Allosteric activation of preformed EGF receptor dimers by a single ligand binding event

Endang R. Purba
Okinawa Institute of Science and Technology
Ei-ichiro Saita
Okinawa Institute of Science and Technology
Reetesh R. Akhouri
Okinawa Institute of Science and Technology
Lars-Göran Öfverstedt
Okinawa Institute of Science and Technology
Gunnar Wilken
Okinawa Institute of Science and Technology Graduate University https://orcid.org/0000-0002-4019-9320
Ulf Skoglund
Okinawa Institute of Science and Technology
Ichiro N. Maruyama (ichi@oist.jp)
Okinawa Institute of Science and Technology https://orcid.org/0000-0002-5112-1342

Research Article

Keywords: EGFR, RTK, Preformed Dimer in vitro, Cryo-ET, Conformational Transition

Posted Date: June 10th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-182557/v2

Title: Allosteric activation of preformed EGF receptor dimers by a single ligand binding event

Authors: Endang R. Purba[1†], Ei-ichiro Saita[1†], Reetesh R. Akhouri[2], Lars-Goran Öfverstedt[2], Gunnar Wilken[2], Ulf Skoglund[2], Ichiro N. Maruyama[1*]

Affiliations: [1]Information Processing Biology Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa 904-0495, Japan. [2]Cellular Structural Biology Unit, Okinawa Institute of Science and Technology Graduate University, Okinawa 904-0495, Japan.

[†]These authors contributed equally to this work.

[*]Corresponding author. Email: ichi@oist.jp

Abstract: Aberrant activation of the epidermal growth factor receptor (EGFR) by mutations has been implicated in a variety of human cancers. Elucidation of the structure of the full-length receptor is essential to understand the molecular mechanisms underlying its activation. Unlike previously anticipated, here, we report that purified full-length EGFR adopts a homodimeric form *in vitro* before and after ligand binding. Cryo-electron tomography analysis of the purified receptor also showed that the extracellular domains of the receptor dimer, which are conformationally flexible before activation, are stabilized by ligand binding. This conformational flexibility stabilization most likely accompanies rotation of the entire extracellular domain and the transmembrane α-helix, resulting in dissociation of the intracellular kinase dimer and, thus, rearranging it into an active form. Consistently, mutations of amino acid residues at the interface of the inactive, symmetric kinase dimer spontaneously activate the receptor *in vivo*. Optical single-molecule observation also demonstrated that binding of only one ligand activates the receptor dimer on the cell surface. Based on these results, we propose an allosteric model for the activation of EGFR dimers by ligand binding. Our results demonstrate how oncogenic mutations spontaneously activate the receptor and shed light on the development of novel cancer therapies.

Introduction: EGFR, a member of the receptor tyrosine kinase (RTK) family, plays vital roles in many cellular processes, including cell survival, proliferation, differentiation, motility, and metabolism (1, 2). The EGFR signaling pathway is one of the most dysregulated pathways in many human cancers (3, 4). RTKs are all synthesized as type-1 single-pass transmembrane proteins and

bind polypeptide ligands, primarily growth factors. It has been assumed that all RTKs, except for the insulin receptor family, are activated by ligand-induced dimerization of the receptor monomers (5). This model was first proposed for EGFR (also known as ErbB1 or HER1) (6), but it has long been controversial whether the receptor is monomeric, dimeric or oligomeric before ligand binding (7-11).

EGFR, ~170 kDa in mass, consists of an extracellular ligand-binding domain, single transmembrane α-helix, intracellular juxtamembrane (JM) region, cytoplasmic kinase domain (KD), and C-terminal tail (Fig. 1*A*) (12). The C-terminal regulatory tail is autophosphorylated following activation and mediates interactions between the receptor and downstream effector molecules, including the Src homology 2 domain-containing transforming protein-1 (Shc1) and growth factor receptor-bound protein-2 (Grb2) adapters (13-15). Crystallographic studies of the isolated extracellular domain (ECD) and intracellular domain (ICD) of EGFR have provided insights into liganded and unliganded forms of the receptor. ECD contains four subdomains (16-18). Subdomains I (also known as L1) and III (L2) have a β-helix solenoid structure and are responsible for ligand binding by simultaneously contacting the same bound ligand. Subdomains II (CR1) and IV (CR2) are cysteine-rich and interact with each other in the unliganded, tethered form. Ligand binding to subdomains I and III breaks this intramolecular interaction for intermolecular association of two β-hairpins of subdomain II in its liganded, extended form. The crystal structures of symmetric and asymmetric KD dimers have also been determined as the inactive and active forms of the receptor, respectively (19). Nonetheless, little is known about how ligand binding to ECD dictates the active form of the KD dimer. In the present study, we analyzed the structures of full-length EGFR in the absence and presence of bound ligands.

Results and Discussion

EGFR adopts a dimeric structure. Full-length human EGFR tagged with eight histidine residues (His tag) at its C-terminus was expressed in human embryonic kidney (HEK)293T cells and was purified by Ni^{2+} Sepharose column chromatography after solubilization with 19 mM *n*-dodecyl-β-D-maltoside (DDM) (Fig. 1*B*). EGFRΔ998, a deletion mutant whose C-terminal residues 999-1186 were replaced with the His tag, was also purified as a reference. This deletion mutant was monomeric when solubilized with the detergent Triton X-100 and became dimeric upon ligand binding (7). Purified full-length EGFR was observed to be phosphorylated at the basal level when analyzed by immunostaining with anti-phosphotyrosine antibody. Upon stimulation with its ligand, epidermal growth factor (EGF), autophosphorylation of full-length EGFR was enhanced markedly *in vitro* in the presence of ATP (Fig. 1*C*).

When analyzed by gel filtration chromatography, the full-length receptor was eluted as a symmetric peak with a Stokes radius of 9.05 ± 0.34 nm or 9.01 ± 0.46 nm before or after

incubation with EGF, respectively. In contrast, the unliganded EGFRΔ998 receptor was eluted as two peaks with the Stokes radii of 5.68 ± 2.14 nm and 7.91 ± 6.01 nm (Fig. 1*D*). The lower

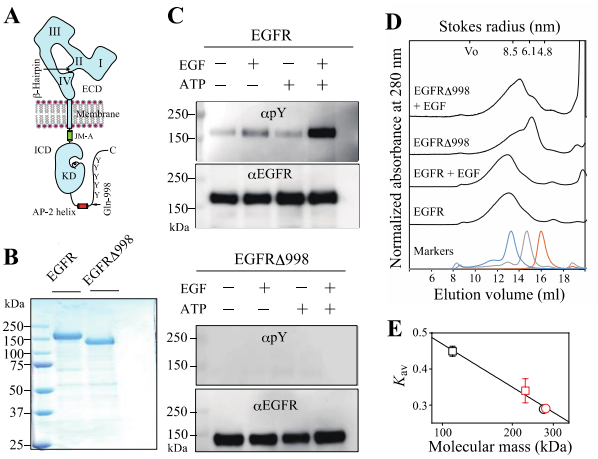


Fig. 1. Purified EGFR adopts a dimeric structure before and after activation. (**A**) Schematic representation of an EGFR monomer. Y, a tyrosine residue as a major phosphorylation site. Not drawn to scale. (**B**) SDS-PAGE of purified full-length EGFR and EGFRΔ998. (**C**) Western blots of purified proteins, which were detected with an antibody specific for EGFR (αEGFR) or phosphotyrosine (αpY). (**D**) Gel filtration chromatograms of purified full-length EGFR and EGFRΔ998 with or without bound EGF. (**E**) Partition coefficient *K*av plots against calculated molecular masses of EGFRΔ998 and full-length EGFR with and without bound EGF. EGFRΔ998 without and with bound EGF and full-length EGFR without and with bound EGF are shown in black square, red square, black circle and red circle, respectively. Data points are means +/- SD.

molecular mass peak shifted to the position of the higher peak after incubation with EGF, indicating that monomeric EGFRΔ998 became its dimeric form upon ligand binding. As shown in Fig. 1*E*, the molecular masses of the full-length unliganded and liganded EGFRs matched with those of the receptor dimer. These results indicated that the full-length EGFR molecules with and without bound EGF adopt homodimeric structures. To exclude the possibility that the His tag contributes to the formation of the dimeric structure, the tag was cleaved from the full-length EGFR by digestion with tobacco etch virus (TEV) endopeptidase (*SI Appendix*, Fig. S1). The Stokes radius of the cleaved EGFR was similar to that of the full-length EGFR with the His tag, indicating that the tag does not contribute to dimer formation. These results demonstrate that prior to ligand binding, full-length EGFR has a homodimeric structure and that the receptor dimer can be activated by ligand binding without changing its dimeric form.

3D density maps of purified EGFR. We collected 15 and 18 tomograms by cryo-electron tomography (Cryo-ET) of the purified full-length EGFR with and without bound EGF, respectively (*SI Appendix*, Table S1). After refinement using the constrained maximum entropy tomography (COMET) software package (20), final three-dimensional (3D) "density maps" of the proteins were calculated from forward-scattered electrons at 300 kV. Within these tomograms, the MINER program of the package was used to extract subtomograms. We reconstructed 474 molecules of full-length unliganded EGFR from the 18 tomograms (*SI Appendix*, Fig. S2). The CORRPAIR program was applied to produce correlation matrices between pairs of all the subtomograms and to classify the 474 subtomograms into 25 clusters (*SI Appendix*, Fig. S3). Using the CORRAVE program, subtomograms of each cluster were averaged to represent the respective cluster (Fig. 2*A*). From the 15 tomograms (*SI Appendix*, Fig. S4), we also reconstructed 557 liganded EGFR molecules activated by EGF binding, which were classified into 25 clusters (*SI Appendix*, Fig. S5). Subtomograms of each cluster were averaged, as shown in Fig. 2*B*.

The 25 averaged density maps of unliganded or liganded EGFRs were analyzed by determining the principal axis of minimum moment of inertia and the molecular center of mass. The mean volumes of slices of both the unliganded and liganded receptors along the principal axis showed two peaks, one of which was larger than the other (Fig. 2*C*). Based on the crystal structures of EGFR ECD and ICD, the results indicated that the large and small peaks correspond to ECD and ICD dimers, respectively, which are separated by the transmembrane α-helices. Furthermore, full lengths and ECD lengths of the liganded receptors were significantly longer than those of the unliganded receptors (Fig. 2*D* and *SI Appendix*, Fig. S6), suggesting that a fraction of the unliganded ECD is likely to take a tethered structure through the interaction of the subdomains II and IV, whereas liganded EGFR ECD has an extended structure (16-18).

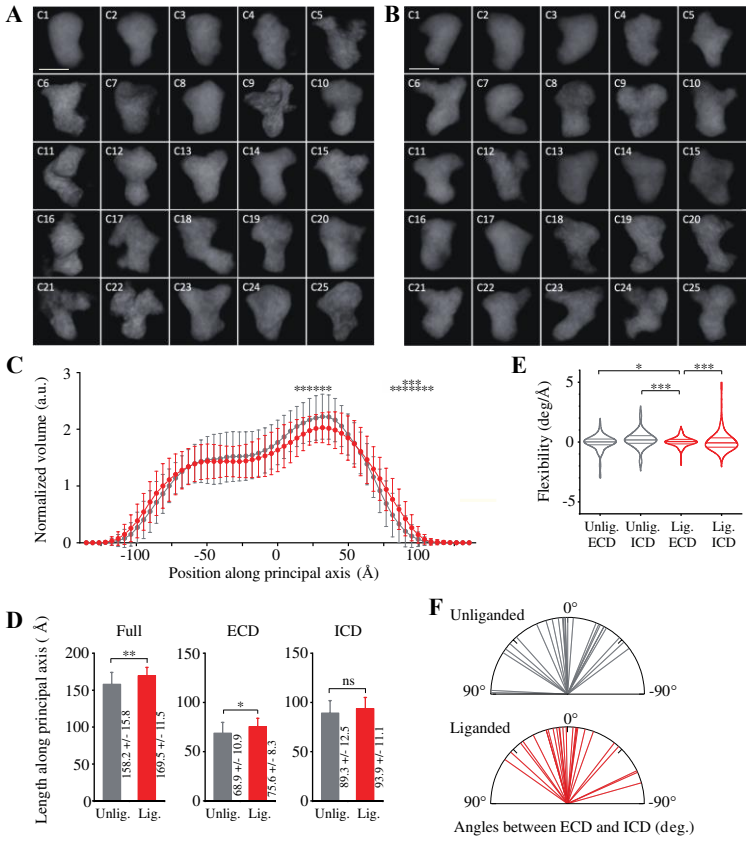


Fig. 2. Conformational variables of EGFR befo re and after activation. (**A** and **B**) Averaged electron density maps of 25 subtomogram clusters of unliganded and liganded EGFRs, respectivel y. (**C**) Plots of averaged volumes of unliganded (gray) and liganded (red) EGFRs along the principal axis with distances from the molecular center of mass. Data points are means +/- SD. Two-sided Student's *t*-test in the range between -140 Å and 75 Å or Mann-Whitney U test in the range over 80 Å along the principal axis (*p < 0.05, **p < 0.01). (**D**) Long-axis lengths of ECD and ICD of unligand - ed and liganded EGFRs were measured as shown in fig. S6. Data points are means +/- SD. Two-sid- ed Student's *t*-test. ns, not significant. (**E**) Flexibility of ECD and ICD of unliganded and liganded EGFRs, which was measured as shown in fig. S7. Lines in the violin plots show 25 th, 50th and 75th percentile values. Asterisks indicate that variances are significantly di fferent (Levene's test; *p < 0.05, ***p < 0.001). (**F**) Angles between long axes of ECD and ICD of unliganded and liganded EGFRs (fig. S8).

When the flexibility of ECD and ICD of unliganded and liganded EGFRs was examined (*SI Appendix*, Fig. S7), it was observed that the ECD dimer was significantly stabilized by ligand binding and had the most rigid structure (Fig. 2*E*). Angles between the long axes, which are perpendicular to the principal axis of minimum moment of inertia, of ICD and ECD of the unliganded EGFRs were variable and ranged from 89° (clockwise) to –54° (counterclockwise) when observing the molecules extracellularly (Fig. 2*F* and *SI Appendix*, Fig. S8). Similarly, angles between the long axes of ICD and ECD of the liganded EGFRs also ranged from 55° to –74°. These rotation angle variabilities were not significantly different between the unliganded and liganded receptors, indicating that the receptor dimers twist flexibly perpendicular to the principal axis before and after ligand binding.

Conformational flexibility transition. Density maps from the 3D reconstruction were sufficient for defining ECD and ICD and their conformational changes, as described above. Using CHIMERA software (21), crystal structures of EGFR domains were manually docked into the envelope of each averaged subtomogram of 25 unliganded and 25 liganded receptor clusters (Fig. 3). Crystal structures of the tethered (PDB ID: 1NQL) (18) or extended (half of 3NJP) (22) ECD monomer were docked into the envelope of unliganded receptors. The ECD dimer (3NJP) (22) was docked into the envelope of liganded EGFRs. The crystal structure (3GT8) (19) of inactive, symmetric KD dimer and the NMR structure (2M0B) (23) of inactive transmembrane α-helices were docked into the envelope of unliganded EGFRs (Fig. 3*A*), whereas the crystal structure (2GS6) (24) of the active, asymmetric KD dimer and the NMR structure (2M20) (25) of active transmembrane α-helices were docked into the envelopes of liganded receptors (Fig. 3*B*).

Then, the 25 docked crystal structures of unliganded or liganded receptors were manually aligned to each other based on two principal axes of minimum and maximum moments of inertia and the molecular center of mass (Fig. 3 *C* and *D*). These alignments show that both ECD and ICD dimers of unliganded and liganded EGFR dimers have flexible structures, which are likely to correspond to large rotation angles between ECD and ICD of the averaged density maps, as shown in Fig. 2*F*. The spontaneous structural transition of unliganded receptor ECD from tethered to extended is also likely to contribute to the flexibility of the domain (26, 27). Furthermore, an NMR study of EGFR in native membranes has shown that ECD of the unliganded receptor is highly dynamic, while ICD is rigid (28). This relative stability of ICD is consistent with its role in the formation of the unliganded receptor dimer, as described below. Incontrast, ECD dimers of the liganded receptor showed the most rigid structures among other domains of unliganded and liganded receptors (Fig. 2*E*). Upon ligand binding, therefore, conformational flexibility transition occurs in ECDs of EGFR dimers from a flexible to rigid structure.

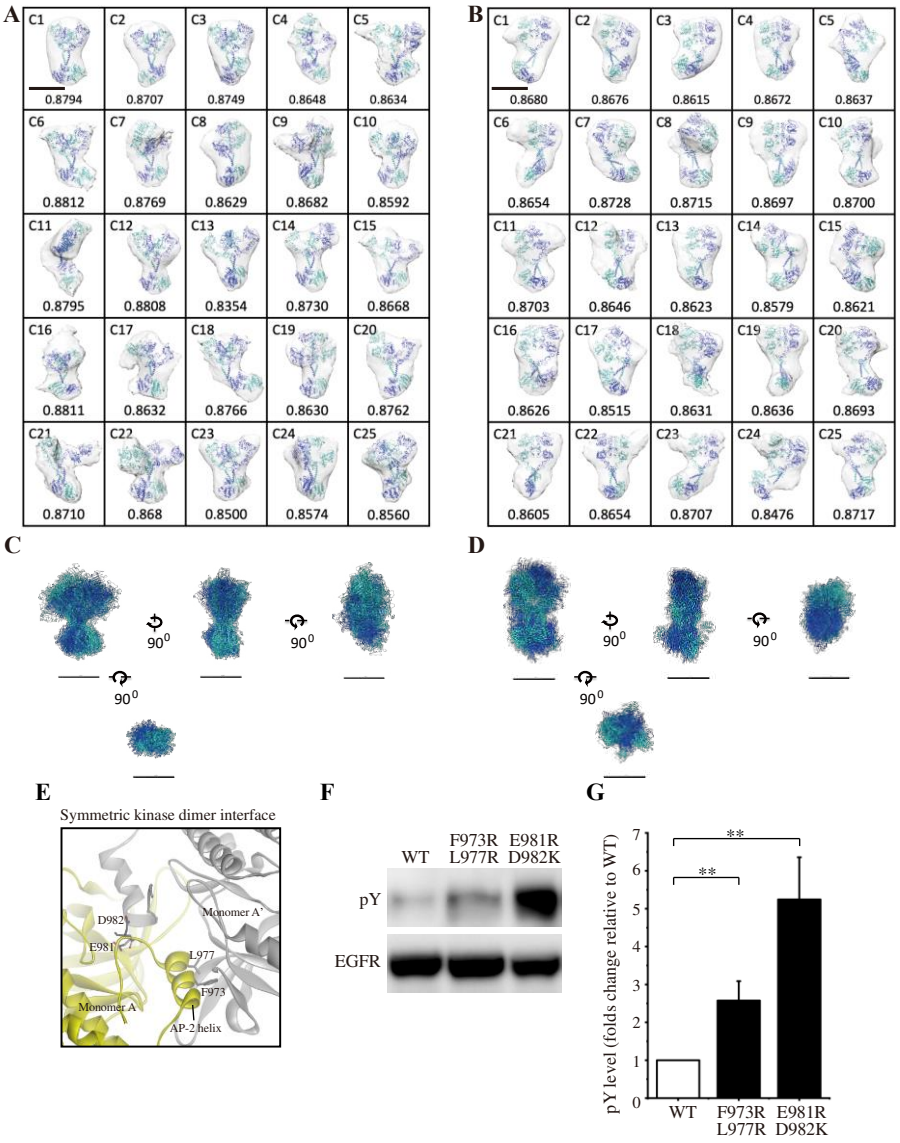


Fig. 3. Activation of EGFR dimers by conform ational flexibility transition or by mutations. (**A** and **B**) Envelopes of averaged density maps of unliganded and liganded EGFRs, respectively, were docked with crystal and NMR structures. Cr oss-correlation coefficients are shown below each map. Scale bar, 10 nm. (**C** and **D**) Alignments of crystal and NMR structures docked into the averaged density maps of unliganded or liganded EGFR, respectively, based on two principal axes of minimum and maximum moments of inertia an d the molecular center of mass. Four orthogonal views are shown, but ECDs are removed from bo ttom views for clarity. Scale bar, 10 nm. (**E**) Crystal interface between protomers of symmetric KD dimer, each of which is shown in gray or yellow. (**F**) Western blots of wild-type EGFR and two mutants with two amino acid substitutions. (**G**) Summary of five independent Western blots. Data points are means +/- SEM. Two-sided Student' s *t*-test (**$p < 0.01$).

Role of ICD in preformed dimers. We also examined whether EGFR adopts in a homodimeric form *in vivo*. Full-length EGFR was expressed in modified HeLa cells that did not express any of the four EGFR family members on the cell surface to prevent heterodimerization with endogenous receptors (*SI Appendix*, Fig. S9). Inactive, symmetric KD dimer is stabilized by the AP-2 helices, which interact with the interfaces of two protomers of the dimer (19) (Figs. 1*A* and 3*E*). The "electrostatic hook", which consists of acidic residues in the turn after the AP-2 helix, also forms ion pairs with residues in the other subunit. When we mutated Phe-973 and Leu-977 of the AP-2 helix, which form hydrophobic interactions with residues of the other protomer, to hydrophilic arginine residues, the mutant receptor was spontaneously autophosphorylated in the absence of bound ligand (Fig. 3 *F* and *G*). When negatively charged Glu-981 and Asp-982 of the "electrostatic hook" were mutated to positively charged arginine and lysine residues, respectively, the mutant receptor was also spontaneously activated. These results indicate that the mutations destabilize the kinase dimer for spontaneous activation, suggesting that ICD plays a major role in the formation of unliganded dimers (29, 30).

Single ligand activates EGFR *in vivo*. Negative cooperativity of interaction between ligand and EGFR has recently been shown (31-33). Therefore, using single-molecule observation, we examined whether a single bound ligand activates EGFR dimers *in vivo* by expressing full-length EGFR and the Shc1 adaptor protein fused with green fluorescent protein (GFP-Shc1) in the modified HeLa cell to prevent potential heterodimerization with other EGFR family members. When fluorescently labeled EGF was incubated with the cell culture, binding of only one EGF molecule to the receptor dimer recruited GFP-Shc1 to the receptor on the cell surface, demonstrating that only one EGF binding induces phosphorylation of EGFR (Fig. 4 *A* and *B* and *SI Appendix*, Fig. S10 and Movie S1). Colocalization of single EGF and Shc1 spots (blue in Fig. 4 *A* and *B*) was faster than that of two EGF spots (pink) or of two EGF and Shc1 spots (gray). Similar results were obtained by expressing Grb2 fused with GFP (Grb2-GFP) in the modified HeLa cell (*SI Appendix*, Fig. S10 *B* and *C*). These results indicate that binding of a single ligand activates the receptor dimer. Interestingly, colocalization of two EGF molecules was not observed when kinase-dead EGFR was expressed on the cell surface (Fig. 4*C*) or the kinase was inhibited by a specific inhibitor (*SI Appendix*, Fig. S10 *D* and *E*). These results indicate that the colocalization of two EGF molecules is due to dimerization of a single-ligand-bound EGFR dimer (tetramerization) after ligand-induced autophosphorylation (30, 34).

Mechanism of activation of EGFR. Our present study demonstrates that, unlike previously assumed, detergent-solubilized full-length EGFR adopts a dimeric form before receptor

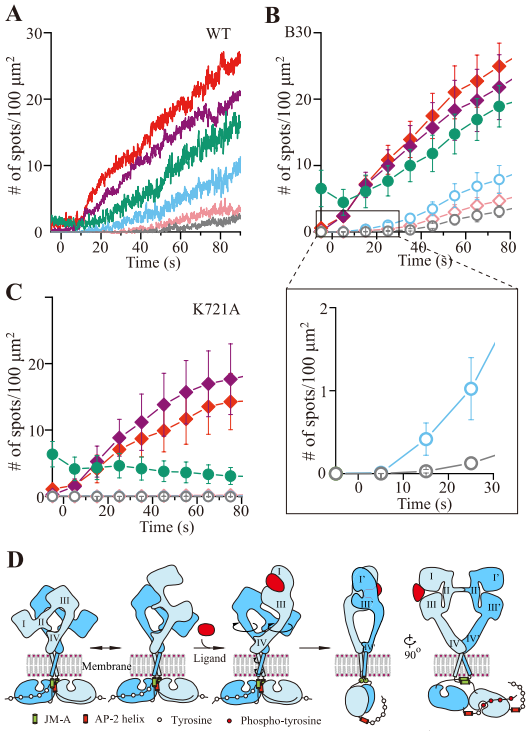


Fig. 4. Single ligand binding activates EGFR dimers. (**A**) Time courses of the number of single spots of Alexa555-labeled EGF (red), Alexa647-labeled EGF (purple) or GFP-Shc1 (green) appeared on the surface of modified HeLa cells exogenously expressing EGFR. Colocalization of GFP-Shc1 and either Alexa555-EGF or Alexa647-EGF, of Alexa555-EGF and Alexa647-EGF, and of all the three is shown in light blue, pink and gray, respectively. (**B**) Same as in (A). Data points are means +/- SEM (n = 4). (**C**) Time courses of the number of the same single spots as in (A), which appeared on the surface of cells exogenously expressing kinase-dead (K721A) EGFR. Data points are means +/- SEM (n = 4). (**D**) Model for the activation of EGFR dimer by ligand binding.

activation *in vitro*. Prior to ligand binding, EGFR adopts a preformed dimeric structure *in vivo* primarily through interactions between ICDs. The transmembrane α-helices (35) and extracellular juxtamembrane domains (18) also seem to contribute to the formation of the unliganded dimer. Our present study also demonstrates that upon ligand binding, the transition from a flexible structure to a stable one occurs in receptor ECD (Fig. 2*E*). Bridging of the extracellular subdomains I and III by the bound ligand brakes tethering between the subdomains II and IV to render the entire ECD a rigid extended structure, wherein β-hairpin of the subdomain II is exposed. This happens concomitantly with an ~90° rotation of the subdomains I/II rigid body about receptor's vertical axis (18). The exposed β-hairpin is likely to interact with β-hairpin of the other subunit to form a rigid ECD dimer, which has lower affinity for the second ligand (negative cooperativity). The formation of the rigid ECD dimer seems to accompany with an ~90° rotation of the rigid subdomains III and IV (18, 22) about its long axis (Fig. 4*D* and *SI Appendix*, Fig. S11). This subdomains III/IV rotation is most likely to accompany with the rotation of the transmembrane α-helices about its long axis since the rotation of extracellular JM segment and transmembrane α-helix in the activation of EGFR has previously been proposed (9) or suggested (35, 36). The rotation of the transmembrane α-helices seems to make the distance between the α-helix C-termini wider (19, 35), and may dissociate the inactive, symmetric kinase dimer and rearranges it to the active, asymmetric kinase dimer. Unlike a monomeric EGFR, an EGFR dimer is able to transmit the outside-in transmembrane signals and inside-out signals through the rotation of the transmembrane α-helices. For example, mutations of the most membrane-proximal portion of ICD abrogate the negative cooperativity of ligand binding to the receptor (31). Many mutations, similar to the mutants shown in the present study, that dissociate the inactive, symmetric ICD dimer are likely to spontaneously activate the receptor, resulting in cancers (37, 38). The similar mechanism may work for the activation of many other cell-surface receptors, such as other RTKs and cytokine receptors (39), and present new opportunities for the development of anti-cancer pharmaceuticals.

Materials and Methods

Plasmid construction. A DNA fragment coding for full-length human EGFR was amplified from the plasmid pNUT/EGFR (9) as a template by PCR using a forward primer, 5'-GGGCTAGCATGCGACCCTCCGGGACG, in which the *Nhe* I site is underlined, and a reverse primer, 5'-GGCTCGAGTCATGCTCCAATAAATTCACTGCTTTG with the *Xho* I site underlined. The resulting PCR product was cloned between the *Nhe* I and *Xho* I sites of a pIRES2-ZsGreen1-Thr-His8 expression vector, a derivate of pIRES2-ZsGreen1 (Clontech) with a thrombin digestion site, LVPRGS, before the His tag. The thrombin digestion site was then replaced with a TEV protease digestion site, ENLYFQG, by inserting an In-Fusion fragment

prepared with two pairs of PCR primers, 5'-CGTCAGATCCGCTAGCATGCGACCCTCCGGGA/3'-**CTGGAAATAGAGGTTCTC**TCCGTTTGTTGCTCCAATAAATTCACT and 5'-**GAGAACCTCTATTTCCAG**GGATCGGATCCGCACCATCACCACCATCACCATCAC/3'-GTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGG, where the *Nhe* I and *Mfe* I restriction enzyme sites are underlined and In-Fusion sites are in bold, using an In-Fusion cloning kit (Takara Bio, Shiga, Japan). After digesting with *Nhe* I and *Mfe* I, the resulting fragment was replaced with the *Nhe* I-*Mfe* I fragment of pIRES2-ZsGreen1-Thr-His8, resulting in pIRES2-ZsGreen1-EGFR-TEV-His8.

cDNA encoding human Shc1was amplified by PCR from pcDNA3.1His p66Shc1 (Plasmid #32574; addgene), with a pair of oligonucleotide primers, 5'-CACCAAGCTTATGAACAAGCTGAGTGGAGGCG and 5'-AACCGCGGCAGTTTCCGCTCCACAGGTTGC, wherein the *Hin*dIII and *Sac* II sites are underlined, respectively. The resulting PCR product was cloned into a pAcGFP1-N1 vector (#632469; Clontech) digested with restriction enzymes *Hin*dIII and *Sac* II to make pAC-N1-GFP-Shc1, in which Shc1 was fused to the C-terminus of AcGFP1. To generate pAc-Grb2-GFP, human *GRB2* cDNA was amplified by PCR using a pair of oligonucleotide primers, 5'-CACCAAGCTTATGGAAGCCATCGCCAAATATG (forward) and 5'-AACCGCGGGACGTTCCGGTTCACGGGGGTG (reverse), from a cDNA library of HEK293T cells and then cloned into pAcGFP1-N1 in which human GRB2 was fused to the N-terminus of AcGFP1. To construct pIRES2-EGFR-GFP-Shc1, a cDNA fragment encoding full-length *EGFR* was amplified from pNUT-EGFR (9), and AcGFP1 fused Shc1 fragment was amplified from pAC-N1-GFP-Shc1. The amplified cDNAs encoding *EGFR* and *AcGFP1-Shc1* were transferred to upstream and downstream of the internal ribosome entry site (IRES) sequence of pIRES2-ZsGreen (Clontech), respectively. To construct pIRES2-EGFR-Grb2-AcGFP1, a DNA fragment encoding *Grb2-AcGFP1* was amplified from pAc-Grb2-AcGFP1 and then replaced the AcGFP1-Shc1fragment of IRES of pIRES2-EGFR-AcGFP1-Shc1. To construct pcDNA3.1-EGFR, full-length *EGFR* was amplified from pIRES2-EGFR-AcGFP1-Shc1, and transferred to pcDNA3.1 (Invitrogen). Two double mutations, F973R/L977R and E981R/D982K, were introduced to pcDNA3.1-EGFR by In-Fusion using oligonucleotide primers encoding the mutations, 5'-CCAACCGGTACCGTGCCCGGATGGATGAAG/5'-CACGGTACCGGTTGGAGTCTGTAG and 5'-GATGAACGAAAGATGGACGACGTGGTGGATGCCGAC/5'-CATCTTTCGTTCATCCATCAGGGCACGGTAGAAGTT, wherein the mutation sites are underlined, respectively. To construct pIRES2-EGFR(K721A)-AcGFP1-Shc1, a mutation, K721A, was introduced to pIRES2-EGFR-AcGFP1-Shc1 by In-Fusion using oligonucleotide

primers, 5'-CGCTATCGCAGAATTAAGAGAAGCAAC and 5'-AATTCTGCGATAGCGACGGGAATTTTAAC, in which the mutation sites are underlined.

Protein expression and purification. Full-length EGFR was expressed by transforming HEK293T cells (American Type Culture Collection) in a 15-cm Petri dish with a mixture of 20 μg pIRES2-ZsGreen1-EGFR-TEV-His8 and polyethylenimine (PEI; Polysciences, Warrington, PA) at a 1:3 (w/w) ratio of DNA to PEI. HEK293T cells were cultured in Dulbecco's modified Eagle medium (DMEM; Gibco) supplemented with 10% (v/v) fetal bovine serum (FBS; Invitrogen) and 2 mM glutamine. This culture medium was replaced with FBS-free DMEM, 5 h prior to transformation. Cells derived from 24 dishes (15 cm in diameter) were harvested 48 h after the transformation and washed once with phosphate-buffered saline (PBS, pH 7.4). Approximately 9×10^8 cells freshly harvested or cells stored at -80°C were incubated for 3 h in 20 ml of bursting buffer, consisting of 20 mM HEPES (pH 7.4), 0.1 mM EDTA, 0.1 mM EGTA, 5.0 mM $MnCl_2$, and an EDTA-free protease inhibitor cocktail (1.0×; Nacalai Tesque, Kyoto, Japan). Cells were disrupted with a Dounce homogenizer on ice and centrifuged at 50,000 ×*g* in a micro ultracentrifuge (model CS150GXL; Hitachi Koki, Tokyo, Japan) for 30 min at 4°C. Pellets were suspended in 20 ml of solubilization buffer, consisting of bursting buffer supplemented with 19.6 mM DDM (Thermo Fisher Scientific) for 3 h and were cleared by centrifugation at 150,000 ×*g* for 30 min at 4°C. The supernatant was then mixed with 1.0 ml of Ni Sepharose resin (GE Healthcare), which was equilibrated with solubilization buffer, on an orbital shaker for 1.0 h in a cold room. The suspension was then collected in a Poly-Prep affinity chromatography column (9 cm in height, 10 ml reservoir volume, and 2 ml bed volume; Bio-Rad) by gravity, and washed with 20-times bed volume of washing buffer (20 mM Tris-HCl, pH 8.0; 400 mM NaCl; and 0.2 mM DDM) supplemented with 20 mM imidazole. The column was then washed with the same volume of washing buffer supplemented with 30 mM imidazole, followed by washing with the same volume of washing buffer supplemented with 40 mM imidazole. EGFR was eluted with two-times the bed volume of washing buffer supplemented with 250 mM imidazole, and the eluate was fractionated into six fractions of 500 μl each. To determine fractions containing EGFR, 15 μl of each fraction was analyzed by polyacrylamide gel electrophoresis (PAGE) using precast 10% Extra PAGE gels (Nacalai Tesque). Collected fractions, which contained 2-4 μM EGFR (~1.5 ml in total), were dialyzed overnight against 1.0 liter of dialysis buffer (20 mM Tris-HCl, pH 8.0; 200 mM NaCl; and 0.2 mM DDM). Dialyzed EGFR was immediately subjected to SDS-PAGE, Cryo-ET, gel filtration, and phosphorylation assays.

Purified full-length EGFR and EGFRΔ998, 15 μl each, were mixed with the same volume of 2 x Laemmli sample buffer (Bio-Rad) containing 5% (v/v) β-mercaptoethanol and then heated at

95°C for 5 min. The samples were separated by using precast 10% Extra PAGE gel with running buffer (25 mM Tris-HCl, pH 8.6; 192 mM glycine; 0.1% (w/v) SDS) at room temperature. Proteins on the gel were fixed in solution [40% (v/v) methanol and 10% (v/v) acetic acid in water] for 15 min at room temperature and then stained with 0.25% (w/v) Coomassie Brilliant Blue G-250 (Nacalai Tesque) in acidic methanol [45% (v/v) methanol and 10% (v/v) acetic acid in water] for 30 min at room temperature. The gel was destained with aqueous 10% (v/v) acetic acid until visible bands appeared.

Gel electrophoresis and phosphorylation assay. In PAGE analysis (Fig. 1*B*), purified EGFR (10 μg in 15 μl of dialysis buffer) was subjected to sodium dodecyl sulfate (SDS)-PAGE analysis as described above. For Western blotting after gel electrophoresis, proteins on the gel were transferred to polyvinylidene difluoride (PVDF) membrane (pore size, 0.45 μm; GE Healthcare) using a TurboBlot dry blotting system (Bio-Rad) and were then immunostained with antibodies, Ab15 (clone H9B4; Thermo Fisher Scientific) for EGFR and 4G10 (Millipore) for phosphorylated EGFR. Horse radish peroxidase (HRP)-conjugated anti-mouse IgG antibody (GE Healthcare) was used as a secondary antibody for the detection of EGFR bound with the first antibody.

For the phosphorylation assay, an aliquot (15 μl of 10 μg/μl) of purified EGFR was reacted at 30°C for 15 min with or without 100 ng EGF in the presence or absence of 1.0 mM ATP in phosphorylation buffer containing 25 mM HEPES (pH 7.4), 20 mM $MgCl_2$, 5 mM β-glycerophosphate, 0.5 mM DTT, and 0.1 mM $NaVO_3$. The reaction was stopped by adding the same volume of two-fold concentrated Laemmli sample buffer and was analyzed by SDS-PAGE and Western blotting as described above. A representative result from three independent experiments is shown in Fig. 1*C*.

For the autophosphorylation analysis of doubly mutated EGFRs (Fig. 3 *F* and *G*), F973R/L977R and E981R/D982K, plasmid constructs, pcDNA3.1-EGFR, pcDNA3.1-EGFR(F973R/L977R) and pcDNA3.1-EGFR(E981R/D982K), were transfected to the modified HeLa cells, which were seeded in a Coster® 6-well plate (Corning, NY) at a density of 1 x 10^5 cells/well in growth media (DMEM supplemented with 10% (v/v) FBS) one day before transfection. Next day, the media in the wells were replaced with 2 ml of fresh media. A transfection mixture, containing 1.5 μg plasmid DNA and 3 μg PEI in 200 μl Opti-MEM (Gibco), was incubated at room temperature for 10 min and was then added to each well. The plate was incubated for 4 h at 37°C in a 5% CO_2 atmosphere. After the incubation, media in wells were replaced with 2 ml of fresh growth media, and the plate was incubated at 37°C in a 5% CO_2 atmosphere. After incubation for 26 h, wells were washed three times with 2 ml of fresh growth media and were covered with 2 ml of Dulbecco's MEM without FBS for serum

starvation, followed by incubation at 37°C for 14 h in a 5% CO_2 atmosphere. After starvation, the plate was placed on ice for 10 min, and washed twice with 2 ml of ice-cold DPBS. Cells in each well were lysed by adding 70 µl Laemmli buffer containing 5% (v/v) β-mercaptoethanol, 1.0 mM Na_3VO_4, a phosphatase inhibitor cocktail (PhosSTOP; Sigma) and a protease inhibitor cocktail (cOmplete EDTA free; Sigma). An aliquot, 15 µg of total proteins, of the lysed cells was incubated at 95°C for 7 min and was subjected to SDS-PAGE analysis, which was performed using 7% (w/v) acrylamide gels in running buffer [25 mM Tris-HCl (pH 8.6), 192 mM glycine, 0.1% (w/v) SDS]. Proteins on the gel were transferred to PVDF by electrophoresis using a Mini Trans-Blot cell (Bio-Rad) at 30 V for 16 h in a cold room. The membrane was probed with primary antibodies, mouse anti-phosphotyrosine monoclonal (pY-20, 1/500 dilution; Santa Cruz Biotechnology) and rabbit anti-EGFR monoclonal (D38B1, 1/3000 dilution; Cell Signaling Technology), and then with secondary antibodies, HRP-conjugated anti-mouse IgG (dilution, 1/3000; GE Healthcare) and HRP-conjugated anti-rabbit IgG (dilution, 1/3000; GE Healthcare), respectively. This autophosphorylation analysis was repeated five times and a representative result is shown in Fig. 3*F*.

Gel filtration chromatography. Full-length EGFR, 2 µM at a final concentration, in dialysis buffer was reacted with or without 20 µM EGF (Recombinant human; Abbiotec, Escondido, CA) for 30 min on ice in the presence of 1.0 mM ATP in 1.0 ml of total reaction volume, and an aliquot (500 µl) was analyzed by gel filtration column chromatography at a flow rate of 0.4 ml/min (24 ml bed volume, 10 mm inner diameter, 300 mm in height; prepacked with Superose 6 Increase; GE Healthcare), which was equilibrated with running buffer (20 mM Tris-HCl, pH 8.0; 200 mM NaCl; and 0.2 mM DDM), by fast protein liquid chromatography (FPLC) using an AKTA-Explorer (GE Healthcare) in a cold room.

Purified EGFR-TEV-His8 (0.15 mg) was digested with 0.05 mg (250 unit) of TEV protease (Accelagen, San Diego, CA) in 0.6 ml of buffer, which contains 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, and 0.2 mM DDM, for 18 h at 4°C. An aliquot (500 µl) was subjected to analysis by gel filtration chromatography using a column prepacked with Superose 6 Increase 10/300 GL as described above. A representative result from three independent experiments is shown in Fig. 1*D*.

Apparent Stokes radii of EGFRΔ998 and full-length EGFR in the presence or absence of bound EGF were determined using the following proteins as standard markers in gel filtration chromatography: bovine thyroid thyroglobulin (Mw, 669 kDa; Stokes radius, 8.5 nm; GE Healthcare), horse spleen ferritin (440 kDa, 6.1 nm; GE Healthcare), and rabbit muscle aldolase (158 kDa, 4.8 nm; GE Healthcare). The partition coefficient K_{av} was calculated using the equation $K_{av} = (V_e - V_o)/(V_t - V_o)$, where V_e, V_t and V_o are the elution volume, total volume, and

void volume, respectively. For K_{av} plots against calculated molecular masses, EGFRΔ998 without bound EGF and others were assumed to be monomeric and dimeric, respectively. The number of EGF bound was assumed to be one molecule per receptor dimer, as discussed in the main text.

Cryo-ET. Colloidal gold particles (10 nm; Amersham Biosciences), 2 µl of a "1.0 OD" solution, were equilibrated with purification buffer (20 mM Tris-HCl, pH 8.0; 200 mM NaCl; and 0.2 mM DDM). Purified EGFR (20 µl, 0.2 mg/ml) before or after incubation with 20 µM EGF at a final concentration on ice for 30 min was mixed with the equilibrated colloidal gold particles at 3:1 ratio for alignment purposes. A Quantifoil holey carbon copper grid (R 1.2/1.3; Electron Microscopy Sciences, Hatfield, PA) was glow-discharged for 60 s, and the EGFR and colloidal gold mixture (3 µl) was spotted on the grid. The grid was blotted with filter paper (grade 595; Ted Pella, Redding, CA) for 3 s and was vitrified at 4°C with 80%-90% humidity (40) using a Vitrobot Mark IV plunge-freezing device (FEI). An FEI Titan Krios equipped with a Falcon II direct electron detector was operated at 300 kV (accelerating voltage) and at a magnification of 37,000 with underfocus values between 1.0 µm and 1.5 µm for data collection (resulting in a pixel size of 2.258 on the specimen scale). Specimens were tilted from 0° to –70° and 0° to +70° with an increment of 1° tilt/image. The total dose for each tilt series did not exceed 90 $e^-/Å^2$ to minimize radiation damage. Tomography software (version 4.0; Thermo Fisher Scientific) was used for data acquisition.

Image processing and docking. Unliganded and liganded EGFR tilt-series were aligned using 10-nm gold particles as fiducial markers with mean errors of 3.5 Å (1.55 pixels) and 3.7 Å (1.6 pixels), respectively. A series of 2D slices perpendicular to the tilt axis were reconstructed by using a radius-weighted back-projection algorithm from a stack of extracted areas from the tilt images. This stack of 2D slices constituted the initial 3D map and was also called a back-projection map or a tomogram. The back-projections were run on multiple selected areas, generating volumes of 800 × 800 × 800 voxels. The tomograms were further improved via a regularizing process by COMET (20, 41, 42), version 6.4.2, which enhances the contrast of density to increase the signal-to-noise in the final tomograms. The MINER program of the package was used to identify coordinates for desired maps in the regularized tomograms using either a 3D voxel or a molecular mass range as a parameter. From the list of coordinates, an automated extraction of subtomograms was performed in a volume of 110 × 110 × 110 voxels that were already low-pass filtered to 15 Å. Individual snapshots of the extracted subtomograms were generated using BOB software (43) with a volume-rendering option. The CORRPAIR program of the package was applied to the individual 3D subtomograms to create a correlation

matrix pairing all the subtomograms. The correlation matrix contained clusters of subtomograms that had similar correlations. The CORRAVE program of the package was then run to maximize the correlations within each cluster and to generate an averaged subtomogram from each cluster. CHIMERA (21), version 1.14.0 (http://www.cgl.ucsf.edu/chimera/), was subsequently used for docking the model of EGFR domains from PDB entries to the averaged subtomograms of the clusters. The crystal structures of EGFR domains were aligned to the individual subtomogram maps by using CHIMERA with the manual option. The PDB entries used for docking were 1NQL (18) 3NJP (22), 2M0B (23), 2M20 (25), 3GT8 (19), and 2GS6 (24). These representative clusters were cross-correlated with the 20 Å-resolution low-pass filtered crystal structures of EGFR domains.

Electron density distribution along the axis of minimum moment of inertia. Electron density maps were represented by a set of vectors (x_i, y_i, z_i, and d_i) ($1 \leq i \leq N$), where x_i, y_i, and z_i are the coordinates of each voxel, and d_i is the density of each voxel. N is the number of voxels in each density map. The coordinates (x_i, y_i, z_i) of each density map were transformed to fit the center of mass of the density map to the origin of the coordinate system (x'_i, y'_i, z'_i), using equation (1).

$$(x'_i, y'_i, z'_i) = \left(x_i - \frac{\sum_i x_i d_i}{\sum_i d_i}, y_i - \frac{\sum_i y_i d_i}{\sum_i d_i}, z_i - \frac{\sum_i z_i d_i}{\sum_i d_i}\right) \quad (1)$$

The inertia tensor I of the electron density map was calculated using equation (2).

$$\mathbf{I} = \begin{bmatrix} \sum_i d_i(y_i'^2 + z_i'^2) & -\sum_i d_i x'_i y'_i & -\sum_i d_i x'_i z'_i \\ -\sum_i d_i y'_i x'_i & \sum_i d_i(z_i'^2 + x_i'^2) & -\sum_i d_i y'_i z'_i \\ -\sum_i d_i z'_i x'_i & -\sum_i d_i z'_i y'_i & \sum_i d_i(x_i'^2 + y_i'^2) \end{bmatrix} \quad (2)$$

The axis of minimum or maximum moment of inertia was determined from an eigenvector that give a minimum or maximum eigenvalue, λ, respectively, in the following equation:

$$\mathbf{I} \cdot \mathbf{n} = \lambda \mathbf{n},$$

where n is a 3D unit vector. The coordinates of each density map, (x'_i, y'_i, z'_i), were rotated around the center of mass so that the eigenvectors of the minimum and maximum moments of inertia can be aligned to the z- and x-axes, respectively. Among the two possible opposite alignments along the z-axis, one of the two alignments was chosen based on the docked models of crystal structures (Fig. 3 *A* and *B*). The data processing above was performed on R (https://www.r-project.org/) with custom scripts.

Each aligned density map was sectioned into two voxels (~4.5 Å) each along the z-axis, and the electron density within each section was plotted against the z-axis. The mean values of the plots of 25 unliganded or 25 liganded EGFR density maps are shown in Fig. 2C, after normalized by total volumes. Error bars indicate standard deviation. Asterisks indicate significant

differences (*p < 0.05, **p < 0.01) of the two groups with two-sided Student's t-test (in the range between -140 Å and 75 Å), or Mann-Whitney U-test (in the range over 80 Å) using R.

HeLa cells lacking EGFR family members. HeLa cells (RIKEN BRC, Saitama, Japan) were cultured in DMEM supplemented with 10% (v/v) FBS and incubated in a humidified incubator containing 5% CO_2 at 37°C. To knockout EGFR family members in HeLa cells, a Cas9-RNA complex transfection system (Alt-R CRISPR-Cas9; Integrated DNA Technologies, Coralville, IA) was used. A custom-made guide RNA was complexed with Cas9 protein, and the resulting RNA and protein complex was transfected to HeLa cells with Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to the manufacturer's instructions. After three days, the transfected cells were transferred to 96-well plates for the isolation of single cells. The isolated cells were further cultured for 1–2 weeks, and their genomes were analyzed by PCR for deletion, which was then confirmed by DNA sequencing. To knockout of multiple EGFR family members, the above knockout procedure was repeated three times to create a triple-knockout cell line lacking EGFR, HER2, and HER4. The target 20-nucleotide genome sequences of guide RNA for EGFR, HER2, and HER4 are 5'-AGGGTTGTTGCTGAACCGCA in exon #4, 5'-TGAGTCCATGCCCAATCCCG in exon #7, and 5'-TGCTGCCATCGAGAATGTGC in exon #6, respectively. All genome deletions introduced to the cell line created stop codons within the extracellular domain regions of each receptor. As HER3 is not expressed on the HeLa cell surface (44), we used a HeLa cell line lacking EGFR, HER2, and HER4, which was confirmed by Western blot analysis and ligand binding observed at the single-molecule level described below (*SI Appendix*, Fig. S9).

Single-molecule observation. To label EGF with fluorescent dyes, 1.0 mg/ml EGF (PeproTech, Cranbury, NJ) solution in water was mixed with the same volume of 100 mM potassium phosphate buffer, pH 7.0. Alexa Fluor 555 NHS (Thermo Fisher Scientific) or Alexa Fluor 647 NHS in dimethyl sulfoxide was added to the EGF solution to a final concentration of 400 μM. The reaction mixture was incubated at room temperature for 70 min and then loaded onto a PD-10 desalting column (GE Healthcare) to remove unbound dye molecules. Eluates from the column were concentrated with a centrifugal filter device (Amicon Ultra 3K; Millipore). Concentrations of EGF and fluorescent dyes in the concentrated samples were determined using a spectrophotometer (NanoDrop; Thermo Fisher Scientific) based on molecular extinction coefficients, 18000 $M^{-1}cm^{-1}$ at 280 nm for EGF, 150000 $M^{-1}cm^{-1}$ at 550 nm for Alexa555, and 239000 $M^{-1}cm^{-1}$ at 650 nm for Alexa647. We used only EGF samples with higher labelling efficiency than 95%.

To prepare a trolox and troloxquinone mixture (TXTQ), 2.5 mg/ml trolox (Cayman Chemical, Ann Arbor, MI) was dissolved in 10 mM potassium phosphate, pH 7.2. After dissolving, the pH of the solution was adjusted to pH 7.0 using 1.0 M NaOH. Approximately 10% of trolox (~0.25 mg/ml) was converted to troloxquinone by oxidization under illumination using a mercury lump (Olympus) on a stereo microscope (SZX16; Olympus). Generation of troloxquinone was monitored by measuring the absorbance at 255 nm using a spectrophotometer. Oxidization was continued until the absorbance at 255 nm reached ~1.2 in 0.1 mm path length (45).

To construct cell lines co-expressing EGFR and GFP-tagged Shc1 or EGFR and GFP-tagged Grb2, a plasmid construct, pIRES2-EGFR-AcGFP1-Shc1, pIRES2-EGFR-Grb2-AcGFP1, or pIRES2-EGFR(K721A)-AcGFP1-Shc1, was transfected into the modified HeLa cell line that did not express EGFR, HER2, HER3, or HER4 on the cell surface. Cells (0.5×10^3) were seeded in a glass-bottom dish (35 mm in diameter; Iwaki, Shizuoka, Japan) one day before transfection, and the next day, medium was replaced with 0.3 ml of fresh growth medium, DMEM supplemented with 10% (v/v) FBS. Transfection mixture, which contained 0.3 μg plasmid DNA and 0.6 μg of PEI in 30 μl of Opti-MEM (Gibco), was incubated at room temperature for 10 min and was then added to each dish. The dish was incubated at 37°C for 4 h under a 5% CO_2 atmosphere. After incubation, growth media in the wells were replaced with fresh media, and the plates were further incubated at 37°C for 40 h in a 5% CO_2 atmosphere. Then, the dish was washed three times with 0.3 ml of DMEM without FBS and filled with 0.3 ml of DMEM without FBS for serum starvation. The dish was further incubated at 37°C for more than 3 h in a 5% CO_2 atmosphere.

Single-molecule observation of EGF binding to the cell surface of the modified HeLa cell was performed using an inverted microscope (Eclipse Ti; Nikon) with an oil-immersion objective (SR APO TIRF ×100/1.49; Nikon) at room temperature. GFP, Alexa555, and Alexa647 were excited by a laser unit (LU-N4; Nikon) with 488 nm, 561 nm, and 640 nm, respectively. The fluorescent signal was split into three EM CCD cameras (DU-897; Andor Technology, Belfast, UK) using dichroic mirrors (FF580-FDi01 and FF662-FDi01; Semrock, Rochester, NY) and bandpass filters (FF01-525/45, FF01-600/37, and FF01-692/40; Semrock). Before observation, cells were washed twice with Hank's balanced salt solution (HBSS; Gibco) and supplemented with 200 μl HBSS containing 1.0 mM TXTQ. The apical surface of the cell was observed under oblique illumination (46). At 10 s after video recording started, 200 μl HBSS containing 1.0 mM TXTQ, 0.5 mg/ml glucose oxidase, 0.04 mg/ml catalase, 1.0 mg/ml glucose, and fluorescently labeled EGF was applied to the dish. Movements of fluorescent spots derived from fluorescently labeled EGF, GFP-Shc1 or Grb2-GFP on the cell surface were recorded at 10 frames/s for 90 s after EGF stimulation of the cell.

The images of three-color channels were shifted and distorted, primarily due to chromatic aberration. We used an ImageJ plugin (DoM_Utrecht, Netherlands) to correct the aberration. The GFP and Alexa647 channels were corrected to fit the Alexa555 channel. Fluorescent bead images of a calibration slide (Tool for calibration Multi Spec #1783-455; Zeiss) were used as references for correction. After correction, fluorescent spots in each channel were detected by an ImageJ plugin (MosaicSuite; MOSAIC Group, Towson, MD) and analyzed using a custom R script. We defined colocalization of spots in different channels when the distance between the spots was less than $\sqrt{2}$ pixel (0.21 μm). The fluorescent spots in different channels sometimes approached each other within the threshold distance by chance without actual colocalization. The number of this "pseudo-colocalization" in different channels was estimated by the colocalization analysis, where one of two images was flipped vertically and horizontally. The mean of two numbers of colocalization with vertically and horizontally flipped images was used as the number of the pseudo-colocalization for every frame and subtracted from the number of colocalization of the original unmodified two images to obtain the corrected numbers of colocalization shown in Fig. 4 and *SI Appendix*, Fig. S10 for every pair of different channel images. In case of the number of colocalizations of all three channels, a colocalized image of Alexa555–EGF and Alexa647–EGF was used for the colocalization analysis with the remaining Shc1 (or Grb2) channel image in the same way.

Statistical analysis. Statistical analysis of data was performed using R (version 3.6.3) or SigmaPlot (version 13.0). All data were checked for normality of distribution and homogeneity of variance using χ^2 goodness of fit test ($p < 0.05$) and were evaluated using two-sided Student's *t*-test for comparisons between pairs of groups. If normality did not hold, Mann-Whitney U test or Levine's test was used. Results are reported as mean ± standard deviation (SD) or standard error of the mean (SEM). Asterisks in figures indicate significant differences (*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$) of two groups.

References

1. P. J. Miettinen, *et al*. Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor. *Nature* **376**, 337-341 (1995).
2. J. C. Chen, *et al*. Expression and function of the epidermal growth factor receptor in physiology and disease. *Physiol. Rev.* **96**, 1025-1069 (2016).
3. Z. An, O. Aksoy, T. Zheng, Q.-W. Fan, W. A. Weiss, Epidermal growth factor receptor and EGFRvIII in glioblastoma: signaling pathways and targeted therapies. *Oncogene* **37**, 1561-1575 (2018).
4. M. London, E. Gallo, Epidermal growth factor receptor (EGFR) involvement in epithelial-

derived cancers and its current antibody-based immunotherapies. *Cell Biol. Int.* **44**, 1267-1282 (2020).

5. M. A. Lemmon, J. Schlessinger, Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117-1134 (2010).
6. Y. Yarden, J. Schlessinger, Self-phosphorylation of epidermal growth factor receptor: evidence for a model of intermolecular allosteric activation. *Biochemistry* **26**, 1434-1442 (1987).
7. L.-Z. Mi, *et al*. Simultaneous visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor. *Nat. Struct. Mol. Biol.* **18**, 984-989 (2011).
8. Y. Huang, *et al*. Molecular basis for multimerization in the activation of the epidermal growth factor receptor. *eLife* **5**, e14107 (2016). doi: 10.7554/eLife.14107
9. T. Moriki, H. Maruyama, I. N. Maruyama, Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. *J. Mol. Biol.* **311**, 1011-1026 (2001).
10. P. Liu, *et al*. Investigation of the dimerization of proteins from the epidermal growth factor receptor family by single wavelength fluorescence cross-correlation spectroscopy. *Biophys. J.* **93**, 684-698 (2007).
11. P. O. Byrne, K. Hristova, D. J. Leahy, EGFR forms ligand-independent oligomers that are distinct from the active state. *J. Biol. Chem*. **295**, 13353-13362 (2020).
12. A. Ullrich, *et al*. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature* **309**, 418-425 (1984).
13. J. Downward, P. Parker, M. D. Waterfield. Autophosphorylation sites on the epidermal growth factor receptor. *Nature* **311**, 483-485 (1984).
14. A. M. Honegger, R. M. Kris, A. Ullrich, J. Schlessinger. Evidence that autophosphorylation of solubilized receptors for epidermal growth factor is mediated by intermolecular cross-phosphorylation. *Proc. Natl Acad. Sci. USA* **86**, 925-929 (1989).
15. K. J. Wilson, J. L. Gilmore, J. Foley, M. A. Lemmon, D. J. Riese, II. Functional selectivity of EGF family peptide growth factors: implications for cancer. *Pharmacol. Ther*. **122**, 1-8 (2009).
16. T. P. J. Garrett, *et al*. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor α. *Cell* **110**, 763-773 (2002).
17. H. Ogiso, *et al*. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110**, 775-787 (2002).
18. K. M. Ferguson, *et al*. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol. Cell* **11**, 507-517 (2003).
19. N. Jura, *et al.* Mechanism for activation of the EGF receptor catalytic domain by the

juxtamembrane segment. *Cell* **137**, 1293-1307 (2009).

20. U. Skoglund, L.-G. Öfverstedt, R. M. Burnett, G. Bricogne, Maximum-entropy three-dimensional reconstruction with deconvolution of the contrast transfer function: A test application with adenovirus. *J. Struct. Biol*. **117**, 173-188 (1996).
21. E. F. Pettersen, *et al*. UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem*. **25**, 1605-1612 (2004).
22. C. Lu, *et al.* Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal growth factor receptor. *Mol. Cell. Biol.* **30**, 5432-5443 (2010).
23. E. V. Bocharov, *et al.* Alternative packing of EGFR transmembrane domain suggests that protein–lipid interactions underlie signal conduction across membrane. *Biochim. Biophys. Acta* **1858**, 1254-1261 (2016).
24. X. Zhang, J. Gureasko, K. Shen, P. A. Cole, J. Kuriyan. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* **125**, 1137-1149 (2006).
25. N. F. Endres, *et al.* Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543-556 (2013).
26. N. J. Bessman, A. Bagchi, K. M. Ferguson, M. A. Lemmon. Complex relationship between ligand binding and dimerization in the epidermal growth factor receptor. *Cell Rep*. **9**, 1306-1317 (2014).
27. N. Kozer, *et al*. Evidence for extended YFP-EGFR dimers in the absence of ligand on the surface of living cells. *Phys. Biol*. **8**, 066002 (12pp) (2011). doi: 10.1088/1478-3975/8/6/066002
28. M. Kaplan, *et al*. EGFR dynamics change during activation in native membranes as revealed by NMR. *Cell* **167**, 1241-1251 (2016). doi: 10.1016/j.cell.2016.10.038
29. X. Yu, K. D. Sharma, T. Takahashi, R. Iwamoto, E. Mekada, Ligand-independent dimer formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced EGFR signaling. *Mol. Biol. Cell* **13**, 2547-2557 (2002).
30. R.-H. Tao, I. N. Maruyama, All EGF(ErbB) receptors have preformed homo- and heterodimeric structures in living cells. *J. Cell Sci.* **121**, 3207-3217 (2008).
31. S. Adak, K. S. Yang, J. Macdonald-Obermann, L. J. Pike, The membrane-proximal intracellular domain of the epidermal growth factor receptor underlies negative cooperativity in ligand binding. *J. Biol. Chem.* **286**, 45146-45155 (2011).
32. D. Alvarado, D. E. Klein, M. A. Lemmon, Structural basis for negative cooperativity in growth factor binding to an EGF receptor. *Cell* **142**, 568-579 (2010).
33. P. Liu, *et al*. A single ligand is sufficient to activate EGFR dimers. *Proc. Natl Acad. Sci. U.S.A.* **109**, 10861-10866 (2012).

34. E. G. Hofman, *et al*. Ligand-induced EGF receptor oligomerization is kinase-dependent and enhances internalization. *J. Biol. Chem*. **285**, 39481-39489 (2010).
35. E. V. Bocharov, *et al*. The conformation of the epidermal growth factor receptor transmembrane domain dimer dynamically adapts to the local membrane environment. *Biochemistry* **56**, 1697-1705 (2017).
36. M. Lelimousin, V. Limongelli, M. S. P. Sansom, Conformational changes in the epidermal growth factor receptor: Role of the transmembrane domain investigated by coarse-grained metadynamics free energy calculations. *J. Am. Chem. Soc*. **138**, 10611-10622 (2016).
37. E. R. Purba, E.-I. Saita, I. N. Maruyama, Activation of the EGF Receptor by Ligand Binding and Oncogenic Mutations: The "Rotation Model". *Cells* **6**, 13 (2017). doi: 10.3390/cells6020013
38. M. Imielinski, *et al*. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**, 1107-1120 (2012).
39. I. N. Maruyama, Activation of transmembrane cell-surface receptors via a common mechanism? The "rotation model". *BioEssays* **37**, 959-967 (2015).
40. M. Adrian, J. Dubochet, J. Lepault, A. W. McDowall, Cryo-electron microscopy of viruses. *Nature* **308**, 32-36 (1984).
41. U. Skoglund, K. Andersson, B. Strandberg, B. Daneholt, Three-dimensional structure of a specific pre-messenger RNP particle established by electron microscope tomography. *Nature* **319**, 560-564 (1986).
42. L.-G. Öfverstedt, K. Zhang, L. A. Isaksson, G. Bricogne, U. Skoglund, Automated correlation and averaging of three-dimensional reconstructions obtained by electron tomography. *J. Struct. Biol.* **120**, 329-342 (1997).
43. K. Chin-Purcell, (motif edition), Bob-brick of Bytes, Minnesota Supercomputer Center, Inc (1993).
44. B. Chen, R. Mao, H. Wang, J. She, Cell line and drug-dependent effect of ERBB3 on cancer cell proliferation, chemosensitivity, and multidrug actions. *Int. J. High Throughput Screen* **1**, 49-55 (2010).
45. T. Cordes, J. Vogelsang, P. Tinnefeld, On the mechanism of Trolox as antiblinking and antibleaching reagent. *J. Am. Chem. Soc*. **131**, 5018-5019 (2009).
46. M. Tokunaga, N. Imamoto, K. Sakata-Sogawa, Highly inclined thin illumination enables clear single-molecule imaging in cells. *Nat. Methods* **5**, 159-161 (2008).

Acknowledgments: We are grateful to T. Sassa for his construction of pAC-N1-GFP-Shc1 and pAc-Grb2-GFP, A. Mugo for his advice on purification of the full-length EGFR, and T. Murayama and H. Iha for comments on the manuscript.

Funding: This work was supported by Okinawa Institute of Science and Technology Graduate University.

Competing interests: The authors declare that they have no competing interests.

Data and materials availability: The authors declare that the data supporting the findings of this study are available in the manuscript or supplementary materials, and the averaged cryo-ET density maps are deposited in the Electron Microscopy Data Bank (EMDB) under accession codes EMD: 30714, 30715, 30716, 30717, 30719, 30721, 30723, 30724, 30725, 30726, 30727, 30728, 30729, 30730, 30731, 30732, 30733, 30734, 30736, 30737, 30738, 30739, 30740, 30741,30742, 30743, 30744, 30745, 30746, 30747, 30748, 30749, 30750, 30751, 30752, 30753, 30754, 30755, 30756, 30757, 30758, 30759, 30760, 30761, 30762, 30763, 30764, 30765, 30862, and 30864.

Supplementary Information

Supplementary Figures S1 to S11

Supplementary Table S1

Supplementary Movie S1

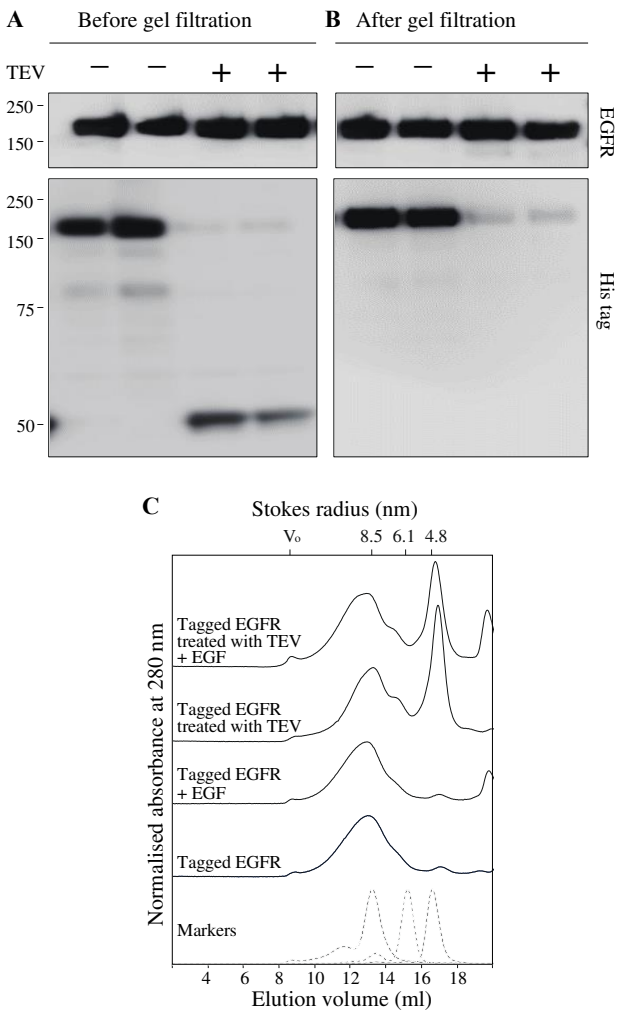


Fig. S1. His tag does not contribute to the EGFR dimer formation. (**A**) His-tagged EGFR molecules were purified by affinity chromatogr aphy, digested with TEV and separated by SDS-PAGE as described in Materials and Meth ods. After blotted to PVDF membrane, the proteins were stained with either antibody speci fic for EGFR (EGFR) or a His tag (His tag) (duplicate experiments). Note that the His-t agged EGFR was stained with His tag before TEV treatment, and that the majority of t he tag was cleaved off after TEV digestion. (**B**) After TEV cleavage, EGFR was purified by gel filtration, separated by SDS-PAGE and analyzed by Western blotting as above. (**C**) Gel filtration chromatograms of tagged EGFRs digested with or without TEV, followed by trea tment with or without EGF.

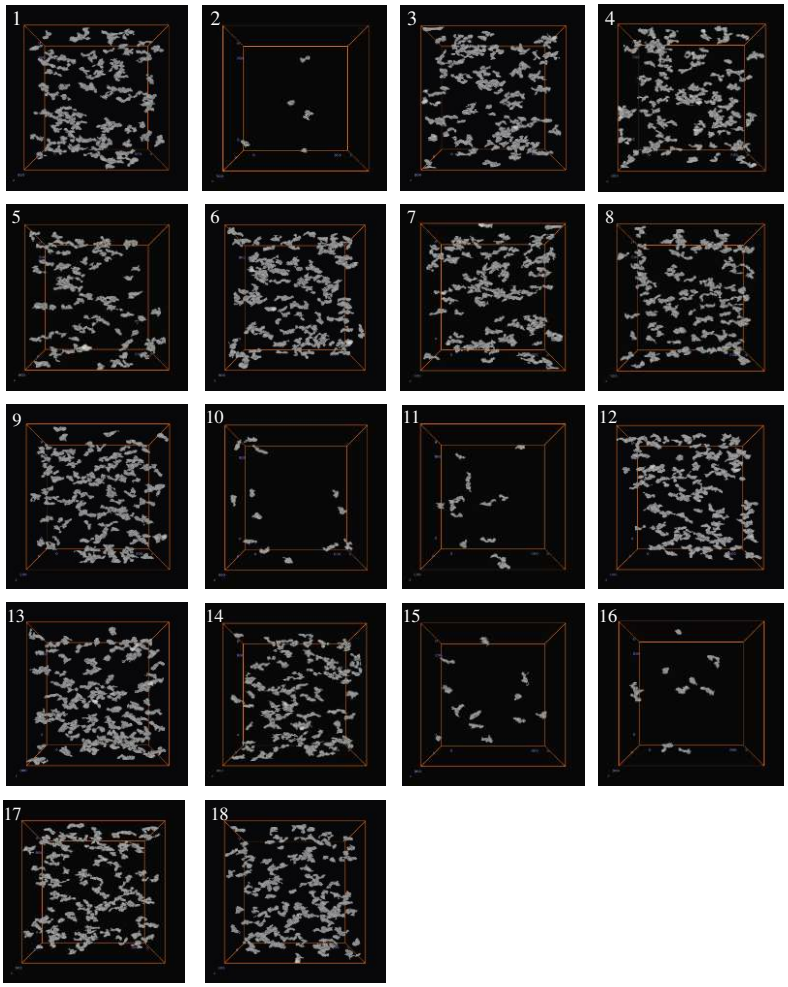


Fig. S2. Tomogram images of purified unligan ded EGFRs. A series of images of a specimen containing purified full-length EGFR molecules a t tilt angles ranging from -70° to +70°, incrementing in 1.0° steps, were recorded, aligned to a global center and used to produce 3D reconstructs of density values (800 × 800 × 800 voxels) lowpassed to 15 Å, using the COMET software package (version 6.4.2).

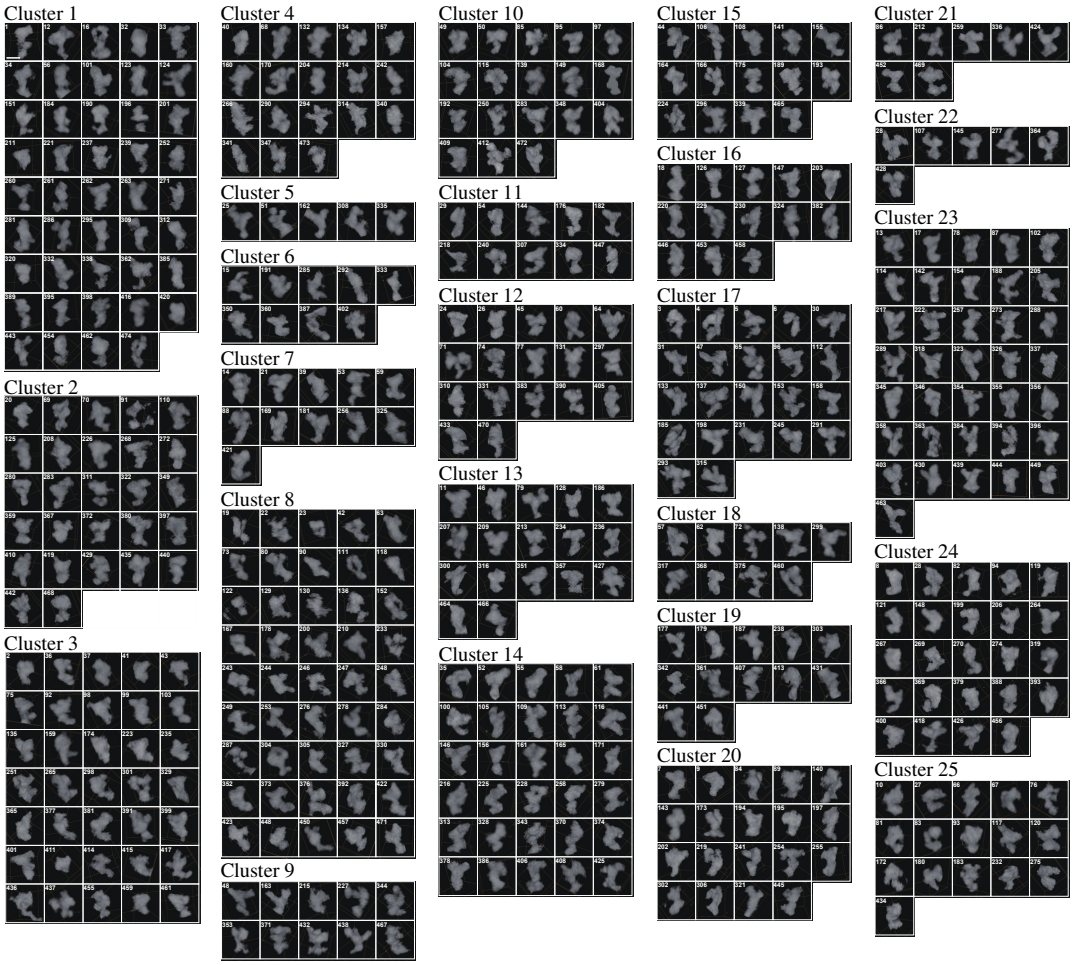


Fig. S3. Subtomogram images of unliganded EGFRs reconstructed. From the 18 tomograms (Fig. S2), the MINER program was used to extract subtomog rams consisting of 110 × 110 × 110 voxels. From the subtomograms, 474 unliganded EGFRs were recon structed and classified into 25 clusters, using CORR - PAIR. Scale, 10 nm.

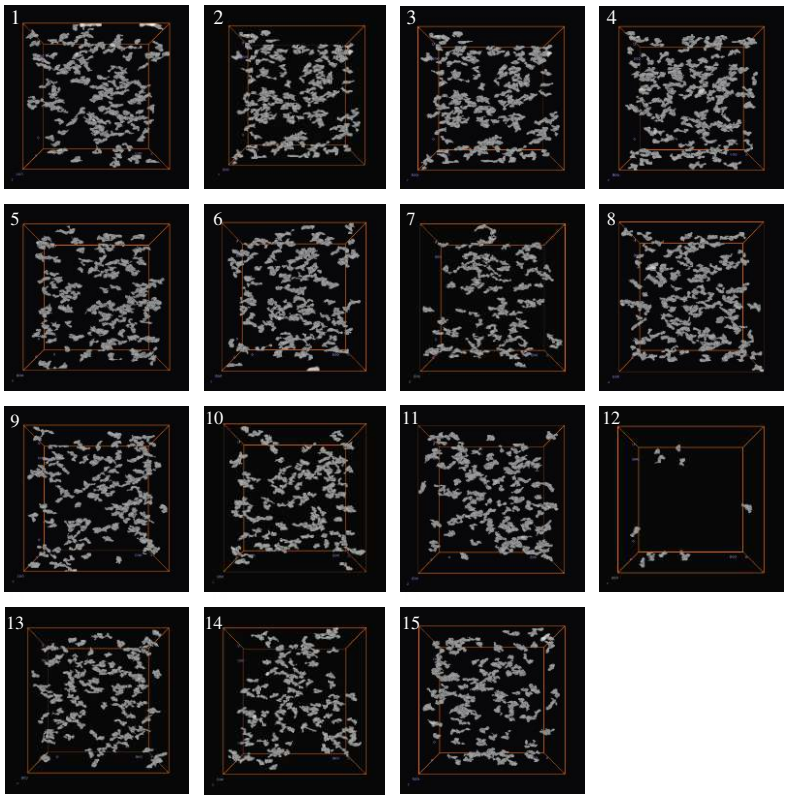


Fig. S4. Tomogram images of liganded EGFRs . A series of images of a specimen containing purified full-length EGFR molecules with bound ligand at tilt angles ranging from -70 ºto +70º, incrementing in 1.0ºsteps, were recorded, aligned to a global center and used to produce 3D reconstructs of density values (800 × 800 × 800 voxels) lowpassed to 15 Å, using the COMET software package (version 6.4.2).

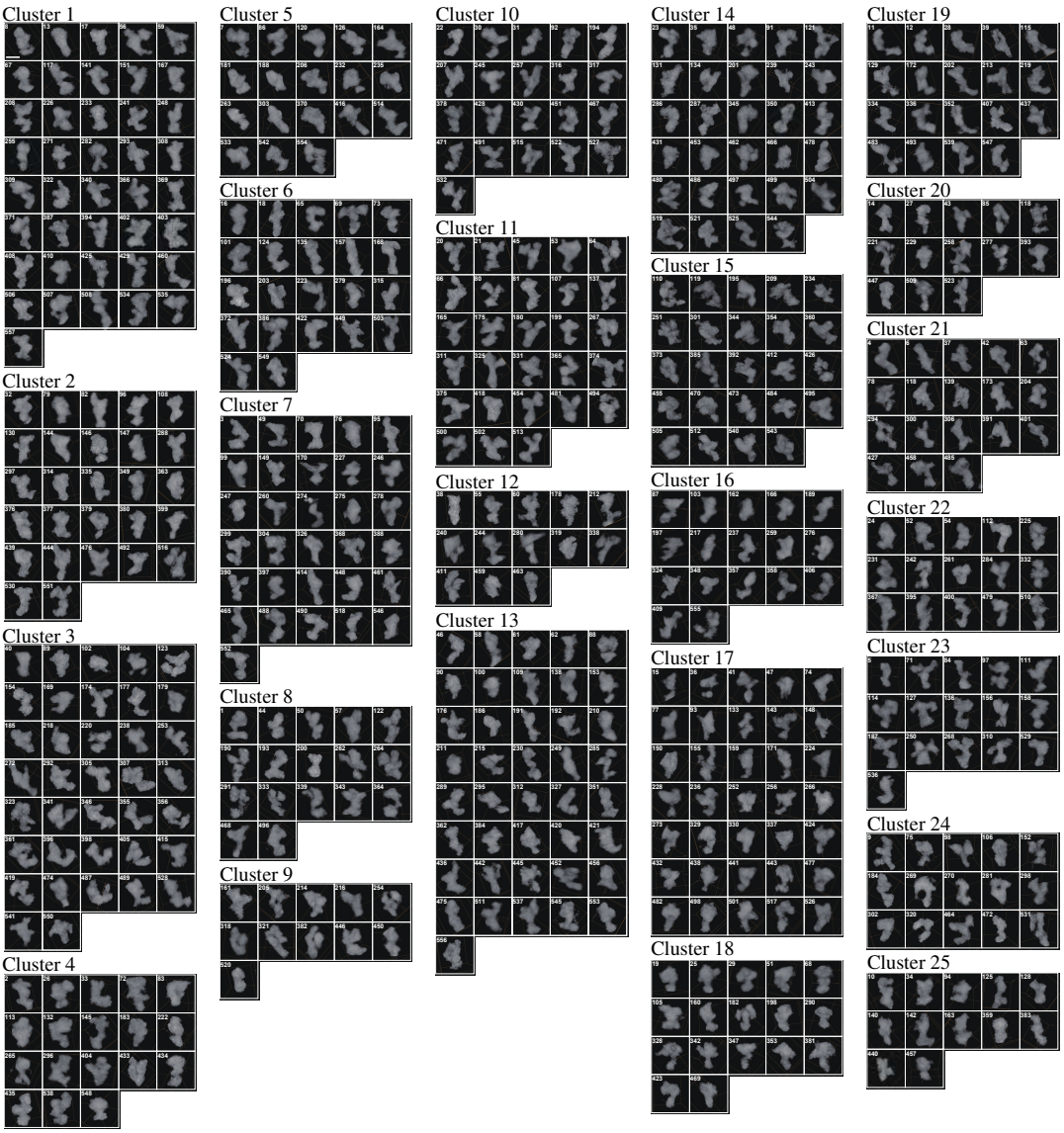


Fig. S5. Subtomogram images of liganded EGFRs reconstructed. From the 15 tomograms (Fig. S4), the MINER program was used to extract subtomogram s consisting of 110 × 110 × 110 voxels. From the subto - mograms, 557 liganded EGFRs were reconstructed and classified into 25 clusters, using CORRPAIR. Scale, 10 nm.

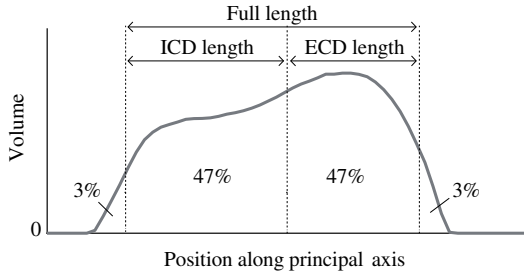


Fig. S6. Lengths of averaged subtomogram density envelopes. The volume of each averaged subtomogram was vertically divided into two halves by drawing a plane through the mid point of the principal axis, after removing 3% each of the total volume from both edges. As described in the main text, right and left halves of density maps were assigned to ECD and ICD, respectively. The sum of ECD and ICD lengths was defined as the full length.

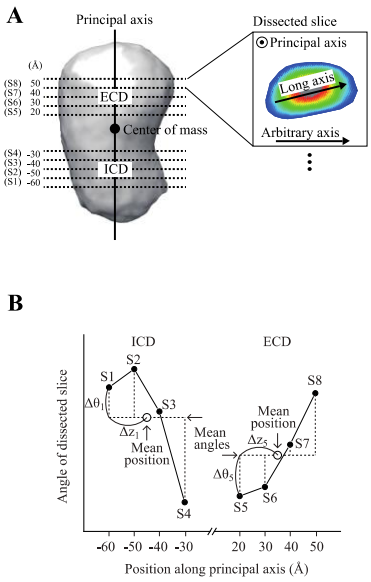


Fig. S7. Flexibility of ICD and ECD. (**A**) Electron density envelopes of 25 averaged unliganded and 25 averaged liganded subtomograms were horizontally dissected along the principal axis at eight places, whose centers were at -60, -50, -40, -30, 20, 30, 40, and 50 Å from the center of mass. Angles between the long axis of the dissected slices and an arbitrary axis were measured. (**B**) In EGFR ICD, the angles of four sections, whose centers were at -60 (S1 position), -50 (S2), -40 (S3) and -30 (S4) Å from the center of mass, were averaged, and differences, $\Delta\theta$i (i = 1, 2, 3, 4), between the measured angle and the mean angle was plotted against positions along the principal axis. The flexibility of averaged density was defined by $\Delta\theta$i/Δzi, where $\Delta\theta$i was normalized by the distance Δzi from the mean position. In EGFR ECD, the angles of four sections, whose centers were at 20 (S5), 30 (S6), 40 (S7) and 50 (S8) Å from the center of mass, were also averaged, and flexibility, $\Delta\theta$i/Δzi (i = 5, 6, 7, 8), was defined as above.

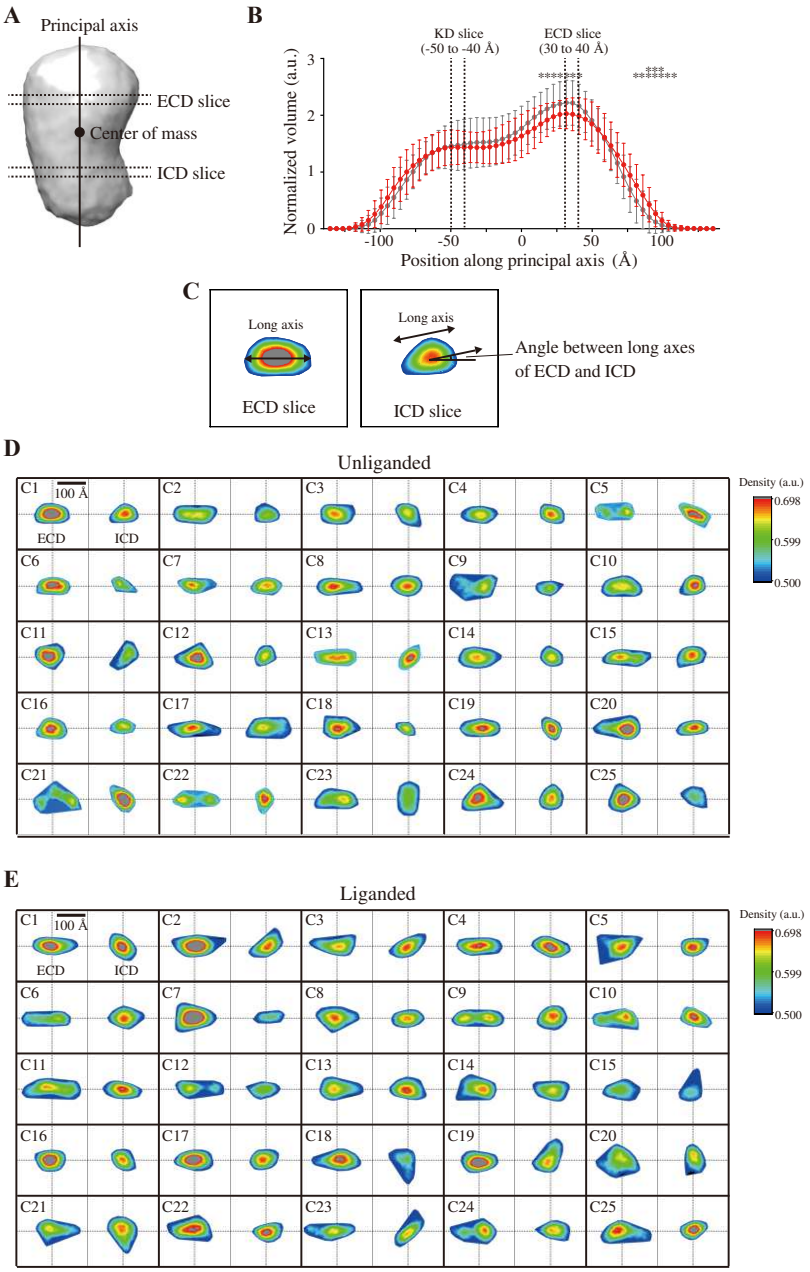


Fig. S8. Measurement of rotation angles betwe en ECD and ICD. (**A**), An image of 3D density maps. Dotted lines indicate regions from which 3 D slices were sampled as shown in (B). (**B**), Averaged density distributions of unliga nded (gray) and liganded (red) EGFRs along the principal axis. Density distributions of ECD and ICD shown in (C) were sampled from two regions, one each from ECD and ICD, indicated by dotted lines per pendicular to the principal axis. (**C**), 2D density maps of ECD (left) and ICD (right) sampled from the regions shown in (B). (**D**), 2D density distributions of slices indicated by the dotted lines from 25 averaged clusters of unliganded EGFRs. (**E**), 2D density distributions of slices from 25 average d clusters of liganded EGFRs.

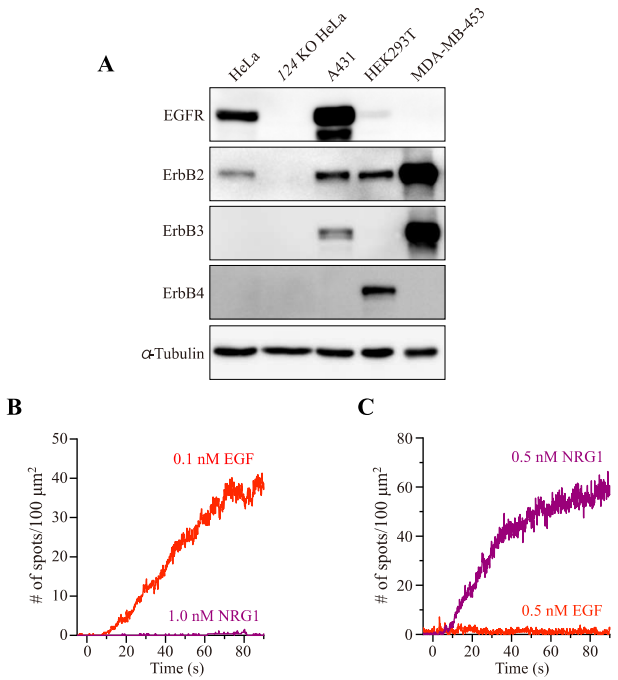


Fig. S9. HeLa cell lacking the expression of all four EGFR family members on the cell surface. Refer to Materials and Methods section for the construction of the "*124* KO HeLa" cell, which has mutations in *EGFR*, *HER2* and *HER4* genes, and for experimental conditions. (**A**) Western blots of original HeLa cells, "*124* KO HeLa" cells, and A431, HEK293T and MDA-MB-453 reference cells. Note that original HeLa cells did not produce HER3 and HER4 proteins at detectable levels, and that the "*124* KO HeLa" cell did not produce all four EGFR family members at detectable levels. (**B**) The ligand NRG1 binding to the "*124* KO HeLa" cell exogenously producing EGFR. NRG1 did not bind to the cell, whereas EGF did. (**C**) NRG1 binding to the "*124* KO HeLa" cell exogenously producing HER4. Note that NRG1 did bind to the cell, whereas EGF did not.

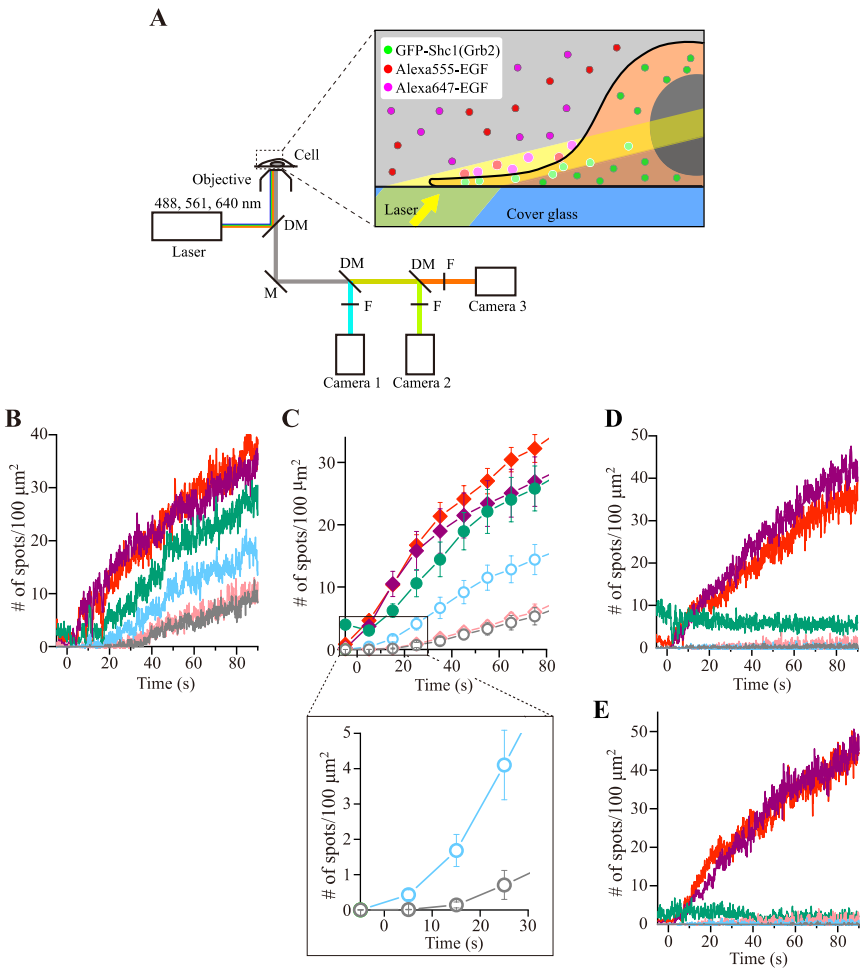


Fig. S10. Observation of EGF binding to EGFR on the cell surface at the single-molecule level. (**A**) A set-up for the observation of single molecules on upper surface of modified HeLa cells, using oblique illumination with lasers. M; mirror, DM; dichroic mirror, F; bandpass filter. (**B**) Time courses of EGF binding to the cell surface of the modified HeLa cell exogenously expressing EGFR. The number of spots of Alexa555-labeled EGF (red), Alexa647-labeled EGF (purple) or Grb2-GFP (green) appeared on the cell surface was plotted against incubation time. Colocalization of red and purple spots, green and either red or purple, and green and both of red and purple are shown in light blue, pink and gray, respectively. (**C**) Summary of four independent experiments. Data points are means +/- SEM. (**D** and **E**) Time courses of EGF binding to the cell surface of the modified HeLa cell exogenously expressing EGFR in the presence of a kinase inhibitor, Erlotinib. The number of spots of Alexa555-labelled EGF (red), Alexa647-labelled EGF (purple) and GFP-Shc1 [green in (D)] or Grb2-GFP [green in (E)] appeared on the cell surface was plotted against incubation time. Colocalization of red and purple spots, green and either red or purple, and green and both of red and purple are shown in light blue, pink and gray, respectively. Erlotinib, 5 μM at a final concentration, was preincubated for 30 min at 37°C before adding EGF at time = 0. Fluorescently labeled EGF molecules, 0.3 nM each at a final concentration, were used in (D); 0.1 nM each of fluorescently labeled EGFs in (E).

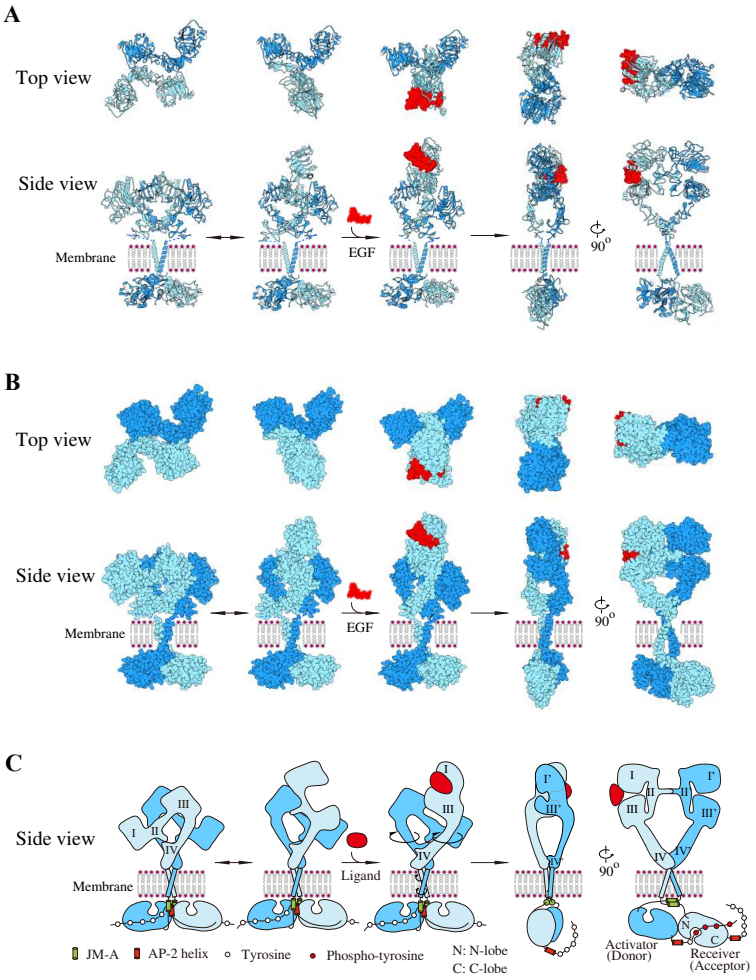
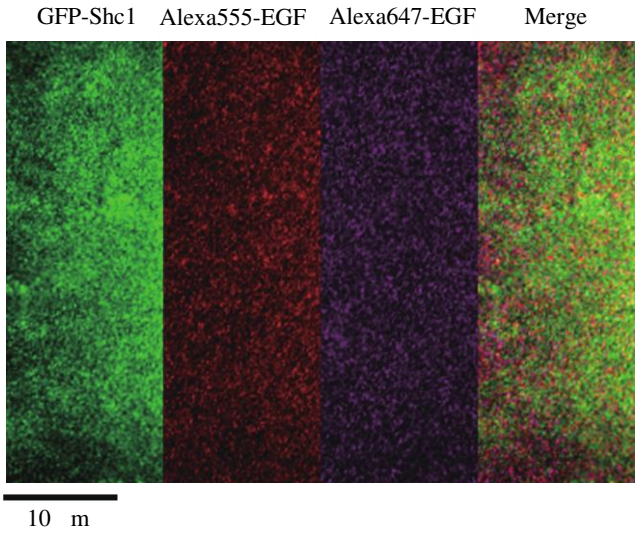


Fig. S11. Model for the activation of EGFR d imers by ligand binding. (**A**) Ribbon representation of the model. (**B**) Space-filling representation of the model. (**C**) Schematic representation of the model. Not drawn to scale . These structures are composites of the following structures: tethered (PDB ID: 1NQL) or extended (half of 3NJP) ECD monomer, ECD dimer (3NJP), inactive symmetric KD dim er (3GT8), NMR structure (2M0B) of inactive transmembrane -helices, active asymmetric KD dimer (2GS6), and NMR structure (2M20) of active transmembrane -helices. The crystal structure of inactive tethered ECD dimer was previously reported. In the active asy mmetric KD dimer, the C-lobe of the "activator" kinase domain interacts with the N-l obe of the "receiver" kinase for the activation of the latter as previously proposed. Upon ligan d binding, it has been proposed that the juxtamembrane segment JM-A forms an anti-pa rallel helical dimer (19, 35).

Table S1. Summary of cryo-ET data acquisition and image processing

Acquisition setup		
	Microscope	Titan Krios
	Voltage (keV)	300
	Detector	FEI FALCON II
	Energy-filter	No
	Å/pixel	2.258
	Magnification	37000×
	Acquisition scheme	Bi-directional
Processing	**Unliganded EGFR**	**Liganded EGFR**
# of tomogram	18	15
# of subtomogram	474	557
Defocus range	-1.0 μm to -1.5 μm	-1.0 μm to -1.5 μm
Total dose ($e^-/Å^2$)	90	90
Tilt range	-70° to 70°	-70° to 70°
# of tilts for reconstruction	134	135

Note: Only a single defocus was set for all the tilt series. A series of tilt angles were at a 1.0° increment. Coupled core-signalling unit and Fourier cell correlation have not been applied.



Movie S1. Single-molecule observation by TIRF microscopy with oblique illumination. Fluorescent spots appeared on the upper surface of cells were observed. The modified HeLa cell expressing GFP-Shc1 and EGFR was incubated with 0.1 nM Alexa555-labeled EGF and 0.1 nM Alexa647-labeled EGF. Scale bar, 10 m.