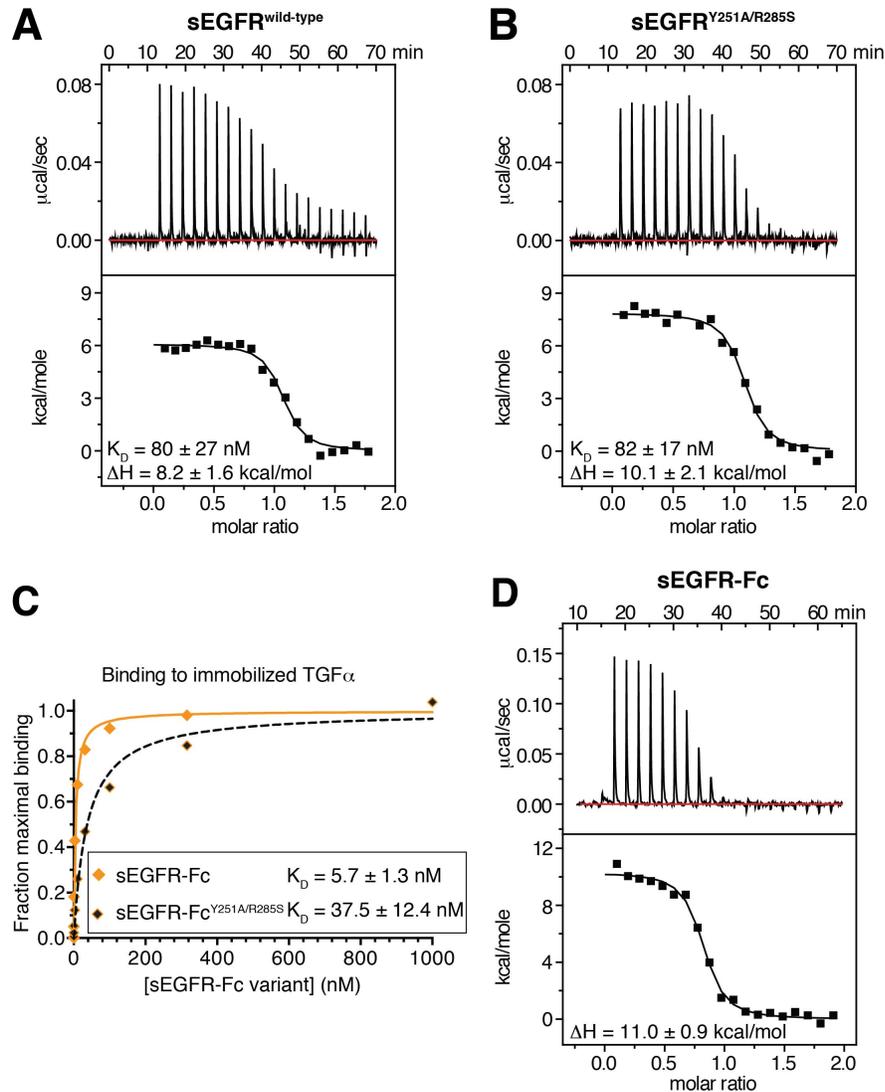
Figure S3, related to Figure 2. TGFα binding to sEGFR and sEGFR-Fc variants.

(**A**) Representative ITC analysis of TGFα binding to sEGFR[wild-type] at 25°C. TGFα was present in the syringe at 70 μM, and sEGFR[wild-type] was present in the calorimeter cell at a concentration of 8 μM. Mean values (± standard deviation) for K_D and ΔH for TGFα binding from three independent experiments are listed in the figure. (**B**) Equivalent experiment to that shown in **A** for sEGFR[Y251A/R285S]. (**C**) SPR analysis of TGFα binding to sEGFR-Fc with wild-type domain II or with the Y251A/R285S mutations in the domain II dimerization interface. Note that, whereas these mutations had no effect on EGF binding to sEGFR-Fc, they appear to reduce TGFα binding by ~6-fold, possibly suggesting a slightly different dependence on dimerization for binding of the two ligands – consistent with the observed structural differences in EGF/sEGFR and TGFα/sEGFR complexes (Liu et al., 2012; Wilson et al., 2009). (**D**) ITC analysis of TGFα binding to sEGFR-Fc. TGFα was present in the syringe at 70 μM, and sEGFR-Fc was present in the calorimeter cell at 9 μM. As with EGF binding, TGFα binding to sEGFR-Fc has a higher (positive or unfavorable) enthalpy than binding to sEGFR[wild-type] – by 2.8 kcal/mol (compared with 3.4 kcal/mol in the EGF case) – affinity is increased due to entropic effects.

sEGFR-Fc + EGF   **sEGFR-Fc alone**
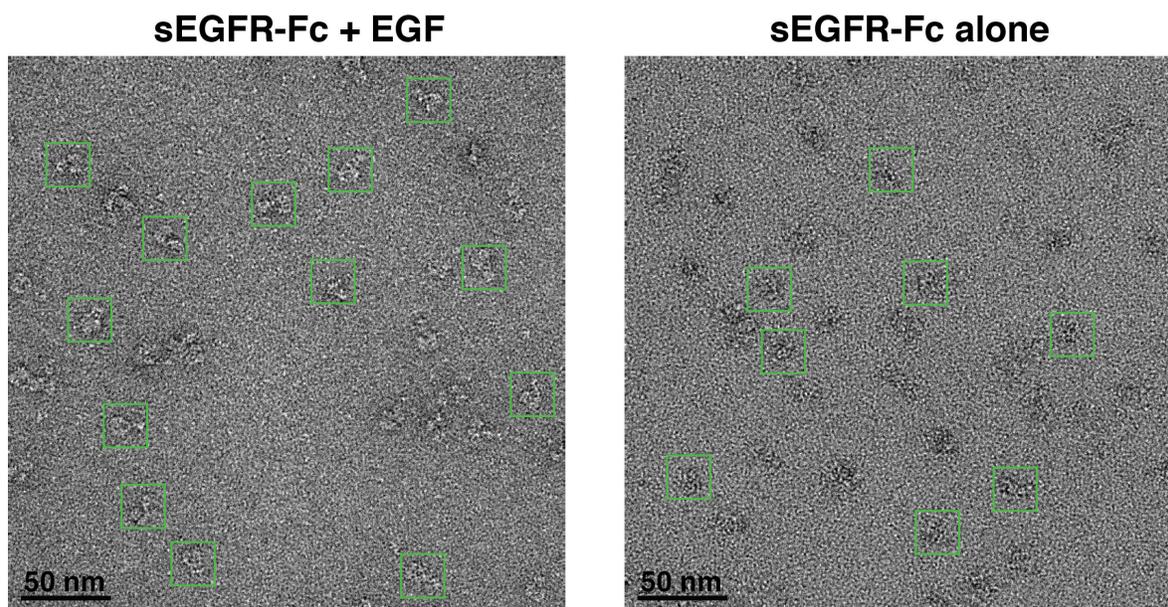
Figure S4, related to Figure 4. Example raw EM images for sEGFR-Fc plus and minus ligand.

Examples of raw image files are shown, from EM studies of negatively stained samples, for sEGFR-Fc plus EGF (left panel) and without ligand (right panel). 150 images such as these were used for single-particle analysis as described in Supplemental Experimental Procedures. Example particles are enclosed in green boxes in each representative image. A 50 nm scale bar is shown.
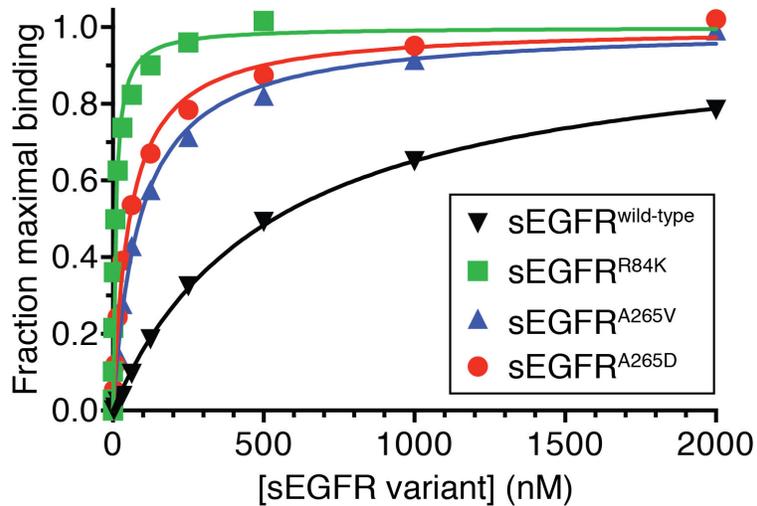
Figure S5, related to Figure 6. TGFα binding to sEGFR variants harboring glioblastoma mutations.

Binding of each glioblastoma-mutated sEGFR variant to TGFα was analyzed by SPR, using the same approach described for EGF binding in Figure 6B, but with immobilized TGFα instead of EGF. Best fit K_D values (± SD) for TGFα binding (from three independent experiments) were 532 ± 39 nM (sEGFR[wild-type]); 8.6 ± 1.5 nM (sEGFR[R84K]); 90 ± 10 nM (sEGFR[A265V]); and 55 ± 7 nM (sEGFR[A265D]). Glioblastoma mutations in EGFR thus enhance TGFα binding to the isolated ECR by 6-62 fold.
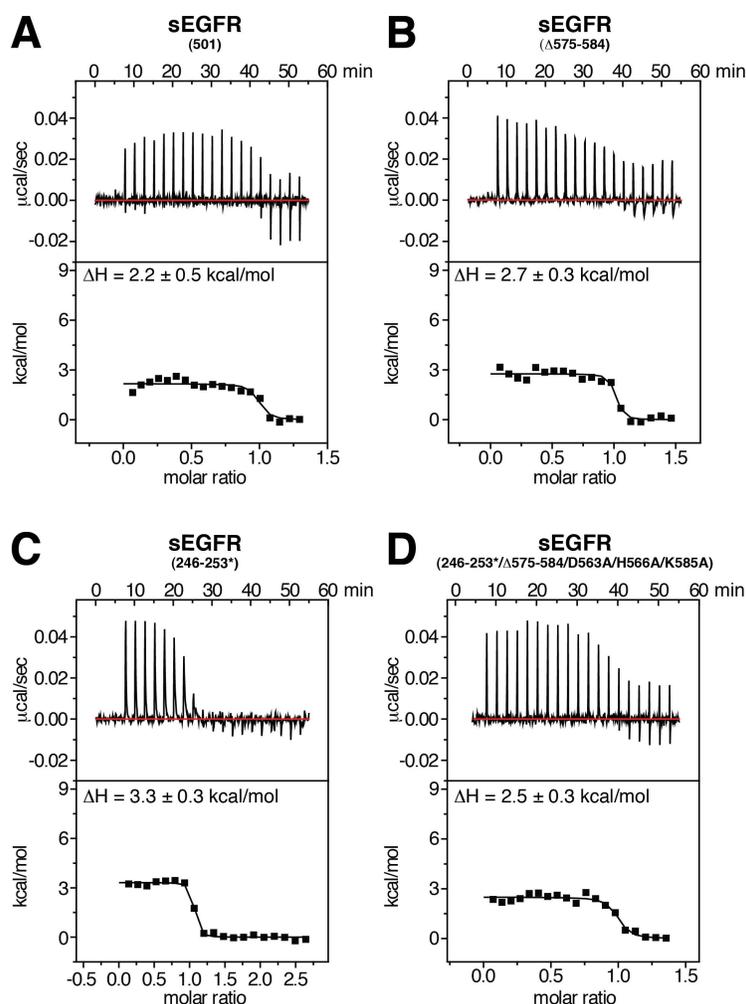
Figure S6, related to Figure 6. ITC studies of EGF binding to sEGFR with tether-disrupting mutations.

(**A**) EGF was titrated at 25°C into a cell containing 10 μM sEGFR(501), a variant of sEGFR that lacks essentially all of domain IV (Elleman et al., 2001), so cannot make the intramolecular tether. K_D for EGF binding to this variant is 7.8 nM (Dawson et al., 2005), and ΔH is 2.2 ± 0.5 kcal/mol. (**B**) The Δ575-584 mutation deletes a prominent loop from the fifth disulfide-bonded module of domain IV in sEGFR, removing a loop that interacts with the dimerization arm in the tethered sEGFR structure – and thus weakening the tether (Dawson et al., 2005). K_D for EGF binding to this variant is 32 nM (Dawson et al., 2005), and ΔH is 2.7 ± 0.3 kcal/mol. (**C**) The 246-253* mutation was initially described by Garrett et al. (2002), and includes the following 6 mutations, designed to break dimerization arm contacts: Y246E, N247A, T249D, Y251E, Q252A, and M253D. These mutations disrupt the intramolecular tether shown in Figure 1A and the dimerization contacts in Figure 1C. K_D for EGF binding to this variant is 260 nM by SPR (Dawson et al., 2005), and ΔH is 3.3 ± 0.3 kcal/mol. (**D**) The 246-253*/Δ575-584/D563A/H566A/K585A-mutated variant includes all of the changes in the other tether mutations described here, and also includes mutations of D563, H566, and K585 in domain IV to alanine. These residues contribute directly to the domain II/IV tether. Thus, this variant has all tether contacts as well as the dimerization site mutated. ΔH for EGF binding to this variant is 2.5 ± 0.3 kcal/mol. Errors are SDs.

Supplemental Experimental Procedures

sEGFR constructs

Construction of sEGFR variants with mutations in the domain II/IV tether or the dimerization arm was described previously (Dawson et al., 2005). Glioblastoma-derived mutations were introduced by standard PCR methods. To generate the sEGFR-Fc fusion, the coding sequence for sEGFR[wild-type] in pFastbac (pFb) was first mutated to introduce an Fse I restriction site close to its 3' end. The cDNA encoding the Fc domain from human IgG1 (IMAGE clone 4575935) was purchased from Open Biosystems, and was used as a PCR template to create a fragment containing the Fc domain coding region flanked by Fse I and Not I sites at the 5' and 3' ends, respectively. This fragment was then cloned into the pFb-sEGFR[wild-type] plasmid containing an Fse I site, yielding a plasmid encoding residues 1-645 of the pro-EGFR protein, followed by an alanine-glycine linker (introduced by the Fse I site), and then by the 231 residue Fc domain and a C-terminal hexa-histidine tag. The sEGFR-Zip pFb plasmid was constructed by first amplifying a fragment encoding the 33-residue coiled-coil domain from yeast GCN4 by a series of four PCR reactions, each extending the length of the fragment at the 3' end. This fragment also contained 20 nucleotides at the 5' end that were complementary to the C-terminus of sEGFR, as well as a hexa-histidine tag and Not I site at the 3' end. This fragment was used as a primer to PCR-amplify (from a pFb-sEGFR[wild-type] template) a fragment encoding sEGFR followed by the 33-residue coiled-coil domain and a C-terminal hexahistidine tag, which was then ligated into the pFb plasmid. All constructs were verified by DNA sequencing.

Protein purification

All sEGFR proteins were purified from the conditioned medium of baculovirus-infected Sf9 cells following the same procedure. Briefly, 4 days post-infection, clarified conditioned media was concentrated to ~1 l and diafiltered against 4 l of 25 mM Tris-HCl, pH 8.0, containing 150 mM

NaCl. Diafiltered media was applied to a Ni-NTA agarose column (Qiagen), and sEGFR protein was eluted with increasing concentrations of imidazole in binding buffer. Fractions containing sEGFR were dialyzed, buffer-exchanged, or diluted into 20 mM MES, pH 6.0, containing 50 mM NaCl and sEGFR was purified by cation-exchange chromatography using a Source S column (GE Healthcare) with a gradient of 0% - 100% of 20 mM MES, pH 6.0, containing 1 M NaCl (over 20 column volumes). The sEGFR-containing fractions were concentrated and protein was further purified by size-exclusion chromatography on a Superose 6 column (GE Healthcare) pre-equilibrated in 20 mM HEPES, pH 8.0, containing 150 mM NaCl. Typical yields were 0.2-1 mg/l of Sf9 cells (depending on the sEGFR variant) and protein was >90% pure by Coomassie stained SDS-PAGE.

Isothermal titration calorimetry (ITC)

All ITC experiments were performed at 25°C (unless specifically stated otherwise) using a MicroCal ITC200 instrument. Ligand and sEGFR proteins were dialyzed overnight into 20 mM HEPES, pH 8.0, containing 150 mM NaCl, and 3.4 mM EDTA. The sEGFR concentration in the calorimeter cell ranged from 8 to 25 μM, and the concentration of ligand in the syringe ranged from 60 to 280 μM. All protein concentrations were determined by measuring absorbance of purified protein at 280 nm and using the extinction coefficient predicted from primary amino acid sequence. A total of 39 μl of ligand (in 2-3 μl aliquots) was injected over the course of each titration. Data from the first (small) injection were discarded to eliminate syringe leakage artifacts. Ligand titrations into receptor-free ITC buffer were performed to determine the heat of ligand dilution, and these heats were subtracted from the ligand-into-receptor titration data. Data were fit to a single-site binding model in the Origin software package. All titrations were performed independently at least three times, and representative titrations are shown, with values for ΔH and other parameters quoted as mean ± standard deviation. Note that K_D values

were only fit for titrations in which c ([sites]/K_D) was less than 250 (Wiseman et al., 1989). Titrations where c > 250 were used for ΔH determination only.

Fluorescence anisotropy (FA)-based binding assays

EGF was labeled with Alexa-488 utilizing a tetrafluorophenyl ester to label primary amines, according to the protocol provided with the Alexa Fluor 488 Protein Labeling Kit from Molecular Probes (Eugene, OR). Labeled EGF (EGF[488]) was purified away from free label by size-exclusion chromatography on a Superdex Peptide (GE Healthcare) column equilibrated in 20 mM HEPES, pH 8.0, containing 150 mM NaCl. Labeling efficiency was calculated by determining the concentration of purified, labeled EGF from its absorbance at 280 nm, using an extinction coefficient of 18,825 cm^{-1}M^{-1} (as predicted from the primary sequence of EGF) and at 490 nm, using the extinction coefficient of the Alexa-488 label (71,000 cm^{-1}M^{-1}). EGF[488] at 10 nM (for sEGFR-Fc and sEGFR-Zip) or 60 nM (for sEGFR[wild-type]) was incubated with varying amounts of sEGFR protein for 30 minutes at room temperature in 20 mM HEPES, pH 8.0, containing 150 mM NaCl. Fluorescence polarization (FP) measurements for each sample were taken on a Beacon instrument at 20˚C. FP values were converted to anisotropy, and binding curves were derived by assuming that the maximal anisotropy response corresponded to [EGF$_{free}$]=0, and that the anisotropy in the absence of receptor corresponded to [EGF$_{free}$]=[EGF$_{total}$]. The resulting curves were fit to binding models using the GraphPad Prism software. sEGFR-Fc and sEGFR[wild-type] binding data were fit to simple single-site binding models, whereas sEGFR-Zip binding data were fit to a model with a Hill coefficient of ~1.7, presumably reflecting the presence of sEGFR-Zip monomers at very low receptor concentrations. Three independent titrations were performed for each receptor variant.

Surface plasmon resonance (SPR)

SPR binding experiments were performed using a Biacore 3000 instrument at 25°C. EGF and TGFα were immobilized as described (Dawson et al., 2005), immobilizing 100-150 response units (RUs). sEGFR variants at various concentrations in 25 mM HEPES, pH 8.0, containing 150 mM NaCl, 3.4 mM EDTA, and 0.005% Nonidet-p20 were injected at a flow rate of 10 µl/min for 10 minutes, which was sufficient to reach equilibrium even at the lowest concentrations. K_D values for binding of sEGFR variants to these surfaces were determined by fitting the equilibrium responses over a range of concentrations to a single-site Langmuir binding equation using GraphPad Prism 6.0. Surfaces were regenerated using 1 minute injections of 10 mM sodium acetate, pH 5.0, containing 1 M sodium chloride. Multiple rounds of regeneration did not impair sEGFR binding. All experiments were repeated independently at least three times.

Analytical ultracentrifugation (AUC)

Samples of sEGFR wild-type and the glioblastoma-derived variants at concentrations of 10 µM, 5 µM, and 2 µM were loaded into 6-hole sample cells in 20 mM HEPES, pH 8.0, containing 150 mM NaCl for sedimentation equilibrium AUC using a Beckman XL-A instrument and an An-Ti 60 analytical rotor at speeds of 6,000, 9,000 and 12,000 r.p.m. at room temperature. For experiments including TGFα, the ligand was present in a 1.2-fold molar excess over sEGFR protein concentration. Data were analyzed using a single species fit in Sedfit (Schuck et al., 2002) and Sedphat (Schuck, 2003), and were also fit to a monomer-dimer association model as described (Dawson et al., 2005), using the program HeteroAnalysis (UConn Biotechnology Bioservices Center).

Electron microscopy (EM)

For the EGF/sEGFR-Fc complex, 755 individual particles, manually picked from 150 images using the EMAN2 software package (Tang et al., 2007), were grouped into 10 classes by a reference-free alignment procedure in the program Spider (Shaikh et al., 2008). For the sEGFR-Fc protein alone, 2,566 particles from 673 images were grouped into 20 classes by the same reference-free alignment procedure. For Figure 4C, a model for the sEGFR-Fc fusion protein bound to EGF (shown in Figure 4B) was created by manually superposing PDB entries 3NJP (Lu et al., 2010) and the Fc portion of PDB entry 1HZH (Saphire et al., 2001). A 12 Å resolution map was calculated from this model using the molmap command in the UCSF Chimera software package (Pettersen et al., 2004). This low-resolution map was then used to create 2D projections from multiple angles using the makeboxref command in the EMAN2 software suite. The projection angle that best matched our experimental class averaging results is displayed.

Small-angle X-ray scattering (SAXS)

Scattering data were radially-averaged and reduced to two-dimensional plots using the SAXSgui software, and intensity data from buffer exposures were then subtracted out. Radius of gyration (R_g) values were determined from Guinier plots using the Primus software package (Konarev et al., 2003). The maximum interatomic distance (D_{max}) was obtained by examining P(r) curves generated by the Gnom software package (Svergun, 1992). Briefly, for each scattering dataset, P(r) curves were calculated for a range of D_{max} from 100 to 250 Å, in 5 Å increments. D_{max} was determined by identifying the value that gave the best fit to the experimental scattering data.

Supplemental References

Alvarenga, M.L., Kikhney, J., Hannewald, J., Metzger, A.U., Steffens, K.-J., Bomke, J., Krah, A., and Wegener, A. (2012). In-depth biophysical analysis of interactions between therapeutic antibodies and the extracellular domain of the epidermal growth factor receptor. Anal. Biochem. *421*, 138-151.

Konarev, P.V., Volkov, V.V., Sokolova, A.V., Koch, M.H.J., and Svergun, D.I. (2003). PRIMUS: a Windows PC-based system for small-angle scattering data analysis. J. Appl. Crystallogr. *36*, 1277-1282.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605-1612.

Saphire, E.O., Parren, P.W., Pantophlet, R., Zwick, M.B., Morris, G.M., Rudd, P.M., Dwek, R.A., Stanfield, R.L., Burton, D.R., and Wilson, I.A. (2001). Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. Science *293*, 1155-1159.

Schuck, P. (2003). On the analysis of protein self-association by sedimentation velocity analytical ultracentrifugation. Anal. Biochem. *320*, 104-124.

Schuck, P., Perugini, M.A., Gonzales, N.R., Howlett, G.J., and Schubert, D. (2002). Size-distribution analysis of proteins by analytical ultracentrifugation: Strategies and application to model systems. Biophys. J. *82*, 1096-1111.

Shaikh, T.R., Gao, H., Baxter, W.T., Asturias, F.J., Boisset, N., Leith, A., and Frank, J. (2008). SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nat. Protocols *3*, 1941-1974

Svergun, D. (1992). Determination of the regularization parameter in indirect-transform methods using perceptual criteria. J. Appl. Crystallogr. *25*, 495-503.

Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., and Ludtke, S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. *157*, 38-46.

Wiseman, T., Williston, S., Brandts, J.F., and Lin, L.N. (1989). Rapid measurement of binding constants and heats of binding using a new titration calorimeter. Anal. Biochem. *179*, 131-137.