A molecular mechanism for the generation of ligand-dependent differential outputs by the epidermal growth factor receptor

Yongjian Huang[1,2,3], Jana Ognjenović[4], Deepti Karandur[1,2,3], Kate Miller[1], Alan Merk[4], Sriram Subramaniam[5*], John Kuriyan[1,2,3,6,7*]

[1] Department of Molecular and Cell Biology, University of California, Berkeley, Berkeley, CA 94720, USA.

[2] California Institute for Quantitative Biosciences, University of California, Berkeley, Berkeley, CA 94720, USA.

[3] Howard Hughes Medical Institute, University of California, Berkeley, Berkeley, CA 94720, USA.

[4] Frederick National Laboratory for Cancer Research, Frederick, MD 21701, USA.

[5] University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada.

[6] Department of Chemistry, University of California, Berkeley, Berkeley, CA 94720, USA.

[7] Divisions of Molecular Biophysics and Integrated Bioimaging, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

*Correspondence: sriram.subramaniam@ubc.ca, kuriyan@berkeley.edu

Abstract

The epidermal growth factor receptor (EGFR) is a receptor tyrosine kinase that couples the binding of extracellular ligands, such as EGF and transforming growth factor-α (TGF-α), to the initiation of intracellular signaling pathways. EGFR binds to EGF and TGF-α with similar affinity but generates different signals from these ligands. To address the mechanistic basis of this phenomenon, we have carried out cryo-EM analyses of human EGFR bound to EGF and TGF-α. We show that the extracellular module adopts an ensemble of dimeric conformations when bound to either EGF or TGF-α. The two extreme states of this ensemble represent distinct ligand-bound quaternary structures in which the membrane-proximal tips of the extracellular module are either juxtaposed or separated. EGF and TGF-α differ in their ability to maintain the conformation with the membrane-proximal tips of the extracellular module separated, and this conformation is stabilized preferentially by an oncogenic EGFR mutation. Close proximity of the transmembrane helices at the junction with the extracellular module has been associated previously with increased EGFR activity. Our results show how EGFR can couple the binding of different ligands to differential modulation of this proximity, thereby suggesting a molecular mechanism for the generation of ligand-sensitive differential outputs in this receptor family.

45 ## Introduction

46 Human EGFR can be activated by seven related ligands that generate different signaling outputs
47 from the receptor. While the classical mechanism for receptor tyrosine kinase activation, ligand-
48 induced dimerization, is essential for all these ligands to activate the receptor (Kovacs et al., 2015;
49 Leahy, 2004; Lemmon et al., 2014), it cannot explain how the extracellular module of EGFR is
50 able to respond differently to different ligands once the receptor is dimerized (Ronan et al., 2016;
51 Singh et al., 2016; Wilson et al., 2009). The extracellular module of EGFR consists of four domains,
52 denoted Domains I, II, III, and IV (Figure 1). Domains I and III sandwich the ligands between
53 them (Garrett et al., 2002; Ogiso et al., 2002). Domain II, which bridges the ligand-binding
54 domains, contains the "dimerization arm", a loop that mediates the principal interaction with the
55 other subunit in an activated dimer (Dawson et al., 2005). Domains I, II, and III form a compact
56 unit that we refer to as the ligand-binding "head", and Domain IV forms a rigid and elongated "leg"
57 that connects the ligand-binding head to the single transmembrane helix (Figure 1).

58 Previous crystallographic analyses of fragments of the EGFR extracellular module have
59 revealed different conformations of the ligand-binding head when bound to EGF, TGF-α,
60 epiregulin, and epigen (Freed et al., 2017; Garrett et al., 2002; Lu et al., 2010; Ogiso et al., 2002).
61 It appeared that these conformational differences might explain how different EGFR ligands
62 generate different signals (Wilson et al., 2009; Scheck et al., 2012), but the significance of these
63 structural difference is unclear because only fragments of the receptor were analyzed. For weakly-
64 binding ligands, such as epiregulin and epigen, the lifetime of the dimeric receptor is shortened
65 relative to that for the tighter-binding ligands, such as EGF, and the reduction in dimer lifetime
66 affects signaling output (Freed et al., 2017). However, EGF and TGF-α are both high-affinity

67 ligands (Jones et al., 1999), and they also generate different outputs from EGFR. How this happens
68 is still unclear.

69 Cell-based studies using fluorescent reporters of EGFR conformation have shown that the
70 binding of EGF and TGF-α results in different conformations of the intracellular juxtamembrane
71 segments of the receptor (Scheck et al., 2012; Doerner et al., 2015; Lowder et al., 2015). The
72 balance between alternative conformations of the juxtamembrane segments can be modulated by
73 point mutations in the transmembrane helices (Sinclair et al., 2018) and oncogenic mutations in
74 the kinase domains (Lowder et al., 2015). A more recent study reveals that simply switching the
75 conformation of the juxtamembrane segments by mutation suffices to switch EGFR signaling
76 outputs (Mozumdar et al., 2020). These studies demonstrate that the extracellular module, the
77 transmembrane helix, the intracellular juxtamembrane segments, and the intracellular kinase
78 domains are coupled allosterically, and that they respond differently to EGF and TGF-α. How the
79 two ligands generate these differences remains unknown. To gain structural insight into this
80 mechanism, we analyzed full-length human EGFR in complex with EGF and TGF-α by cryo-EM.

81 Results

82 Linkages between the extracellular modules and the transmembrane helices are not
83 rigidly defined

84 We purified full-length human EGFR expressed in a human HEK293S GnTI$^-$ cell line(Reeves et
85 al., 2002), and reconstituted the complex of EGFR with EGF into four membrane-mimetic
86 environments: detergent micelles, lipid nanodiscs (Bayburt and Sligar, 2010), amphipols (Gohon
87 et al., 2006), and peptidiscs (Angiulli et al., 2020; Carlson et al., 2018) (Figure 2-figure supplement
88 1). Cryo-EM reconstructions of the full-length EGFR:EGF complex in all four reconstitutions look

89 very similar, and only the extracellular modules are resolved in the cryo-EM densities (Figure 2).
90 The lack of resolvable density for the transmembrane helices and the intracellular kinase domains
91 suggests that the linkage between the extracellular modules and the transmembrane helices is not
92 rigidly defined, consistent with 2D class averages of negative-stain EM images (Mi et al., 2011)
93 (Figure 2-figure supplement 1). Thus, mechanisms for ligand-dependent differential responses in
94 EGFR that require strict stereochemical coupling between the extracellular ligand-binding
95 modules and the transmembrane helices appear to be ruled out. Recently published cryo-EM
96 analyses of intact insulin receptor and intact type 1 insulin-like growth factor receptor also did not
97 resolve the transmembrane and intracellular modules (Li et al., 2019; Uchikawa et al., 2019; Zhang
98 et al., 2020), suggesting that a lack of rigidity in the connection between the extracellular module
99 and the transmembrane helices is a feature common to receptor tyrosine kinases.

100 Cryo-EM single-particle analysis reveals a range of conformations for the EGF-
101 bound extracellular module

102 After focusing the cryo-EM reconstructions on the extracellular module alone, the dataset for the
103 full-length EGFR:EGF complex reconstituted in peptidiscs yielded the highest resolution
104 reconstruction of the dimeric extracellular module, with an overall resolution of 2.9 Å (Figure 3A,
105 Figure 3-table supplement 1). In this reconstruction, cryo-EM density for the Domain IV legs is
106 less well resolved than for the heads. To better understand the origin of this apparent structural
107 heterogeneity, we used 3D classification to further partition the entire dataset of 765,883 particles
108 into 10 subclasses (Punjani et al., 2017). Cryo-EM reconstructions corresponding to individual
109 subclasses resulted in overall resolutions ranging between 3.1 Å and 3.7 Å (Figure 3A, Figure 3-
110 figure supplement 1).

111 The resulting structures for different subclasses differ from one another in the extent of a
112 "scissor-like" rotation between the two ligand-binding heads of the dimeric extracellular module
113 (Figure 3A, Figure 3-video supplement 1-3). The rotation between the two ligand-binding heads
114 is due to a rigid-body rotation of each head about an axis parallel to the dimerization arms, and is
115 correlated with a slight twist in the dimerization arm of each EGFR subunit (Figure 3B-D, Figure
116 3-video supplement 1-3). The extent of this intersubunit rotation can be estimated by defining a
117 dihedral angle corresponding to rotation about a virtual bond connecting the Cα atoms of Thr 249
118 in each EGFR subunit (we number amino acid residues in EGFR without including the 24-residue
119 signal sequence). The positions of the Cα atoms of Ile 190 in each subunit are used to further
120 specify the dihedral angle, and the value of this dihedral angle ranges from ~10° to ~25° in the 10
121 representative structures (Figure 3A).

122 The scissor-like rotation between ligand-binding heads in the cryo-EM structures is
123 reminiscent of differences noted previously between crystal structures of fragments of EGFR
124 extracellular modules bound to different ligands (Liu et al., 2012). At one extreme, the EM-derived
125 structure with the largest intersubunit rotation (~25°) resembles closely the crystal structure of the
126 ligand-binding head of EGFR bound to EGF (PDB code: 3NJP (Lu et al., 2010; Ogiso et al., 2002),
127 root-mean-square deviations (RMSDs) in Cα positions calculated over EGFR residues 1-501: 0.8
128 Å, Figure 4A). The EM-derived structure with the smallest intersubunit rotational angle (~10°)
129 resembles the crystal structure of the EGFR ligand-binding head bound to TGF-α (PDB code:
130 1MOX (Garrett et al., 2002), RMSDs in Cα positions calculated over EGFR residues 1-501: 1.5
131 Å, Figure 4B). Thus, the partitioning of the cryo-EM data into subclasses reflects an intrinsic
132 conformational variability of the ligand-bound EGFR extracellular module. Importantly, the
133 cryoEM results demonstrate that the conformations ascribed previously to specific binding of EGF

134 or TGF-α to the receptor are actually part of an ensemble of conformations that are readily
135 accessible to EGFR when bound to either one of these ligands.

136 Different dimeric conformations of the ligand-bound extracellular module have
137 different separations between the tips of the two legs

138 The two Domain IV legs in the dimer behave differently when the scissor-like rotational angle
139 between the ligand-binding heads changes. The structure of the Domain IV leg from a previously
140 determined crystal structure of the EGFR extracellular module (PDB code: 3NJP (Lu et al., 2010))
141 can be docked into our cryo-EM densities (Figure 3-figure supplement 2A, 2B). This allowed us
142 to determine the position and orientation of the Domain IV legs associated with each representative
143 conformation from the ensemble revealed by cryo-EM. When the rotational angle between the two
144 heads is at its largest value (~25°), the two legs adopt similar conformations, and the membrane-
145 proximal tips of the two legs are juxtaposed, with a separation of ~5 Å between the Cα atoms of
146 the last residue in the legs (Thr 614) (the "juxtaposed" conformation, Figure 3B, Figure 5). When
147 the rotation between the two subunits is small, maintenance of the same conformation in both legs
148 would lead to steric clashes between the two legs (Figure 3-figure supplement 2C). The clashes
149 are relieved by one of the two Domain IV legs undergoing a hinge rotation around the connection
150 between the head and the leg, and we refer to this leg as "flexible". This hinge rotation causes the
151 flexible leg to swing upwards, closer to the ligand-binding head, increasing the separation between
152 the membrane-proximal tips of the two Domain IV legs in the dimer. When the angle between the
153 heads is at its minimal value (~10°), the flexible leg swings maximally towards the ligand-binding
154 head, increasing the distance between the membrane-proximal tips of the extracellular module by
155 nearly 15 Å (the "separated" conformation, Figure 3D, Figure 5).

156 **The tips-separated conformation of the dimeric extracellular module can couple more**
157 **readily to the N-terminal dimer of the transmembrane helices**

158 The EGFR transmembrane helix contains two dimerization motifs, one located toward the N-
159 terminal end of the transmembrane helix, near the extracellular face of membrane, and the other
160 located toward the C-terminal end (Mendrola et al., 2002). These dimerization motifs are involved
161 in close association between the transmembrane helices in alternative structures defined by NMR,
162 in which the helices are splayed apart at either the N-terminal or the C-terminal ends (Bocharov et
163 al., 2016, 2017; Endres et al., 2013). To understand how these alternative dimeric configurations
164 of the transmembrane helices connect to the different conformations of the extracellular module,
165 with different separation between the tips of the Domain IV legs, we analyzed two structures of
166 EGFR transmembrane helix dimers determined previously by NMR. One structure shows N-
167 terminal association of the transmembrane helices (PDB code: 5LV6 (Bocharov et al., 2017)) and
168 the other shows C-terminal association (PDB code: 2M0B (Bocharov et al., 2016)) (Figure 6A).

169 The seven-residue linker connecting the extracellular module and the transmembrane
170 helices (Asn 615-Ser 621) is modeled in both NMR structures. The three C-terminal residues of
171 the linker (Ile 619-Ser 621), located at the junction with the transmembrane helix, adopt well-
172 defined conformations in both NMR structures. The hydrophobic sidechain of Ile 619 packs
173 against the sidechain of Ala 623 in the first turn of the transmembrane helix. Ile 619 is followed
174 by a 3_{10}-helical turn spanning Pro 620 to Ala 623 (Figure 6B). These structural constraints cause
175 the linker-transmembrane helix junction to orient parallel to the membrane surface in both
176 structures (Figure 6A). As a consequence, in the N-terminal dimer of the transmembrane helices,
177 the two linkers are oriented away from each other, with the two Ile 619 residues separated by ~20
178 Å. In contrast, the C-terminal dimer has the two linkers oriented towards each other, with the

179 distance between the two Ile 619 residues at ~10 Å. Thr 614 is at the junction between Domain IV
180 leg of the extracellular module and the linker. In the ensemble of NMR structures for the N-
181 terminal dimer of the transmembrane helices (Bocharov et al., 2017), the two threonine residues
182 (Thr 614) are separated by an average distance of ~35 Å. In comparison, the separation between
183 the two threonine residues is ~20 Å in the ensemble of structures for the C-terminal dimer of the
184 transmembrane helices (Bocharov et al., 2016) (Figure 6A). These observations suggest that the
185 junctions between the extracellular module and the transmembrane helices are more separated in
186 the N-terminal dimer of the transmembrane helices than the connection in the C-terminal dimer.

187 We used molecular dynamics simulations to explore the constraints on the separation
188 between the two threonine residues at position 614 in the linker, starting either the N-terminal
189 dimer configuration (PDB code: 5LV6 (Bocharov et al., 2017)) or the C-terminal dimer
190 configuration (PDB code: 2M0B (Bocharov et al., 2016)). Following equilibration (see Materials
191 and Methods for full details), we initiated twenty-five independent molecular dynamics trajectories
192 for both configurations, each 100 ns long, for a total simulation time of 2.5 μs for each
193 configuration. The configuration of the transmembrane helices remains close to the starting
194 configuration over this timescale. The conformation of the junction between the linker and the
195 transmembrane helices is also very stable in both sets of simulations, with the sidechain of Ile 619
196 remaining associated with Ala 623 over the course of the simulations (Figure 6E, 6F). In the
197 simulations of the N-terminal dimer configuration, the linkers explore a range of configurations
198 but are oriented predominantly such that they point away from one another, and the distance
199 between Thr 614 from each subunit ranges between 20 Å and 35 Å (Figure 6C). In the simulations
200 starting from the C-terminal dimer configuration, the Thr 614 residues from each subunit are much
201 closer to each other, predominantly sampling distances between 7 Å and 12 Å (Figure 6C). These

202 results further support the idea that the junction between the linker and the transmembrane helices
203 imposes a constraint on the orientation of the linkers. When the transmembrane helices are in the
204 N-terminal dimer configuration, this linker favors a connection to the tips-separated conformation
205 of the extracellular module, and when the transmembrane helices are in the C-terminal dimer
206 configuration, a connection to the tips-juxtaposed conformation of the extracellular module is
207 indicated.

208 We computed the free-energy change of the system when the distance between the Thr 614
209 residues from each subunit is driven from 27 Å to 8 Å, when the transmembrane helices are in the
210 N-terminal dimer configuration. We used umbrella-sampling simulations (Torrie and Valleau,
211 1977), with sampling windows spaced 1 Å apart (for a total of 20 windows) and starting structures
212 for each window taken from a steered molecular dynamics simulation. Three independent
213 umbrella-sampling calculations were performed, with the starting structures for each taken from
214 different steered molecular dynamics trajectories. These calculations converged (i.e., the potential
215 of mean force does not change) when each window sampled ~60-90 ns of simulation. All three
216 umbrella-sampling simulations yield similar results (Figure 6D), with the free-energy being lowest
217 when the distance between the Cα atoms of Thr 614 residues from each subunit is ~26 Å.
218 Decreasing this distance to 8 Å leads to the free energy of the system increasing by ~9 $kJ.mol^{-1}$
219 (~3.5 $k_B T$ at room temperature, Figure 6D). This indicates that in the N-terminal dimer
220 configuration, it is thermodynamically unfavorable for the linkers between the extracellular
221 modules and the transmembrane helices to sample conformations where the distance between the
222 Thr 614 residues is small. Thus, when the transmembrane helices are in the N-terminal dimer
223 configuration, it is likely that the tips-separated conformation of the extracellular module dimer,
224 with the tips of Domain IV at a larger distance, is favored over the tips-juxtaposed conformation.

225 **An oncogenic mutation in the kinase domain stabilizes the tips-separated**
226 **conformation of the extracellular module**

227 Dimerization of the transmembrane helices through the N-terminal dimerization motif is
228 associated with increased kinase activity (Miloso et al., 1995; Lu et al., 2010; Pahuja et al., 2018;
229 Sinclair et al., 2018), so our structural analysis of the connection between the extracellular module
230 and the transmembrane helices suggests that the tips-separated conformation of the extracellular
231 module is associated with a more active conformation of the kinase domains, through its coupling
232 to the N-terminal dimer of the transmembrane helices.

233 To test this hypothesis, we examined an oncogenic variant of EGFR seen frequently in
234 cancer patients, in which a leucine residue in the activation loop of the intracellular kinase domain
235 is substituted by arginine (L834R; this is residue 858 when the signal sequence is included). The
236 activation loop is a key allosteric control element at the active site of the kinase domain, and the
237 L834R mutation promotes the formation of the catalytically-active asymmetric dimer of the kinase
238 domains, leading to hyper-activation of EGFR (Red Brewer et al., 2013; Shan et al., 2012). We
239 carried out cryo-EM analysis of full-length EGFR(L834R) bound to EGF, using the same
240 procedures as for wild-type EGFR. The two structures representing the tips-juxtaposed and the
241 tips-separated conformations of the EGFR(L834R):EGF complex were determined at resolution
242 of 3.3 Å and 3.4 Å, respectively (Figure 3-table supplement 1).

243 The overall structures of the tips-juxtaposed and tips-separated conformations of the
244 EGFR(L834R):EGF complex are very similar to that seen for the corresponding complex of the
245 wild-type receptor. The effect of the activating mutation, L834R, becomes apparent when
246 comparing the quality of the cryo-EM densities in the Domain IV legs for the wild-type and mutant
247 EGFR. For the tips-juxtaposed conformations, the cryo-EM densities are of comparable quality

248 (Figure 7). An important result emerges when comparing the tips-separated conformations for
249 wild-type EGFR and the oncogenic variant. Although the oncogenic L834R mutation involves a
250 residue that is located in the intracellular kinase domain, this mutation markedly stabilizes the
251 flexible leg in the tips-separated conformation of the extracellular module on the other side of the
252 membrane, compared to wild-type EGFR. The cryo-EM density for the flexible leg is much better
253 resolved in the EGFR(L834R):EGF complex than for wild-type EGFR, with secondary structural
254 features clearly visible (Figure 7). Thus, activation of the intracellular kinase domains by
255 oncogenic mutation preferentially stabilizes the flexible leg in the tips-separated conformation of
256 the extracellular module, providing direct evidence for conformational coupling between the
257 extracellular and intracellular modules of EGFR. These results are in agreement with the idea that
258 the tips-separated conformation of the extracellular module is coupled to the more active
259 conformations of the transmembrane helices and the intracellular kinase domains.

260 EGF and TGF-α differ in their ability to stabilize the flexible Domain IV leg in the 261 tips-separated conformation

262 Given the association of the tips-separated conformation of the extracellular module with
263 increased activity of the receptor, we wondered if EGF and TGF-α can differentially modulate
264 different conformations of the extracellular module. We carried out cryo-EM analysis of EGFR
265 bound to TGF-α using the same procedures as before. Compared to the EGFR:EGF complex, the
266 flexible leg in the tips-separated conformation of the EGFR:TGF-α complex is destabilized. The
267 cryo-EM density for the flexible leg is poorly defined, and there is no interpretable density for this
268 leg in the reconstruction of the tips-separated conformation of the EGFR:TGF-α complex (Figure
269 7). For the tips-juxtaposed conformation, the reconstructions for the EGFR:EGF and EGFR:TGF-

270 α complexes are similar (Figure 7), with near-symmetrical densities for the Domain IV legs in
271 both complexes. The only noticeable difference is that the cryo-EM density at the very end of the
272 Domain IV legs in the EGFR:TGF-α complex is slightly weaker than the corresponding density
273 in the EGFR:EGF complex (Figure 7).

274 The structures we have determined provide a plausible explanation for why EGF and TGF-
275 α differ in their ability to stabilize the flexible leg in the tips-separated conformation. Several key
276 interactions with Domain I of the receptor are different between EGF and TGF-α (Figure 8A). As
277 a result, the ligand-binding pocket is slightly compressed when TGF-α is bound. We used
278 molecular dynamics simulations to investigate this difference. We initiated four trajectories from
279 the crystal structure of the EGFR extracellular module bound to EGF (PDB code: 3NJP (Lu et al.,
280 2010)). We also initiated four trajectories starting from this crystal structure (PDB code: 3NJP),
281 but with the EGF replaced by TGF-α, taken from the crystal structure of truncated EGFR bound
282 to TGF-α (PDB code: 1MOX (Garrett et al., 2002)). Each trajectory was run for between ~350 and
283 ~550 ns, for a total simulation time of ~1.8 μs for EGFR bound to EGF and ~2.0 μs for EGFR
284 bound to TGF-α. We measured the angle between the centroids of Domain I, part of Domain II
285 and Domain III (Figure 8B) over the course of the simulations. In EGFR bound to EGF, this angle
286 is slightly larger, with the mode occurring at ~80°. In the simulations of EGFR bound to TGF-α,
287 this angle is slightly smaller, with the mode occurring at ~75°, indicating that the ligand-binding
288 pocket is slightly compressed when EGFR is bound to TGF-α (Figure 8C). This compression
289 between Domains I and III causes Domain II to bend in the TGF-α complex relative to its structure
290 in the EGF complex (Figure 8A), leading to a small difference in the orientation of the dimerization
291 arm with respect to Domains III and IV. EGF and TGF-α both maintain the same interface between
292 the dimerization arms of the two EGFR subunits in the dimer, but the additional bending of

293 Domain II in the TGF-α complex requires the two Domain IV legs to rotate towards each other.
294 In the tips-separated conformation of the TGF-α complex, this inward motion would cause a steric
295 clash between the two legs if the legs maintained the conformation seen in the EGF complex
296 (Figure 9A). The increased disorder in the Domain IV leg when TGF-α is bound appears to be a
297 result of this close contact being avoided.

298 These results demonstrate that the structural difference between the two ligand-bound
299 EGFR complexes can propagate from the ligand-binding head to the disposition of the Domain IV
300 legs. Due to the more compressed ligand-binding pocket in the TGF-α complex, which leads to
301 potential steric clashes between the Domain IV legs, TGF-α is less effective than EGF at
302 stabilizing the Domain IV legs in the tips-separated conformation of the extracellular module.

303 EGF and TGF-α induce different levels of phosphorylation in EGFR and Erk

304 Our cryo-EM analysis of the oncogenic EGFR variant (L834R) show that the stabilization
305 of the flexible Domain IV leg in the separated conformation is correlated with the presence of an
306 oncogenic mutation that results in increased levels of kinase activity. We speculate that the
307 observed differences in the ability of EGF and TGF-α to stabilize the Domain IV leg in the tips-
308 separated conformation underlies the ability of these two ligands to trigger different levels of
309 kinase activity and downstream signaling from EGFR. Specifically, our structural model suggests
310 that the EGF-bound EGFR complex is likely to be more active than the TGF-α bound EGFR.

311 Although previous studies have demonstrated that these two ligands can induce different
312 long-term cellular responses (Wilson et al., 2009), it is unclear how this is manifested at the level
313 of receptor autophosphorylation and the activation of proximal signaling events. To test the effect
314 of these two ligands on the kinase activity of EGFR, we first purified detergent-solubilized full-

315 length human EGFR from mammalian cells (HEK293FT), followed by stimulation of EGFR
316 activity through addition of EGF and TGF-α, and detection of phosphorylation on EGFR by
317 western-blot analysis. Indeed, the results show that EGF induces stronger EGFR
318 autophosphorylation than TGF-α (Figure 9-figure supplement 1A). Next, we transiently
319 transfected mammalian cells with full-length human EGFR and stimulated the cells with EGF and
320 TGF-α, respectively. EGFR phosphorylation levels in the stimulated cells were analyzed by
321 western blot. Consistent with the results obtained with the purified EGFR protein, the EGFR
322 phosphorylation level is higher in the EGF-stimulated cells than that in the TGF-α-stimulated cells
323 (Figure 9-figure supplement 1B). Importantly, ligand-dependent differences can also be detected
324 in the downstream signaling events following receptor activation, by monitoring the
325 phosphorylation level of the extracellular signal–regulated kinase (Erk). As shown in Figure 9C,
326 EGF stimulates higher levels of Erk phosphorylation than TGF-α does, in cells transiently
327 transfected with EGFR.

328 According to our structural model, for these two ligands to differentially activate EGFR,
329 the structural difference originating from the ligand-binding head needs to be transmitted to the
330 tips of the extracellular module, and thereby to the transmembrane helices. We hypothesize that
331 disrupting the conformational coupling between the ligand-binding head and the Domain IV leg,
332 or the coupling between the Domain IV leg and the transmembrane helices, will interfere with the
333 differential effects of EGF and TGF-α. To disrupt the connection between the ligand-binding head
334 and the Domain IV leg, we mutated a conserved residue (Trp 492) buried in a hydrophobic pocket
335 between Domain III and Domain IV (Figure 9B). Mutating this tryptophan to glycine (W492G) is
336 expected to loosen the connection between Domain III and Domain IV, and the Erk

337 phosphorylation assay shows that this mutation results in levels of Erk phosphorylation that are
338 the same for activation by EGF and TGF-α (Figure 9C).

339 Since our structural model suggests that the conformation of the linker at the junction
340 between the extracellular module and the transmembrane helices is important for the coupling
341 between the two domains, replacing the linker (Asn 615-Pro 620) with a flexible sequence, Gly-
342 Gly-Ser-Gly-Gly-Ser (GGSGGS), should compromise this coupling. This mutation (replacement
343 of the linker by GGSGGS) breaks the differential Erk activation by EGF and TGF-α (Figure 9D,
344 9E). In another construct, we inserted a flexible segment, with sequence Gly-Gly-Ser-Gly-Gly-
345 Ser-Gly-Gly-Ser (GGSGGSGGS), into the linker connecting the extracellular module to the
346 transmembrane helix, between Pro 617 and Lys 618. The insertion of this flexible segment results
347 in a similar extent of Erk activation by EGF and TGF-α (Figure 9D, 9E), which is consistent with
348 the importance of this linker in coupling the conformations of the extracellular module and the
349 transmembrane helix.

350 Conclusions

351 Our cryo-EM analysis provides a molecular explanation for the ability of EGFR to respond
352 differently to two high-affinity ligands, EGF and TGF-α (Figure 10). The differential activation
353 of EGFR by these two ligands arises from structural differences in the receptor-ligand interactions.
354 As a result, EGF and TGF-α differ in their abilities to stabilize the tips-separated conformation of
355 the extracellular module that couples to the N-terminal dimer configuration of the transmembrane
356 helices, which is further coupled to a specific configuration of the juxtamembrane segments
357 (Arkhipov et al., 2013; Endres et al., 2013; Jura et al., 2009; Sinclair et al., 2018), ultimately
358 promoting the formation of kinase dimers with higher activity (Zhang et al., 2006). In this way,

359 EGF and TGF-α can generate different responses from EGFR by differentially modulating the
360 structure of the junction point between the extracellular module and the transmembrane helices,
361 altering the balance between alternative dimeric configurations of the transmembrane helices
362 (Figure 10B). A lack of rigidity in the linkage between the extracellular module and the
363 transmembrane helices is emerging as a common feature in receptor tyrosine kinases (Diwanji et
364 al., 2021; Li et al., 2019; Lu et al., 2010; Uchikawa et al., 2019; Zhang et al., 2020), and the
365 mechanism we have defined, involving modulation of the separation of the tips of the extracellular
366 modules, could potentially be utilized by other receptor tyrosine kinases to generate ligand-specific
367 signaling outputs (Ahmed et al., 2020).

368 Members of the EGFR family are important targets for the development of cancer
369 therapeutics, including antibodies that block receptor dimerization and activation (Arteaga and
370 Engelman, 2014; Cai et al., 2020). Our structural analysis suggests new strategies for the
371 modulation of EGFR signaling. For example, antibodies or peptides that can bind to the Domain
372 IV legs and stabilize the juxtaposed conformation would allow the receptor to dimerize but send
373 weaker signals. Such a strategy may circumvent resistance mechanisms that arise in response to
374 antibodies and inhibitors that shut down EGFR signaling completely (Brand et al., 2011; Chong
375 and Jänne, 2013; Cai et al., 2020).

376 Materials and Methods

377 Cell culture

378 Insect cells

379 *Spodoptera frugiperda* Sf9 cells were maintained as suspension cultures at 27 ^{o}C in ESF-
380 921™ Insect Cell Culture Medium (Expression Systems). Isolation of bacmid DNA, transfection

381 of Sf9 cells, and amplification of baculovirus were performed following the protocols detailed in
382 the Bac-to-Bac Baculovirus Expression System (ThermoFisher Scientific).
383
384 **Mammalian cells**
385 Human HEK293S GnTI$^-$ cells were cultured as suspension at 37°C in FreeStyle293
386 expression medium (Invitrogen) supplemented with 2% (vol/vol) FBS, in the presence of 8% CO_2.
387 The HEK293S GnTI$^-$ suspension cells were maintained between 0.2 and 3 x 10^6 cells/ml.
388 HEK293T and HEK293FT cells were cultured at 37°C in complete Dulbecco's modified Eagle's
389 medium (DMEM) supplemented with GlutaMAX™ (ThermoFisher) and 10% (vol/vol) FBS, in
390 the presence of 5% CO_2.
391
392 **Full-length EGFR protein expression and purification for cryo-EM studies**
393 **Protein expression**
394 A cDNA fragment encoding full-length human EGFR (UniProt accession no. P00533) was
395 cloned into the pEG_BacMam vector (Goehring et al., 2014). To facilitate affinity purification, a
396 FLAG-tag (with sequence DYKDDDDK) was inserted after the signal sequence of EGFR. To
397 monitor protein expression and to reduce non-specific degradation of the C-terminal tail of EGFR,
398 EGFP was fused after the C-terminus of EGFR. An oncogenic EGFR variant (L834R) was
399 generated by site-directed mutagenesis.
400 Plasmid DNA carrying the EGFR sequence was transformed into DH10Bac competent
401 cells (ThermoFisher Scientific) to produce the bacmid DNA, following the protocol detailed in the
402 Bac-to-Bac Baculovirus Expression System. To produce the recombinant baculovirus, purified
403 bacmid DNA was transfected into Sf9 cells using Cellfectin II (ThermoFisher Scientific). After

404 two rounds of amplification, the recombinant baculovirus was used to infect HEK293S GnTI⁻ cells
405 at a 1:20 (vol/vol) ratio. After 24 hours, the infected HEK293S GnTI⁻ suspension culture was
406 supplemented with 2 μM of Erlotinib, an EGFR inhibitor, and 8 mM sodium butyrate, for optimal
407 protein expression and stability. Cells were harvested by centrifugation 72 hours after the infection.
408

409 **Protein purification**

410 Harvested cells were resuspended in lysis buffer (25 mM Tris pH 8.0, 150 mM NaCl, 2
411 μM Erlotinib and cOmplete Protease Inhibitor Cocktail (Roche)) and were lysed by sonication in
412 pulse mode for 10 minutes, with 2s intervals between each 1s pulse. Large debris and unbroken
413 cells were removed from the crude lysate by centrifugation for 12 mins at 19,000g. The membrane
414 fraction was pelleted by ultra-centrifugation at 190,000g for 1 hr and resuspended in lysis buffer.
415 To extract EGFR from the membrane, 1% DDM (Dodecyl-β-D-Maltoside, Anatrace) and 0.1%
416 CHS (Cholesteryl Hemisuccinate, Anatrace) were added to the suspended membrane fraction.
417 After mixing for 2 hrs at 4 ^{o}C, detergent solubilized proteins were collected by ultra-centrifugation
418 at 190,000g for 30 mins. The supernatant was incubated with Anti-FLAG M2 Affinity agarose gel
419 (Sigma-Aldrich) for 4 hrs at 4 ^{o}C. The gel was washed with lysis buffer supplemented with 0.05%
420 DDM and 0.005% CHS. EGFR was eluted from the gel by lysis buffer containing 0.16 mg/ml
421 FLAG peptide, 0.05% DDM and 0.005% CHS.

422 Affinity-purified full-length EGFR was concentrated to 50 μM and treated with 100 μM
423 human EGF (Novoprotein) or TGF-α (Sino Biological) for at least 5 mins on ice, before further
424 fractionation by size-exclusion chromatography using a Superose 6 Increase 10/300 GL column
425 (GE Healthcare), with buffer containing 25 mM Tris pH 8.0, 150 mM NaCl, 2 μM Erlotinib, 0.02%
426 DDM and 0.002% CHS). The peak fractions containing ligand-bound dimeric EGFR were

427 collected and immediately used for reconstitution and the reconstituted receptor was applied to
428 EM grids. The fractions were also flash-frozen in liquid nitrogen and used for reconstitution at
429 later times. Freezing and thawing the fractions more than once leads to deterioration in cryo-EM
430 sample quality.
431

432 **Reconstitution of ligand-bound full-length EGFR complex**

433 **Reconstitution into lipid nanodiscs**

434 Purified protein complexes of EGFR bound to EGF or TGF-α were reconstituted into lipid
435 nanodiscs as previously described, with modifications (Ritchie et al., 2009). Briefly, membrane
436 scaffold proteins MSP1D1 and MSP1E3D1 were expressed and purified from *Escherichia coli*. 10
437 mg soybean PC (L-α-phosphatidylcholine (soy), Avanti) dissolved in chloroform was dried using
438 nitrogen stream and residual chloroform was further removed by desiccation under vacuum
439 overnight. Rehydration buffer containing 25 mM Tris pH 8.0, 150 mM NaCl, 0.2 mM DTT, 1%
440 DDM and 1% sodium cholate was added to the dried lipid and vortexed briefly, resulting in a semi-
441 clear solution at 10 mM lipid concentration. After 3 cycles of freezing (in liquid nitrogen)-and-
442 thawing (warm water bath), the lipid solution was bath sonicated for 1 hr. The resulting clear lipid
443 stock (10 mM) was immediately used for lipid nanodiscs assembly or stored at -80 ^{o}C in small
444 aliquots. Purified complexes of EGFR bound to EGF or TGF-α (~10 μM, solubilized in 0.02%
445 DDM and 0.002% CHS) was supplemented with 20 μM extra EGF or TGF-α to keep the ligand-
446 binding site on the receptor saturated, before it was mixed with the soybean PC stock and
447 membrane scaffold proteins at molar ratios of 1:6:240 (EGFR monomer:MSP1D1:soy PC) and
448 1:6:600 (EGFR monomer:MSP1E3D1:soy PC). The reaction mixture (~500 μl) was incubated on
449 ice for 30 mins, before 20 mg Bio-beads SM2 absorbent (Bio-Rad) were added to initiate the

450 removal of detergents from the mixture and to start the nanodiscs assembly process. The mixture
451 was incubated at 4 ^{o}C by continuous rotation. After 1 hour, a second batch of Bio-beads (20 mg)
452 was added to the mixture to replace the first batch, and the sample was incubated for overnight at
453 4 ^{o}C. A third and final batch of Bio-beads (20 mg) was added to replace the second batch, 2 hrs
454 before the reconstitution mixture was cleared by removing the Bio-beads through centrifugation.
455 After another centrifugation at 14,000 rpm for 10 mins, the reconstitution mixture was fractionated
456 by size-exclusion chromatograph using Superose 6 Increase 10/300 GL column (GE Healthcare)
457 in detergent-free buffer containing 25 mM Tris pH 8.0, 150 mM NaCl and 5 μM Erlotinib. The
458 peak fraction corresponding to dimeric EGFR:ligands complexes reconstituted in lipid nanodisc
459 was collected and analyzed by SDS-PAGE, before being applied to negative-stain and cryo-EM
460 grids.
461
462 **Reconstitution into amphipols**
463 Samples containing purified complexes of EGFR bound to EGF or TGF-α (~10 μM,
464 solubilized in 0.02% DDM and 0.002% CHS) were supplemented with 20 μM extra EGF or TGF-
465 α, before the amphipol A8-35 (Anatrace) was added to reach a mass ratio of 1:3 (EGFR:A8-35).
466 After incubation at 4 ^{o}C overnight, free amphipol molecules were removed by size-exclusion
467 chromatography using a Superose 6 Increase 10/300 GL column in detergent-free buffer
468 containing 25 mM Tris pH 8.0, 150 mM NaCl and 5 μM Erlotinib. Peak fractions corresponding
469 to the reconstituted EGFR:ligand complexes in amphipol were analyzed by SDS-PAGE, before
470 EM studies.
471
472 **Reconstitution into peptidiscs**

473 As for the reconstitution into amphipols, samples containing purified complexes of EGFR
474 bound to EGF or TGF-α (~10 μM, solubilized in 0.02% DDM and 0.002% CHS) were
475 supplemented with 20 μM extra EGF or TGF-α, before the addition of NSP_r peptide (Carlson et
476 al., 2018) to the samples at a mass ratio of 1:1.5 (EGFR:NSP_r peptide). After overnight incubation
477 at 4 ^{o}C, the sample was fractionated using a Superose 6 Increase 10/300 GL column (detergent-
478 free buffer: 25 mM Tris pH 8.0, 150 mM NaCl and 5 μM Erlotinib) and free NSP_r peptides were
479 removed during fractionation. Peak fractions corresponding to the reconstituted EGFR:ligand
480 complexes in peptidiscs were analyzed by SDS-PAGE, before applied to EM grids.
481

482 **Preparation of electron microscopy (EM) grids and data acquisition**

483 For negative-stain EM, a 5 μL sample of reconstituted full-length human EGFR (~0.05
484 mg/mL protein in 25 mM Tris pH 8.0, 150 mM NaCl and 2 μM Erlotinib) was placed on the
485 continuous-carbon side of a glow-discharged copper grid (Ted Pella, Redding, CA, USA), and the
486 excess sample was removed by wicking with filter paper after 1 min incubation. The bound
487 particles were stained by floating the grids on four consecutive 30 μL droplets of freshly prepped
488 2% uranyl formate solution and incubating the grids with each droplet for 15 s. The excess stain
489 was removed by blotting with filter paper and the grids were air-dried for at least 5 min. Images
490 of negative-stained full-length human EGFR reconstituted in lipid nanodisc, amphipol, and
491 peptidisc were recorded on a 4049 x 4096 pixels CMOS camera (TVIPS TemCam-F416) using
492 the automated Leginon data collection software (Suloway et al., 2005). Samples were imaged
493 using a Tecnai 12 transmission electron microscope (FEI, Hillsboro, OR, USA) at 120 keV with a
494 nominal magnification of 49,000x (2.18 Å calibrated pixel size at the specimen level) using a

495 defocus range of -0.8 to -1.5 μm. All data were acquired under low dose conditions, allowing a
496 dose at around 35 $e^-/Å^2$.

497 For cryo-EM, UltrAuFoil R 1.2/1.3 300 mesh grids (Electron Microscopy Sciences,
498 Q350AR13A) were rendered hydrophilic by pretreatment with a O_2/H_2 gas mixture using a Solarus
499 950 plasma cleaning system (Gatan, Inc., Pleasanton, CA) for 15s at 50W. After application of 3
500 μL sample of reconstituted full-length human EGFR (~0.5 mg/ml protein in 25 mM Tris pH 8.0,
501 150 mM NaCl and 2 uM Erlotinib), grids were plunge-frozen using a Leica EM GP (Leica
502 Microsystems Inc., Buffalo Grove, IL) with blotting time of 8 seconds, chamber temperature and
503 humidity of +4°C and 95%, and a liquid ethane temperature of -184°C. The grids were imaged
504 using a Titan Krios transmission electron microscope (Thermo Fisher Scientific, Waltham, MA)
505 operated at 300 kV. Images were recorded on a K3 camera (Gatan, Inc.) operated in super-
506 resolution counting mode with a physical pixel size of 0.54 Å. The detector was placed at the end
507 of a GIF Quantum energy filter (Gatan, Inc.), operated in zero-energy-loss mode with a slit width
508 of 20 eV. Due to strongly preferred particle orientations, all movies were collected at 40° tilt, at a
509 nominal magnification of 81,000x and a defocus range between -0.75 and -1.5 μm., using the
510 automated data collection software Latitude (Gatan, Inc.). The nominal dose per frame was ~1 e^-
511 $/Å^2$ and the total dose per movie is ~50 $e^-/Å^2$. For the full-length EGFR:EGF complex reconstituted
512 in peptidiscs, a total of 4,331 movies were collected. For the full-length EGFR:TGF-α complex
513 reconstituted in peptidiscs, a total of 7,406 movies were collected. For the full-length
514 EGFR(L834R):EGF complex reconstituted in peptidiscs, a total of 5,013 movies were collected.
515

516 **EM image processing**

517 For negative-stain EM images of full-length EGFR:EGF complexes reconstituted in lipidic
518 nanodiscs (MSP1E3D1), the initial image processing and classification steps were performed
519 using the Appion image processing environment (Lander et al., 2009). Particles were first selected
520 from micrographs using DoG Picker (Voss et al., 2009). The contrast transfer functions (CTFs) of
521 the micrographs were estimated using the CTFFIND3 (Mindell and Grigorieff, 2003). CTF
522 correction of the micrographs was performed by Wiener filter using ACE2 (Mallick et al., 2005).
523 A total of 112,408 particles were extracted using a 192 x 192 box size and binned by a factor of 2.
524 Each particle was normalized to remove pixels whose values were above or below 4.5 σ of the
525 mean pixel value using the XMIPP normalization program (Scheres et al., 2008). In order to
526 remove incorrectly selected protein aggregates or other artifacts, particles with extreme intensity
527 values were removed. The remaining 84,260 particles were subjected to 2D iterative reference-
528 free alignment and classification using IMAGIC multi-reference alignment (MRA) and a
529 topology-representing network classification (van Heel et al., 1996; Ogura et al., 2003), to produce
530 the 200 2D class averages.

531 For Cryo-EM images of full-length EGFR:EGF complex reconstituted in peptidiscs,
532 Extended Data Figure 3 shows a summary of the image processing workflow. In brief, movies
533 were motion-corrected using the MotionCor2 wrapper (Zheng et al., 2017) in Relion3 (Zivanov et
534 al., 2018). Corrected micrographs were imported to cryoSPARC v2 (Punjani and Fleet, 2020;
535 Punjani et al., 2017) for further processing. After local CTF determination using Patch-Based CTF
536 Estimation, a total of 4,234,635 particles from 4,431 micrographs were automatically selected
537 using template-free Blob_picker and were extracted with a box size of 320 x 320 pixels. Due to
538 strongly preferred particle orientations, a relatively low threshold value was used during the
539 particle picking to include particles with less contrast than the preferred views. The majority of the

540 initial 4,234,635 picks are from the background. After iterative 2D classification, most of the
541 background particles were removed based on 2D class averages. The remaining 1,587,872 particles
542 were further cleaned-up by ab-initio reconstruction into 4 classes. The resulting 3D maps showed
543 that 3 of the ab-initio reconstruction classes were mostly "junk" particles without recognizable
544 structural features, and the other class showed clear resemblance to the dimeric extracellular
545 module of EGFR. This clean class of particles (814,089) were selected and subjected to a final
546 round of 2D classification. After visual inspection of the final 2D class averages, 765,883 particles
547 were retained and ab-initio reconstruction was performed to generate the initial model, which was
548 used as the reference for a round of homogeneous refinement as well as a non-uniform refinement
549 (Punjani et al., 2019). Based on the gold-standard Fourier shell correlation (FSC) 0.143 criterion,
550 the overall resolution of the reconstructions from homogeneous refinement and non-uniform
551 refinement are determined as 3.0 Å and 2.9 Å, respectively. After visual inspection of both
552 reconstructions, the map generated from non-uniform refinement with better quality was chosen
553 for further analysis and used as the reference in the subsequent round of heterogeneous refinement
554 (3D classification) using the 765,883 particles, resulting in 10 classes. A final round of
555 homogeneous refinement and non-uniform refinement were performed for each of the resulting 10
556 classes from the heterogeneous refinement run. After visual inspection, the 10 maps resulting from
557 non-uniform refinement were chosen for model building and refinement for each class. The overall
558 resolution of these 10 maps ranges between 3.1 Å and 3.7 Å.

559 Cryo-EM data processing of full-length EGFR(L834R):EGF complex and full-length
560 EGFR:TGF-α complex reconstituted in peptidiscs followed similar procedures described above.
561 For full-length EGFR(L834R):EGF complex, after local CTF estimation by Patch-Based CTF
562 Estimation, a total of 3,250,237 particles were automatically picked by Blob_Picker. After iterative

563 2D classification and ab-initio reconstruction, background and spurious particles were removed.
564 987,367 cleaned particles were subjected to homogeneous as well as non-uniform refinement,
565 using the initial model generated from ab-initio reconstruction as reference. After visual inspection,
566 the map generated by non-uniform refinement (overall resolution: 3.0 Å) was chosen as the
567 reference for the subsequent heterogeneous refinement, resulting in 10 classes. A final round of
568 non-uniform refinement was performed for each of the 10 classes and the resulting 10 maps were
569 used for model building and refinement. The resolutions of these 10 maps were determined to be
570 in the range of 3.2-4.2 Å.

571 For the full-length EGFR:TGF-α complex, a total of 6,846,834 particles were initially
572 picked by Blob_Picker. Iterative 2D classification and ab-initio reconstruction were performed to
573 remove background and spurious particles, and 1,019,024 cleaned particles were retained. These
574 cleaned particles were subjected to homogeneous refinement and non-uniform refinement. The
575 map resulting from non-uniform refinement was at an overall resolution of 3.0 Å and it was used
576 as the reference for the subsequent heterogeneous refinement, resulting in 10 classes. A final round
577 of non-uniform refinement was performed for each of the 10 classes and the resulting 10 maps
578 were determined at resolutions between 3.2 Å and 4.2 Å. These maps were used for model building
579 and refinement.

580 3D variability analysis, local resolution estimation, and B-factor sharpening of all refined
581 maps were performed using the cryoSPARC v2.

582

583 **Model building and refinement**

584 To generate models for different conformations of the EGFR(WT):EGF and
585 EGFR(L834R):EGF complexes, a crystal structure of the isolated EGFR extracellular module

586 bound to EGF (PDB code: 3NJP) was split into four fragments: two heads and two legs. The two
587 heads contain EGFR residues 1-500 and EGF residues 5-51 from each side of the dimeric structure,
588 respectively. The two legs contain EGFR residues 501-614 from each EGFR molecule,
589 respectively. These four fragments were fit into the reconstructed cryo-EM maps of the
590 EGFR(WT):EGF and EGFR(L834R):EGF complexes, using UCSF Chimera (Pettersen et al.,
591 2004). UCSF chimera-fitted models were subjected to one round of real-space refinement (Afonine
592 et al., 2018) in *Phenix* (Liebschner et al., 2019). To prevent overfitting due to the lower local
593 resolution in the map region corresponding to the flexible Domain IV legs, the coordinates of the
594 flexible Domain IV leg was not refined against the cryo-EM maps. The models were manually
595 adjusted in *Coot* (Emsley et al., 2010), followed by iterative rounds of real-space refinement in
596 *Phenix* and manual fitting in *Coot*. Model quality was assessed using the comprehensive model
597 validation tools from *Phenix* (Extended Data Table 1). All maps and models were visualized using
598 UCSF Chimera and PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger,
599 LLC).

600 Similar procedures were followed to generate models for the different conformations of
601 the EGFR:TGF-α complex, except that a crystal structure of the truncated EGFR extracellular
602 domains bound to TGF-α (PDB code: 1MOX) was used to generate the fragments of the heads for
603 the initial model fitting in UCSF Chimera.

604

605 **Molecular dynamics simulations**

606 **System preparation for molecular dynamics simulations of extracellular module**

607 We built a system comprising of the extracellular module bound to EGF starting from the
608 crystal structure (PDB code: 3NJP (Lu et al., 2010)). The N-terminal ends of each EGFR subunit

609 are the true termini and were left uncapped. The C-terminal ends of EGFR were capped with a
610 methyl group. In the crystal structure four residues are missing from the N-terminal ends of the
611 either ligand, and two residues are missing from the C-terminal ends. We initiated one trajectory
612 starting with these structures while capping the N-terminal ends with acetyl groups and the C-
613 terminal ends with methyl groups. For three other trajectories, we built in the missing residues
614 with PyMOL (Schrödinger, LLC., 2015) and left the termini uncapped in these structures. These
615 systems were solvated such that the water layer extended for 15 Å beyond the longest dimension
616 of the protein, and ions were added such that the ionic strength was 0.15 M with VMD (Humphrey
617 et al., 1996).

618 We built a second system of EGFR bound to TGFα. For this structure, EGFR was taken
619 from the crystal structure 3NJP (Lu et al., 2010), and TGFα was taken from the structure of
620 truncated EGFR bound to TGFα (PDB code: 1MOX (Garrett et al., 2002)). Similar to the system
621 of EGFR bound to EGF, the EGFR subunits were left uncapped at the N-terminal ends and capped
622 at the C-terminal ends. The ligands were uncapped in all the simulations. These systems were
623 solvated such that the water layer extended for 15 Å beyond the longest dimension of the protein
624 and ions were added such that the ionic strength was 0.15 M with VMD (Humphrey et al., 1996).

625 **Minimization, equilibration and production protocols for simulations of extracellular**
626 **module**

627 The energies of each system were minimized, first for a 1000 steps while holding protein
628 atoms fixed, and then for a 1000 steps while allowing all the atoms to move. The systems were
629 then equilibrated, first for 1 ns while holding all the heavy atoms fixed, then for 1 ns while holding
630 the protein backbone atoms fixed, and then for 1 ns without any constraints. The four trajectories

631 for each system were then allowed to run for between ~350 and ~550 ns, for a total simulation
632 time of ~ 1.8 μs for EGFR bound to EGF and ~ 2.0 μs for EGFR bound to TGF-α. All equilibration
633 and production simulations were carried out at a temperature of 300 K and pressure of 1 atm.
634

635 **System preparation for molecular dynamics simulations of transmembrane helices in POPC**
636 **membrane**

637 We built two systems – the first of the two transmembrane helices (residues 610-653)
638 associating at the N-terminal ends starting from model 1 of the NMR structure (PDB code: 5LV6
639 (Bocharov et al., 2017)), and the second of the two transmembrane helices (residues 614-653)
640 associating at the C-terminal ends starting from model 6 of the NMR structure (PDB code: 2M0B
641 (Bocharov et al., 2016)), We used the Membrane Builder module on the CHARMM-GUI website
642 (Jo et al., 2008; Wu et al., 2014) to insert each protein system into a lipid bilayer comprising of
643 neutral POPC lipids. The x- and y- dimensions of the bilayer was set to 100 Å each. This lipid
644 bilayer comprised of 148 lipids in the upper leaflet, and 146 lipids in the lower leaflet in the N-
645 terminal dimer system, and 142 lipids in each leaflet in the C-terminal dimer system. The N-termini
646 of each protein chain were capped with acetyl groups and the C-termini were capped with methyl
647 groups. The system was solvated such that the solvent layer extended to 20 Å beyond the lipid
648 bilayer on either side, and ions were added such that the ionic strength was 0.1 M, using VMD
649 (Humphrey et al., 1996). The final edge dimensions of these systems were ~103 Å x 103 Å x 118
650 Å.

651 **Minimization, equilibration and production protocols**

652 For each system, first the energy was minimized while holding the protein, waters, ions,
653 and lipid heads fixed, and allowing only the lipid tails to move, for 1000 steps, then while holding

654 the protein and lipid heads fixed for 5000 steps, then while holding only the protein fixed for 1000
655 steps, and finally for a 1000 steps while allowing all the atoms to move. This system was then
656 equilibrated using molecular dynamics for 5 ns while holding the protein fixed, to allow the lipids
657 to relax about the protein.

658 For the N-terminal dimer system, a longer trajectory of 175 ns then was generated from
659 this equilibrated structure, during which all the atoms were allowed to move. During the course
660 of this simulation, the N-terminal ends of the linkers (residues 610-613), which are disordered in
661 this structure, interacted frequently with each other. In the full-length structure, these residues are
662 a part of the ordered region in Domain IV of the extracellular module and would not be available
663 to form such interactions. We therefore deleted these 4 residues, such that the protein chains started
664 at Thr 614, and capped these N-terminal ends with acetyl groups. The energy of this system was
665 minimized with the same multi-step protocol as before and used to initiate 25 independent
666 molecular dynamics trajectories.

667 For the simulations starting from the C-terminal dimer, in the initial structure the gap
668 between the N-terminal ends is too small to accommodate lipid tails. Additionally, the helices tilt
669 away from each other such that the distance between the Thr 627 residues at the N-terminal ends
670 of the helices changes from ~18 Å to ~23 Å, in order to optimally pack the hydrophobic residues
671 of the protein into the membrane (Killian and Nyholm, 2006). In order the maintain the N-terminal
672 ends of the helices, a multi-step equilibration protocol was used. First, the distance between the
673 Thr 627 residues of either subunit was increased from ~18 Å to ~23 Å over 5 ns. During this
674 phase, restraints applied to the protein backbone atoms of residues 620-626 to maintain the Ile619-
675 Ala623 association and the 3_{10}-helical turn at the N-terminal ends of the helices. The system was
676 equilibrated for another 25 ns with these restraints applied. The system was then run for 20 ns

677 without restraints, to ensure that the N-terminal ends were stable. This structure was then used to
678 start 25 productions runs which ran for 100ns each for a total of 2.5µs. All equilibration and
679 production simulations were carried out at a temperature of 300 K and pressure of 1 atm.

680 **Simulation parameters**

681 The NAMD package was used to run the minimization and equilibration simulations
682 (Phillips et al., 2005) with the CHARMM36m force field (Huang et al., 2017). The velocity Verlet
683 algorithm was used to calculate the trajectories of the atoms. A time step of 2 fs was used. Particle
684 Mesh Ewald was used to calculate long-range electrostatic interactions, with a maximum space of
685 1 Å between grid points (Darden et al., 1993). Long-range electrostatics were updated at every
686 time step. Van der Waals interactions were truncated at 12 Å. Hydrogen atoms bonded to the heavy
687 atoms were constrained using the ShakeH algorithm (Ryckaert et al., 1977). The temperature was
688 controlled with Langevin dynamics with a damping coefficient of 1/ps, applied only to non-
689 hydrogen atoms. Pressure was controlled by the Nose-Hoover method with the Langevin piston,
690 with an oscillation period of 200 fs and a damping time scale of 50 fs (Feller et al., 1995; Martyna
691 et al., 1994).

692 Distance and angle variations over the course of the simulations were calculated with the
693 AmberTools20 package (Case et al., 2020).

694 **Free energy calculations with umbrella sampling**

695 Umbrella sampling simulations (Torrie and Valleau, 1977) were carried out to measure the
696 free energy change as the distance between the Thr614 residues was decreased from 27 Å to 8 Å
697 when the transmembrane helices are in the N-terminal dimer conformation. We performed three
698 umbrella sampling calculations, starting from initial paths generated from three different steered

699 molecular dynamics simulations. For each umbrella sampling calculation, 20 windows of
700 simulations were run, where each window was separated by 1 Å, and the starting structure in each
701 window had distances between Thr614 distances ranging from 8 Å to 27 Å. A force constant of 1
702 $kcal.mol^{-1}.Å^{-2}$ was used to maintain the distance between Thr 614 residues in each window, using
703 the colvars module (Fiorin et al., 2013). Two of the umbrella sampling simulations converged over
704 60 ns while the third converged over 90 ns. The biased probability distribution thus generated was
705 reweighted and the free energy change was calculated using the Weighted Histogram Analysis
706 Method (Kumar et al., 1992) using the implementation provided by Alan Grossfield (Grossfield).
707

708 **Western-blot analysis of EGFR autophosphorylation with purified EGFR proteins**

709 **Expression and purification of FLAG-tagged full-length human EGFR from HEK293FT**
710 **cells**

711 A cDNA fragment encoding full-length human EGFR (UniProt accession no. P00533) was
712 cloned into the pEGFP-N1 plasmid (Clontech, Mountain View, CA). To facilitate affinity
713 purification, a FLAG-tag (with sequence DYKDDDDK) was inserted after the signal sequence of
714 EGFR. To monitor protein expression and to reduce non-specific degradation of the C-terminal
715 tail of EGFR, EGFP was fused after the C-terminus of EGFR. HEK293FT cells were transfected
716 with the plasmid using FuGENE6 (Promega). After 24 hours, 10 mM sodium butyrate was added
717 to the culture medium. The cells were harvested from culture dishes 36 hours after transfection
718 and were resuspended in lysis buffer (25 mM Tris pH 8.0, 150 mM NaCl, and cOmplete Protease
719 Inhibitor Cocktail). To extract EGFR from the membrane, 1% DDM and 0.1% CHS were added
720 to the cell suspension. After mixing for 1 hour at 4 ^{o}C, detergent solubilized proteins were collected
721 by ultra-centrifugation at 190,000g for 40 mins. The supernatant was incubated with Anti-FLAG

722 M2 Affinity agarose resin for at least 4 hours at 4 ^{o}C and applied to a Micro Bio-Spin column. The
723 resin was washed with lysis buffer supplemented with 0.05% DDM and 0.005% CHS. EGFR was
724 eluted from the resin with lysis buffer containing 0.16 mg/ml FLAG peptide, 0.05% DDM and
725 0.005% CHS. Protein concentration was estimated by the band intensities on a SDS-PAGE gel
726 stained with Coomassie Brilliant Blue and by UV absorbance at 280 nm measured using NanoDrop
727 (ThermoFisher).
728
729 **Western-blot assays using purified EGFR protein**
730 ~1 μM purified EGFR protein prepared as described above, was mixed with ~16 μM EGF
731 or TGF-α in a buffer containing 25 mM Tris pH 8.0, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM sodium
732 orthovanadate, and 0.2 mM DTT. The mixture was incubated for 5 minutes, and phosphorylation
733 reactions were initiated by addition of ATP to a final concentration of 1 mM. Reactions were
734 typically carried out at 25 °C. At various time points, aliquots of the reaction solution were mixed
735 with quenching solution, SDS-PAGE loading buffer supplemented with 20 mM EDTA, to achieve
736 final concentration of 1% SDS and 10 mM EDTA in the quenched solutions. The samples
737 quenched at different time points were run on two 12% acrylamide SDS-PAGE gels at 250 V, 400
738 mA for 35 min at room temperature.
739 The first SDS-PAGE gel was stained using Coomassie Brilliant Blue to monitor the total
740 amount of EGFR protein loaded in each lane. The second SDS-PAGE gel was transferred to PVDF
741 membrane using a semi-dry transfer apparatus TRANS-BLOT (Bio-Rad) at 25 V, 150 mA for 40
742 minutes at room temperature, in transfer buffer containing 25 mM Tris, 192 mM glycine, 0.02%
743 SDS, and 10% Methanol. The PVDF membranes were blocked for 1 hour at room temperature in
744 5% (w/v) dry milk dissolved in Tris-buffered saline with 0.1% Tween-20 (TBST). A pan-

745 phosphotyrosine antibody (pY20, catalog no. sc-12351, Santa Cruz Biotechnology) was diluted
746 (1:2000) with 5% (w/v) dry milk in TBST and incubated with the membranes at 4°C overnight.
747 The membranes were then washed 3 times with TBST, and incubated with horseradish peroxidase
748 (HRP)-conjugated anti-mouse secondary antibody (catalog no. 7074, Cell Signaling Technology)
749 at a 1:3000 dilution in TBST with 5% dry milk at room temperature for 1 hour. The membranes
750 were then washed 3 times with TBST, incubated with WesternBright® Quantum®
751 chemiluminescent reagents (Advansta) for 2 minutes, and were imaged using ImageLab (Bio-Rad).
752

753 **Cell-based western-blot assays for EGFR autophosphorylation and Erk phosphorylation**

754 **EGFR autophosphorylation in cells**

755 N-terminal Flag-tagged full-length human EGFR was cloned into the pEGFP-N1 plasmid.
756 HEK293T cells were transiently transfected with EGFR using polyethylenimine (PEI) in a 6-well
757 plate format. 24 hours after transfection, cells were serum starved overnight (~12 hours) in serum
758 free medium. For stimulation, cells were incubated with fresh medium without ligand, with EGF
759 (16.7 nM), or with TGF-α (16.7 nM) for 1 hour at 4 ^{o}C (on ice). Following ligand stimulation,
760 cells were pelleted and then immediately lysed in ice-cold cell lysis buffer (catalog no. 9803, Cell
761 Signaling Technology) supplemented with cOmplete Protease Inhibitor Cocktail and PhosSTOP
762 phosphatase inhibitors (Roche) for 30 minutes on ice. Cell lysates were clarified by centrifugation
763 and western-blot analyses were carried out as described above for the purified EGFR proteins. A
764 pan-phosphotyrosine antibody (pY20) was used to detect EGFR autophosphorylation. An anti-
765 FLAG M2 antibody (catalog no. F1804, Sigma-Aldrich) was used to monitor the total amount of
766 EGFR protein in the cell lysate loaded to each lane.
767

768 **Erk phosphorylation in cells**

769 N-terminal Flag-tagged full-length human EGFR was cloned into the pEGFP-N1 plasmid.
770 The mutations were introduced by QuikChange site-directed mutagenesis (Stratagene). 24 hours
771 after transfection, HEK293T cells transiently transfected with EGFR and EGFR variants were
772 serum starved overnight (~12 hours) in serum free medium. For stimulation, cells were incubated
773 with fresh medium without ligand, with EGF (16.7 nM), or with TGF-α (16.7 nM) for 5 minutes
774 at 37 ^{o}C. Following ligand stimulation, cells were immediately placed on ice and then lysed in ice-
775 cold cell lysis buffer supplemented with cOmplete Protease Inhibitor Cocktail and PhosSTOP
776 phosphatase inhibitors for 30 minutes on ice. Cell lysates were clarified by centrifugation and
777 western-blot analyses were carried out as described above. An anti-phospho-Erk antibody (catalog
778 no. 4370, Cell Signaling Technology) is used to detect phosphorylation of p44 and p42 MAP
779 kinase (Erk1 and Erk2). An anti-FLAG M2 antibody was used to monitor the total amount of
780 EGFR protein in the cell lysate loaded to each lane.
781 All western blots were reproduced at least three times and the representative ones were presented.

782 Acknowledgments

783 We thank the Cell Culture Facility at the University of California, Berkeley, for insect cell and
784 mammalian cell cultures. We thank the Eric Gouaux group at the Vollum Institute at the Oregon
785 Health & Science University, for the pEG_BacMam plasmid. This research was in part supported
786 by the National Cryo-EM Facility of the National Cancer Institute at the Frederick National
787 Laboratory for Cancer Research, under contract HSSN261200800001E. We also acknowledge the
788 Berkeley Bay Area Cryo-EM Facility and the HHMI Janelia Cryo-EM Facility. Molecular
789 graphics and analyses performed with UCSF Chimera, developed by the Resource for

790 Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with
791 support from NIH P41-GM103311. This work used the Extreme Science and Engineering
792 Discovery Environment (XSEDE), which is supported by National Science Foundation grant ACI-
793 1548562. This work is in part supported by an award from the Canada Excellence Research Chair
794 program (to S.S.) and by a philanthropic gift to the VGH and UBC Hospital Foundation (to S.S.).

795 Author Contributions

796 Y.H., S.S., and J.K. conceived the project; Y.H. designed and performed experiments; Y.H., J.O.,
797 and A.M. performed EM experiments; Y.H. analyzed data; Y.H., D.K., and K.M. performed
798 molecular dynamics simulations; S.S. and J.K. supervised the project; Y.H. and J.K. drafted the
799 manuscript; Y.H., S.S., and J.K. edited the manuscript. All authors commented on the manuscript.

800 Competing interests

801 J.K. is a cofounder of Nurix Therapeutics and is on the scientific advisory boards of Carmot and
802 Revolution Medicine. S.S. is Founder and Chief Executive Officer of Gandeeva Therapeutics Inc.
803 S.S. is a reviewing editor for eLife.

804 References

805 Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and
806 Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta
807 Crystallogr. D Struct. Biol. *74*, 531–544.
808
809 Ahmed, F., Paul, M., and Hristova, K. (2020). The biophysical basis of receptor tyrosine kinase
810 ligand functional selectivity: Trk-B case study. Biochem J.
811
812 Angiulli, G., Dhupar, H.S., Suzuki, H., Wason, I.S., Duong Van Hoa, F., and Walz, T. (2020).
813 New approach for membrane protein reconstitution into peptidiscs and basis for their adaptability
814 to different proteins. Elife *9*.

815
816 Arkhipov, A., Shan, Y., Das, R., Endres, N.F., Eastwood, M.P., Wemmer, D.E., Kuriyan, J., and
817 Shaw, D.E. (2013). Architecture and membrane interactions of the EGF receptor. Cell *152*, 557–
818 569.
819
820 Arteaga, C.L., and Engelman, J.A. (2014). ERBB receptors: from oncogene discovery to basic
821 science to mechanism-based cancer therapeutics. Cancer Cell *25*, 282–303.
822
823 Bayburt, T.H., and Sligar, S.G. (2010). Membrane protein assembly into Nanodiscs. FEBS Lett
824 *584*, 1721–1727.
825
826 Bocharov, E.V., Lesovoy, D.M., Pavlov, K.V., Pustovalova, Y.E., Bocharova, O.V., and Arseniev,
827 A.S. (2016). Alternative packing of EGFR transmembrane domain suggests that protein-lipid
828 interactions underlie signal conduction across membrane. Biochim Biophys Acta *1858*, 1254–
829 1261.
830
831 Bocharov, E.V., Bragin, P.E., Pavlov, K.V., Bocharova, O.V., Mineev, K.S., Polyansky, A.A.,
832 Volynsky, P.E., Efremov, R.G., and Arseniev, A.S. (2017). The conformation of the epidermal
833 growth factor receptor transmembrane domain dimer dynamically adapts to the local membrane
834 environment. Biochemistry *56*, 1697–1705.
835
836 Brand, T.M., Iida, M., and Wheeler, D.L. (2011). Molecular mechanisms of resistance to the EGFR
837 monoclonal antibody cetuximab. Cancer Biol Ther *11*, 777–792.
838
839 Cai, W.-Q., Zeng, L.-S., Wang, L.-F., Wang, Y.-Y., Cheng, J.-T., Zhang, Y., Han, Z.-W., Zhou,
840 Y., Huang, S.-L., Wang, X.-W., et al. (2020). The latest battles between EGFR monoclonal
841 antibodies and resistant tumor cells. Front Oncol *10*, 1249.
842
843 Carlson, M.L., Young, J.W., Zhao, Z., Fabre, L., Jun, D., Li, J., Li, J., Dhupar, H.S., Wason, I.,
844 Mills, A.T., et al. (2018). The Peptidisc, a simple method for stabilizing membrane proteins in
845 detergent-free solution. Elife *7*.
846
847 Case, D.A., Belfon, K., Ben-Shalom, I.Y., Brozell, S.R., Cerutti, D.S., Cheatham, III, T.E.,
848 Cruzeiro, V.W.D., Darden, T.A., Duke, R.E., Giambasu, G., et al. (2020). AMBER 20.
849
850 Chong, C.R., and Jänne, P.A. (2013). The quest to overcome resistance to EGFR-targeted therapies
851 in cancer. Nat Med *19*, 1389–1400.
852
853 Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for
854 Ewald sums in large systems. J. Chem. Phys. *98*, 10089.
855
856 Dawson, J.P., Berger, M.B., Lin, C.-C., Schlessinger, J., Lemmon, M.A., and Ferguson, K.M.
857 (2005). Epidermal growth factor receptor dimerization and activation require ligand-induced
858 conformational changes in the dimer interface. Mol Cell Biol *25*, 7734–7742.
859
860 Diwanji, D., Trenker, R., Thaker, T.M., Wang, F., Agard, D.A., Verba, K.A., and Jura, N. (2021).

861 Structures of the active HER2/HER3 receptor complex reveal dynamics at the dimerization
862 interface induced by binding of a single ligand. BioRxiv.
863
864 Doerner, A., Scheck, R., and Schepartz, A. (2015). Growth Factor Identity Is Encoded by Discrete
865 Coiled-Coil Rotamers in the EGFR Juxtamembrane Region. Chem Biol *22*, 776–784.
866
867 Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot.
868 Acta Crystallogr D Biol Crystallogr *66*, 486–501.
869
870 Endres, N.F., Das, R., Smith, A.W., Arkhipov, A., Kovacs, E., Huang, Y., Pelton, J.G., Shan, Y.,
871 Shaw, D.E., Wemmer, D.E., et al. (2013). Conformational coupling across the plasma membrane
872 in activation of the EGF receptor. Cell *152*, 543–556.
873
874 Feller, S.E., Zhang, Y., Pastor, R.W., and Brooks, B.R. (1995). Constant pressure molecular
875 dynamics simulation: The Langevin piston method. J Chem Phys *103*, 4613–4621.
876
877 Fiorin, G., Klein, M.L., and Hénin, J. (2013). Using collective variables to drive molecular
878 dynamics simulations. Mol Phys *111*, 3345–3362.
879
880 Freed, D.M., Bessman, N.J., Kiyatkin, A., Salazar-Cavazos, E., Byrne, P.O., Moore, J.O., Valley,
881 C.C., Ferguson, K.M., Leahy, D.J., Lidke, D.S., et al. (2017). EGFR ligands differentially stabilize
882 receptor dimers to specify signaling kinetics. Cell *171*, 683–695.e18.
883
884 Garrett, T.P.J., McKern, N.M., Lou, M., Elleman, T.C., Adams, T.E., Lovrecz, G.O., Zhu, H.-J.,
885 Walker, F., Frenkel, M.J., Hoyne, P.A., et al. (2002). Crystal structure of a truncated epidermal
886 growth factor receptor extracellular domain bound to transforming growth factor alpha. Cell *110*,
887 763–773.
888
889 Goehring, A., Lee, C.-H., Wang, K.H., Michel, J.C., Claxton, D.P., Baconguis, I., Althoff, T.,
890 Fischer, S., Garcia, K.C., and Gouaux, E. (2014). Screening and large-scale expression of
891 membrane proteins in mammalian cells for structural studies. Nat Protoc *9*, 2574–2585.
892
893 Gohon, Y., Giusti, F., Prata, C., Charvolin, D., Timmins, P., Ebel, C., Tribet, C., and Popot, J.-L.
894 (2006). Well-defined nanoparticles formed by hydrophobic assembly of a short and polydisperse
895 random terpolymer, amphipol A8-35. Langmuir *22*, 1281–1290.
896
897 Grossfield, A. An implementation of WHAM: the Weighted Histogram Analysis Method Version
898 2.0.9.
899
900 van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. (1996). A new generation of
901 the IMAGIC image processing system. J Struct Biol *116*, 17–24.
902
903 Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., Grubmüller, H., and
904 MacKerell, A.D. (2017). CHARMM36m: an improved force field for folded and intrinsically
905 disordered proteins. Nat Methods *14*, 71–73.
906

907 Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD – Visual Molecular Dynamics. J. of
908 Mol. Graph. *14*, 33–38.
909
910 Jo, S., Kim, T., Iyer, V.G., and Im, W. (2008). CHARMM-GUI: a web-based graphical user
911 interface for CHARMM. J Comput Chem *29*, 1859–1865.
912
913 Jones, J.T., Akita, R.W., and Sliwkowski, M.X. (1999). Binding specificities and affinities of egf
914 domains for ErbB receptors. FEBS Lett *447*, 227–231.
915
916 Jura, N., Endres, N.F., Engel, K., Deindl, S., Das, R., Lamers, M.H., Wemmer, D.E., Zhang, X.,
917 and Kuriyan, J. (2009). Mechanism for activation of the EGF receptor catalytic domain by the
918 juxtamembrane segment. Cell *137*, 1293–1307.
919
920 Killian, J.A., and Nyholm, T.K.M. (2006). Peptides in lipid bilayers: the power of simple models.
921 Curr Opin Struct Biol *16*, 473–479.
922
923 Kovacs, E., Zorn, J.A., Huang, Y., Barros, T., and Kuriyan, J. (2015). A structural perspective on
924 the regulation of the epidermal growth factor receptor. Annu Rev Biochem *84*, 739–764.
925
926 Kumar, S., Rosenberg, J.M., Bouzida, D., Swendsen, R.H., and Kollman, P.A. (1992). THE
927 weighted histogram analysis method for free-energy calculations on biomolecules. I. The method.
928 J Comput Chem *13*, 1011–1021.
929
930 Lander, G.C., Stagg, S.M., Voss, N.R., Cheng, A., Fellmann, D., Pulokas, J., Yoshioka, C., Irving,
931 C., Mulder, A., Lau, P.-W., et al. (2009). Appion: an integrated, database-driven pipeline to
932 facilitate EM image processing. J Struct Biol *166*, 95–102.
933
934 Leahy, D.J. (2004). Structure and function of the epidermal growth factor (EGF/ErbB) family of
935 receptors. Adv Protein Chem *68*, 1–27.
936
937 Lemmon, M.A., Schlessinger, J., and Ferguson, K.M. (2014). The EGFR family: not so
938 prototypical receptor tyrosine kinases. Cold Spring Harb Perspect Biol *6*, a020768.
939
940 Li, J., Choi, E., Yu, H., and Bai, X.-C. (2019). Structural basis of the activation of type 1 insulin-
941 like growth factor receptor. Nat Commun *10*, 4567.
942
943 Liebschner, D., Afonine, P.V., Baker, M.L., Bunkóczi, G., Chen, V.B., Croll, T.I., Hintze, B.,
944 Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using
945 X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct. Biol.
946 *75*, 861–877.
947
948 Liu, P., Cleveland, T.E., Bouyain, S., Byrne, P.O., Longo, P.A., and Leahy, D.J. (2012). A single
949 ligand is sufficient to activate EGFR dimers. Proc Natl Acad Sci U S A *109*, 10861–10866.
950
951 Lowder, M.A., Doerner, A.E., and Schepartz, A. (2015). Structural Differences between Wild-
952 Type and Double Mutant EGFR Modulated by Third-Generation Kinase Inhibitors. J Am Chem

953 Soc *137*, 6456–6459.
954
955 Lu, C., Mi, L.-Z., Grey, M.J., Zhu, J., Graef, E., Yokoyama, S., and Springer, T.A. (2010).
956 Structural evidence for loose linkage between ligand binding and kinase activation in the epidermal
957 growth factor receptor. Mol Cell Biol *30*, 5432–5443.
958
959 Mallick, S.P., Carragher, B., Potter, C.S., and Kriegman, D.J. (2005). ACE: automated CTF
960 estimation. Ultramicroscopy *104*, 8–29.
961
962 Martyna, G.J., Tobias, D.J., and Klein, M.L. (1994). Constant pressure molecular dynamics
963 algorithms. J. Chem. Phys. *101*, 4177.
964
965 Mendrola, J.M., Berger, M.B., King, M.C., and Lemmon, M.A. (2002). The single transmembrane
966 domains of ErbB receptors self-associate in cell membranes. J Biol Chem *277*, 4704–4712.
967
968 Mi, L.-Z., Lu, C., Li, Z., Nishida, N., Walz, T., and Springer, T.A. (2011). Simultaneous
969 visualization of the extracellular and cytoplasmic domains of the epidermal growth factor receptor.
970 Nat Struct Mol Biol *18*, 984–989.
971
972 Miloso, M., Mazzotti, M., Vass, W.C., and Beguinot, L. (1995). SHC and GRB-2 are constitutively
973 by an epidermal growth factor receptor with a point mutation in the transmembrane domain. J Biol
974 Chem *270*, 19557–19562.
975
976 Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen
977 tilt in electron microscopy. J Struct Biol *142*, 334–347.
978
979 Mozumdar, D., Quach, K., Doerner, A., and Schepartz, A. (2020). Coiled coil control of diverse
980 EGFR functions. BioRxiv.
981
982 Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J.-H., Saito, K., Sakamoto, A.,
983 Inoue, M., Shirouzu, M., et al. (2002). Crystal structure of the complex of human epidermal growth
984 factor and receptor extracellular domains. Cell *110*, 775–787.
985
986 Ogura, T., Iwasaki, K., and Sato, C. (2003). Topology representing network enables highly
987 accurate classification of protein images taken by cryo electron-microscope without masking. J
988 Struct Biol *143*, 185–200.
989
990 Pahuja, K.B., Nguyen, T.T., Jaiswal, B.S., Prabhash, K., Thaker, T.M., Senger, K., Chaudhuri, S.,
991 Kljavin, N.M., Antony, A., Phalke, S., et al. (2018). Actionable activating oncogenic
992 ERBB2/HER2 transmembrane and juxtamembrane domain mutations. Cancer Cell *34*, 792–
993 806.e5.
994
995 Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and
996 Ferrin, T.E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis.
997 J Comput Chem *25*, 1605–1612.
998

999 Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D.,
1000 Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J Comput Chem *26*,
1001 1781–1802.
1002
1003 Punjani, A., and Fleet, D.J. (2020). 3D Variability Analysis: Directly resolving continuous
1004 flexibility and discrete heterogeneity from single particle cryo-EM images. BioRxiv.
1005
1006 Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for
1007 rapid unsupervised cryo-EM structure determination. Nat Methods *14*, 290–296.
1008
1009 Punjani, A., Zhang, H., and Fleet, D.J. (2019). Non-uniform refinement: Adaptive regularization
1010 improves single particle cryo-EM reconstruction. BioRxiv.
1011
1012 Red Brewer, M., Yun, C.-H., Lai, D., Lemmon, M.A., Eck, M.J., and Pao, W. (2013). Mechanism
1013 for activation of mutated epidermal growth factor receptors in lung cancer. Proc Natl Acad Sci U
1014 S A *110*, E3595–E3604.
1015
1016 Reeves, P.J., Callewaert, N., Contreras, R., and Khorana, H.G. (2002). Structure and function in
1017 rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation
1018 by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable
1019 mammalian cell line. Proc Natl Acad Sci U S A *99*, 13419–13424.
1020
1021 Ritchie, T.K., Grinkova, Y.V., Bayburt, T.H., Denisov, I.G., Zolnerciks, J.K., Atkins, W.M., and
1022 Sligar, S.G. (2009). Chapter 11 - Reconstitution of membrane proteins in phospholipid bilayer
1023 nanodiscs. Meth Enzymol *464*, 211–231.
1024
1025 Ronan, T., Macdonald-Obermann, J.L., Huelsmann, L., Bessman, N.J., Naegle, K.M., and Pike,
1026 L.J. (2016). Different epidermal growth factor receptor (EGFR) agonists produce unique
1027 signatures for the recruitment of downstream signaling proteins. J Biol Chem *291*, 5528–5540.
1028
1029 Ryckaert, J.-P., Ciccotti, G., and Berendsen, H.J.. (1977). Numerical integration of the cartesian
1030 equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys
1031 *23*, 327–341.
1032
1033 Scheck, R.A., Lowder, M.A., Appelbaum, J.S., and Schepartz, A. (2012). Bipartite tetracysteine
1034 display reveals allosteric control of ligand-specific EGFR activation. ACS Chem Biol *7*, 1367–
1035 1376.
1036
1037 Scheres, S.H.W., Núñez-Ramírez, R., Sorzano, C.O.S., Carazo, J.M., and Marabini, R. (2008).
1038 Image processing for electron microscopy single-particle analysis using XMIPP. Nat Protoc *3*,
1039 977–990.
1040
1041 Schrödinger, LLC., P. (2015). PyMOL (Schrödinger, LLC.).
1042
1043 Shan, Y., Eastwood, M.P., Zhang, X., Kim, E.T., Arkhipov, A., Dror, R.O., Jumper, J., Kuriyan,
1044 J., and Shaw, D.E. (2012). Oncogenic mutations counteract intrinsic disorder in the EGFR kinase

1045 and promote receptor dimerization. Cell *149*, 860–870.
1046
1047 Sinclair, J.K.L., Walker, A.S., Doerner, A.E., and Schepartz, A. (2018). Mechanism of Allosteric
1048 Coupling into and through the Plasma Membrane by EGFR. Cell Chem. Biol. *25*, 857–870.e7.
1049
1050 Singh, B., Carpenter, G., and Coffey, R.J. (2016). EGF receptor ligands: recent advances. [version
1051 1; peer review: 3 approved]. F1000Res *5*.
1052
1053 Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S.,
1054 and Carragher, B. (2005). Automated molecular microscopy: the new Leginon system. J Struct
1055 Biol *151*, 41–60.
1056
1057 Torrie, G.M., and Valleau, J.P. (1977). Nonphysical sampling distributions in Monte Carlo free-
1058 energy estimation: Umbrella sampling. J Comput Phys *23*, 187–199.
1059
1060 Uchikawa, E., Choi, E., Shang, G., Yu, H., and Bai, X.-C. (2019). Activation mechanism of the
1061 insulin receptor revealed by cryo-EM structure of the fully liganded receptor-ligand complex. Elife
1062 *8*.
1063
1064 Voss, N.R., Yoshioka, C.K., Radermacher, M., Potter, C.S., and Carragher, B. (2009). DoG Picker
1065 and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy.
1066 J Struct Biol *166*, 205–213.
1067
1068 Wilson, K.J., Gilmore, J.L., Foley, J., Lemmon, M.A., and Riese, D.J. (2009). Functional
1069 selectivity of EGF family peptide growth factors: implications for cancer. Pharmacol Ther *122*, 1–
1070 8.
1071
1072 Wu, E.L., Cheng, X., Jo, S., Rui, H., Song, K.C., Dávila-Contreras, E.M., Qi, Y., Lee, J., Monje-
1073 Galvan, V., Venable, R.M., et al. (2014). CHARMM-GUI Membrane Builder toward realistic
1074 biological membrane simulations. J Comput Chem *35*, 1997–2004.
1075
1076 Zhang, X., Gureasko, J., Shen, K., Cole, P.A., and Kuriyan, J. (2006). An allosteric mechanism
1077 for activation of the kinase domain of epidermal growth factor receptor. Cell *125*, 1137–1149.
1078
1079 Zhang, X., Yu, D., Sun, J., Wu, Y., Gong, J., Li, X., Liu, L., Liu, S., Liu, J., Wu, Y., et al. (2020).
1080 Visualization of Ligand-Bound Ectodomain Assembly in the Full-Length Human IGF-1 Receptor
1081 by Cryo-EM Single-Particle Analysis. Structure *28*, 555–561.e4.
1082
1083 Zheng, S.Q., Palovcak, E., Armache, J.-P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017).
1084 MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron
1085 microscopy. Nat Methods *14*, 331–332.
1086
1087 Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H.
1088 (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3.
1089 Elife *7*.
1090

Figure 1

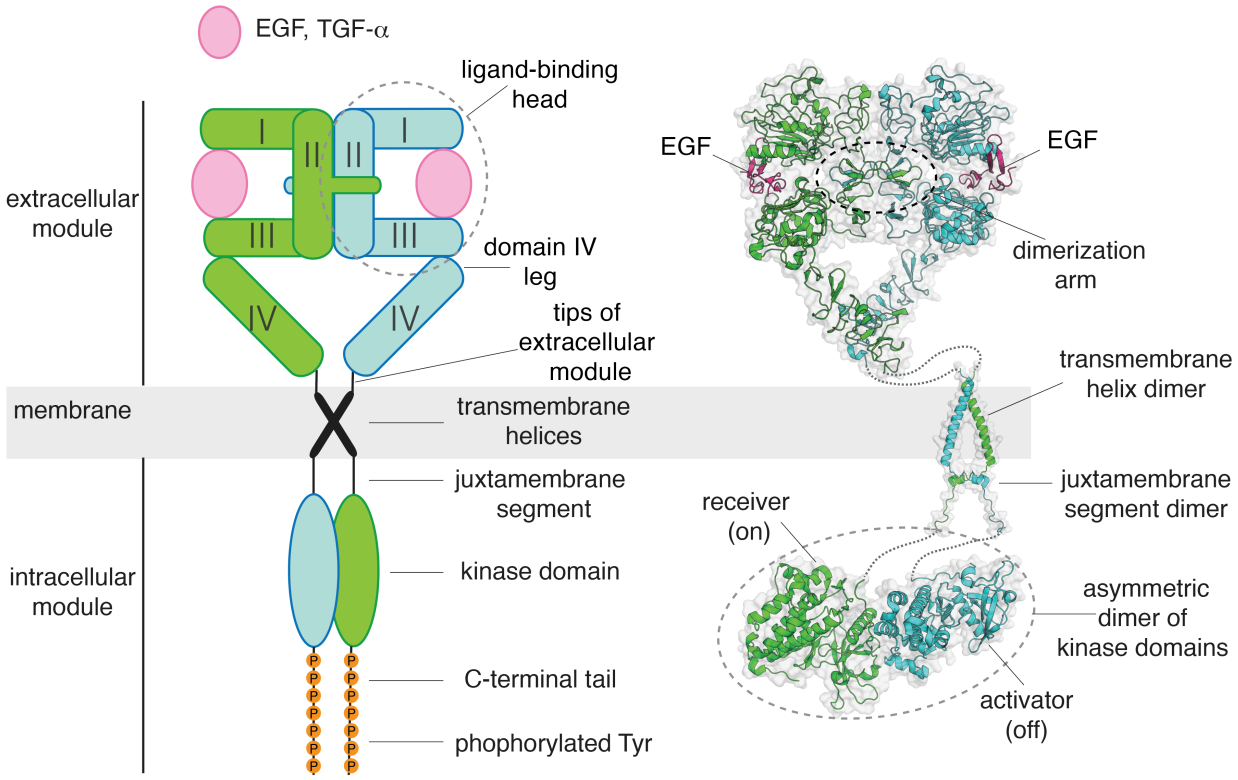
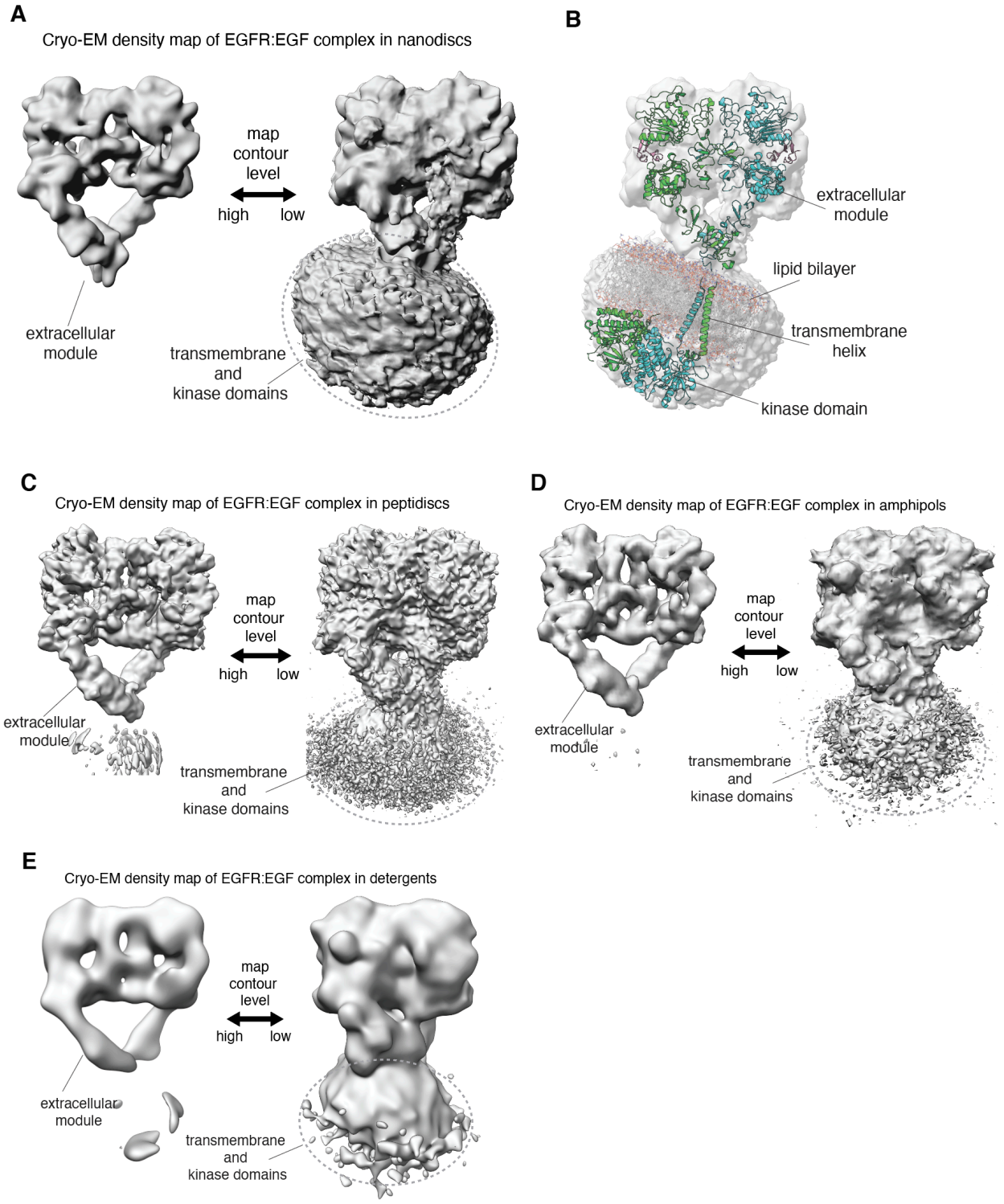


Figure 1 A structural collage of a ligand-bound EGFR dimer

Left, a schematic diagram of ligand-bound dimeric EGFR. Right, structures of EGFR fragments, determined previously: a dimeric extracellular module bound to EGF (PDB code: 3NJP (Lu et al., 2010)), a transmembrane helix dimer (PDB code: 2M20 (Endres et al., 2013)), and an asymmetric dimer of kinase domains (PDB code: 2GS6 (Zhang et al., 2006)).

Figure 2



1099

1100 **Figure 2 Linkage between the extracellular module and transmembrane helices in full-length**
1101 **human EGFR is not rigid**

1102 **A, C-E,** In each panel, the cryo-EM density map for a reconstituted full-length EGFR:EGF
1103 complex is shown at two contour levels. Note the lack of EM density for the transmembrane and
1104 kinase domains of EGFR at higher contour level (left) in all these complexes.

1105 **A**, Full-length EGFR:EGF complex reconstituted in nanodiscs.

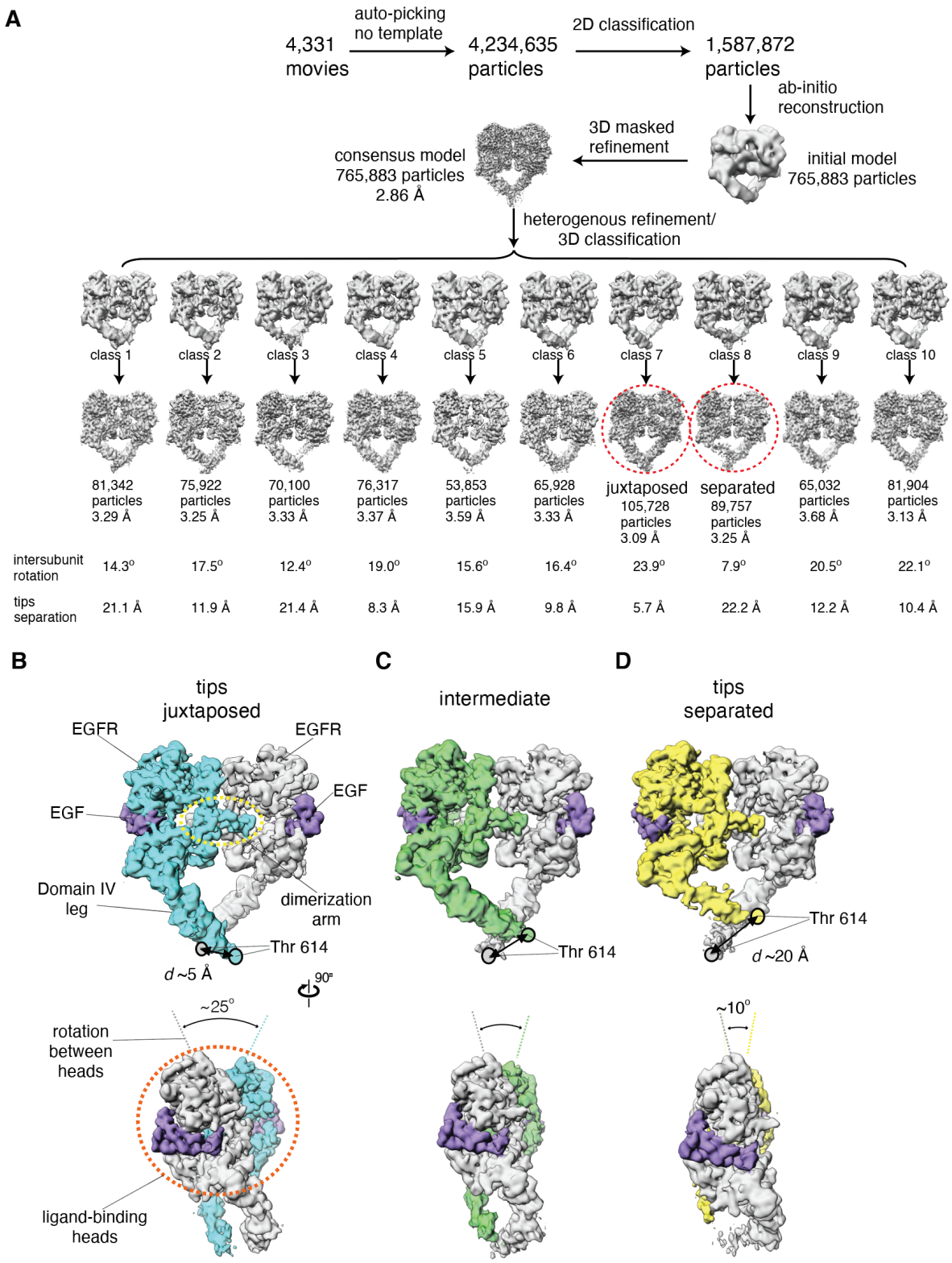
1106 **B**, Superposition of the cryo-EM map shown in **A** with a model for full-length EGFR embedded
1107 in the lipid bilayer, described previously (Arkhipov et al., 2013).

1108 **C**, Full-length EGFR:EGF complex reconstituted in peptidiscs.

1109 **D**, Full-length EGFR:EGF complex reconstituted in amphipols.

1110 **E**, Full-length EGFR:EGF complex reconstituted in detergents.

Figure 3



1111

1112

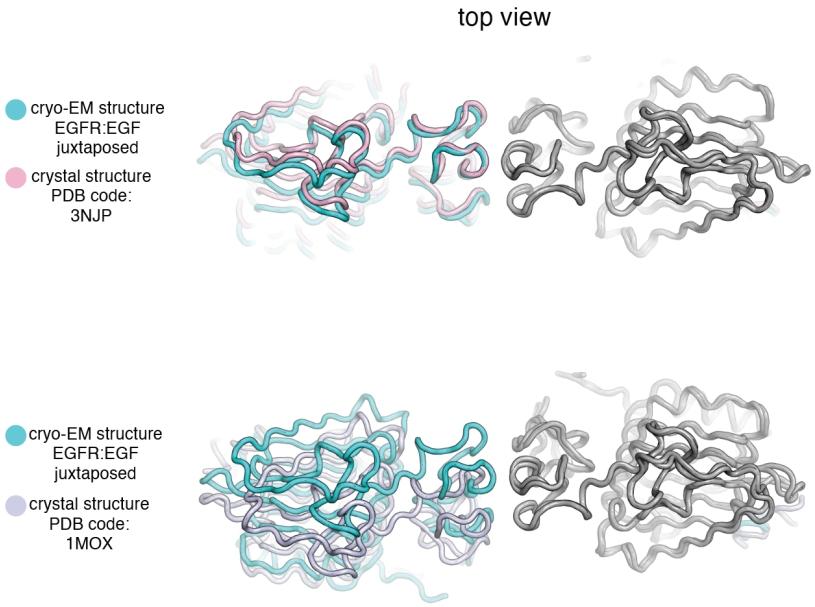
1113 **Figure 3 Cryo-EM analysis reveals a range of conformations for the EGF-bound**
1114 **extracellular module**

1115 **A,** Summary of the cryo-EM data processing workflow for the full-length EGFR:EGF complex
1116 reconstituted in peptidiscs. The two extreme conformations of the ensemble, the juxtaposed and
1117 the separated, are highlighted by red dotted circle. The intersubunit rotational angles and the
1118 separations between the tips of the extracellular module are summarized below each subclass.
1119 See Materials and Methods for more details.

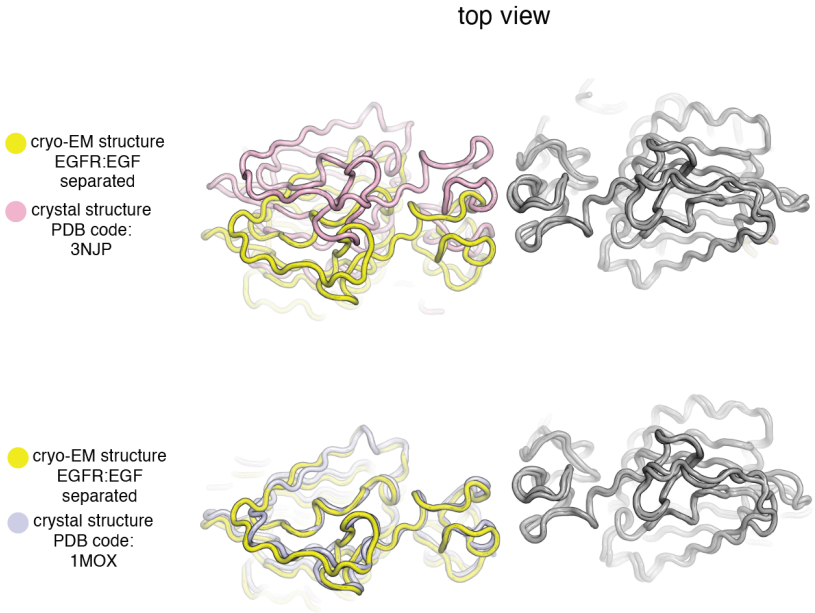
1120 **B-D**, Two orthogonal views of the cryo-EM maps for three representative states of the
1121 conformational ensemble of the EGF-bound extracellular module: the tips-juxtaposed
1122 conformation (**B**), an intermediate conformation (**C**), and the tips-separated conformation (**D**).
1123 These conformations differ from one another in the rotations between the ligand-binding heads, as
1124 well as in the distances between the two Thr 614 residues located at the membrane-proximal ends
1125 of the Domain IV legs.

Figure 4

A



B



1126

1127 **Figure 4 Comparison of Cryo-EM structures of EGFR:EGF complex in two extreme**
1128 **conformations with previous crystal structures of the ligand-binding head of EGFR bound**
1129 **to EGF and TGF-α.**

1130 In each panel, a cryo-EM structure of the extracellular module of the full-length EGFR:EGF
1131 complex (juxtaposed or separated) is compared to a crystal structure of the ligand-binding head of
1132 EGFR bound to EGF or TGF-α. In each comparison, structures are superimposed on the right-
1133 hand EGFR subunits (grey).

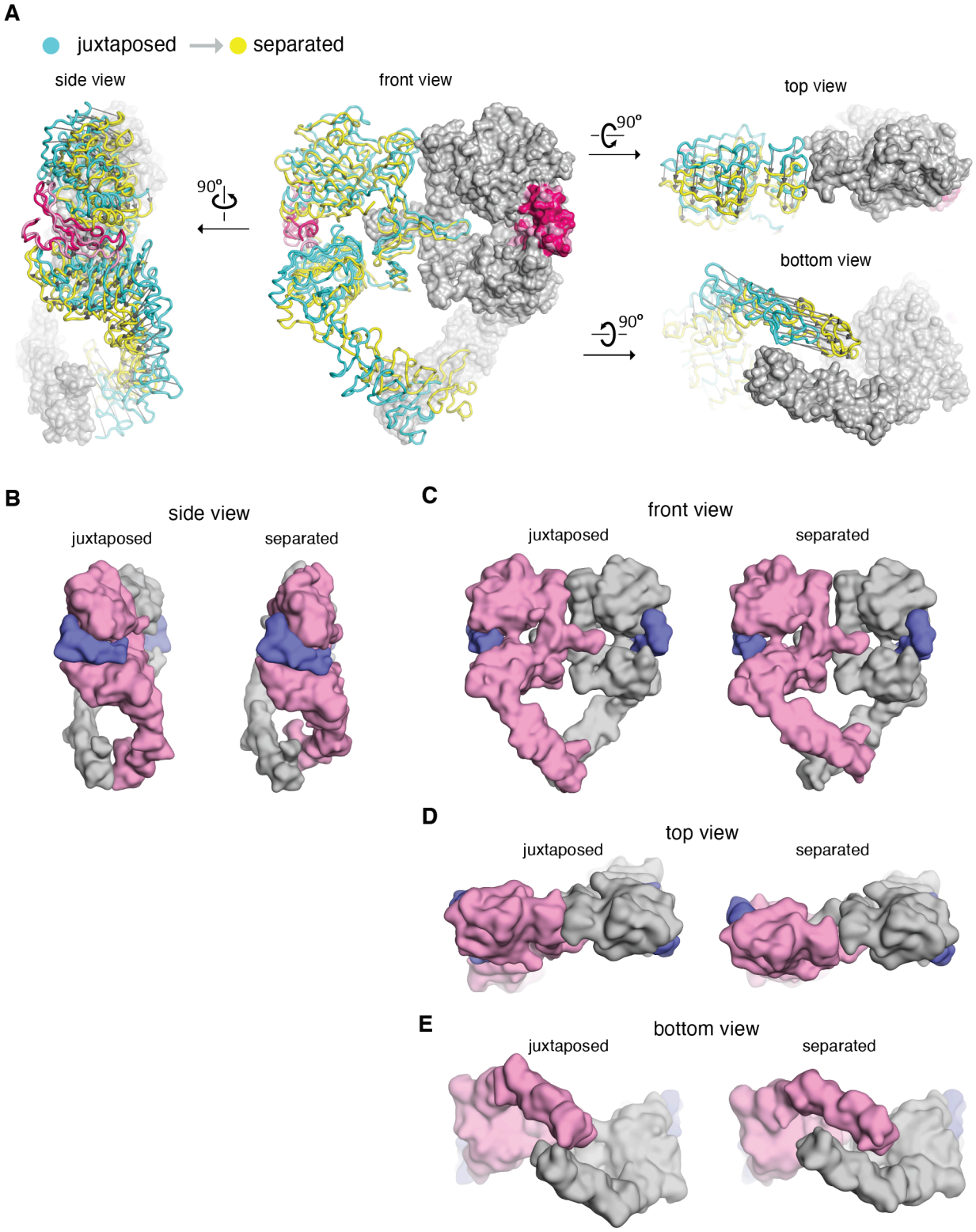
1134 **A**, Top, the cryo-EM structure of juxtaposed conformation of EGFR:EGF (cyan) is compared to a
1135 crystal structure of the ligand-binding head of EGFR bound to EGF (pink, PDB code: 3NJP).
1136 Bottom, the cryo-EM structure of juxtaposed conformation of EGFR:EGF (cyan) is compared to
1137 a crystal structure of the ligand-binding head of EGFR bound to TGF-α (light purple, PDB code:
1138 1MOX).
1139 Note that the cryo-EM structure of the juxtaposed conformation of the EGFR:EGF complex
1140 resembles closely the crystal structure of the ligand-binding head of EGFR bound to EGF (PDB
1141 code: 3NJP), but deviates from the crystal structure of the ligand-binding head of EGFR bound to
1142 TGF-α (PDB code: 1MOX).

1143 **B**, Top, the cryo-EM structure of separated conformation of EGFR:EGF (yellow) is compared to
1144 a crystal structure of the ligand-binding head of EGFR bound to EGF (pink, PDB code: 3NJP).
1145 Bottom, the cryo-EM structure of separated conformation of EGFR:EGF (yellow) is compared to
1146 a crystal structure of the ligand-binding head of EGFR bound to TGF-α (light purple, PDB code:
1147 1MOX).
1148 Note that the cryo-EM structure of the separated conformation of the EGFR:EGF complex
1149 resembles closely the crystal structure of the ligand-binding head of EGFR bound to TGF-α (PDB

1150 code: 1MOX), but deviates from the crystal structure of the ligand-binding head of EGFR bound

1151 to EGF (PDB code: 3NJP).

Figure 5



1152

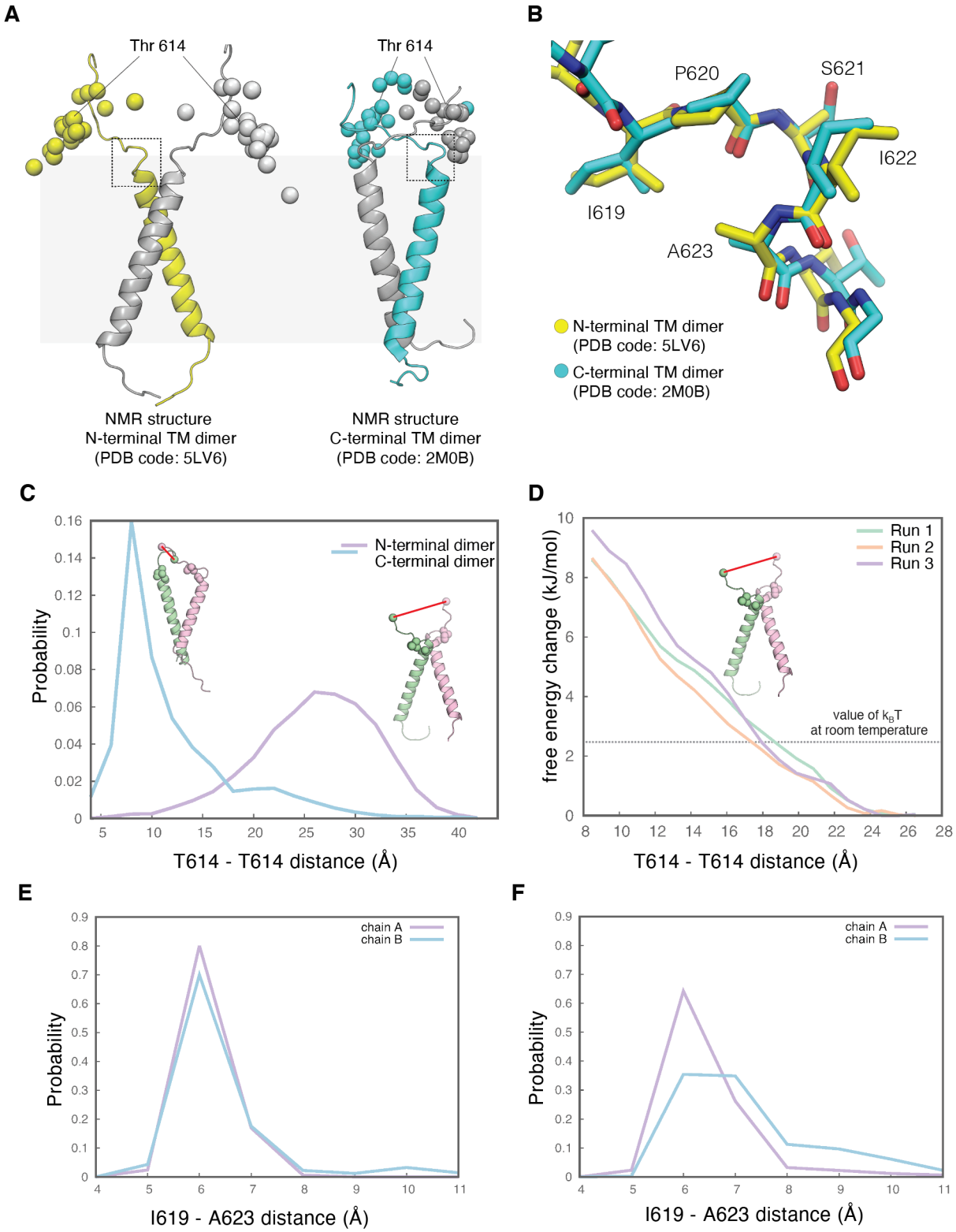
1153 **Figure 5 Comparison between the cryo-EM structures of the juxtaposed and the separated**
1154 **conformations of the extracellular module of the EGFR:EGF complex**

1155 **A**, The cryo-EM structures representing the juxtaposed conformation (cyan) and the separated
1156 conformation (yellow) are superimposed on the right-hand EGFR subunits which have similar
1157 Domain IV leg conformations (light grey). A vector (dark grey) is used to indicate the displacement
1158 of Cα atoms between the two structures, pointing from the juxtaposed structure to the separated
1159 structure.

1160 **B-E**, Side by side comparison between the juxtaposed conformation and the separated
1161 conformation of the extracellular module of the EGFR:EGF complex. Molecular surface of the
1162 cryo-EM models of the juxtaposed and separated conformations are shown from the front (**A**), the
1163 side (**B**), the top (**C**), and the bottom (**D**). EGF molecules are shown in slate blue. The same EGFR
1164 subunit used for superposition in **A** is colored grey and the other EGFR subunit is light pink.

1165

Figure 6



1166

1167 **Figure 6 The tips-separated conformation of the extracellular module can couple more**
1168 **readily to the N-terminal dimer of the transmembrane helices**

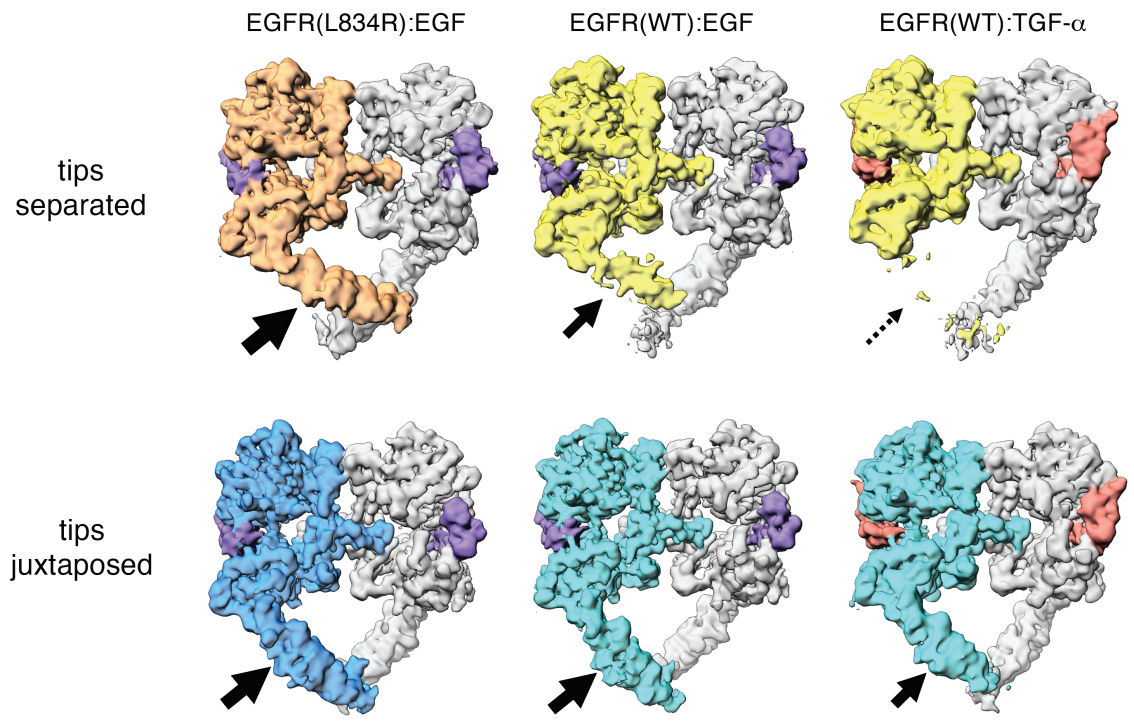
1169 **A**, Models of the transmembrane helices in two configurations based on previous NMR structures.
1170 One of the structures in the NMR-based ensemble for the N-terminal dimer configuration (PDB
1171 code: 5LV6) is shown in yellow (chain A) and grey (chain B), and the structure for the C-terminal
1172 dimer configuration (PDB code: 2M0B) is shown in cyan (chain A) and grey (chain B). Shown in
1173 spheres are superimposed Cα atoms of the residue Thr 614, located at the end of the extracellular
1174 module, from the 20 structures in the NMR-based ensemble for the N-terminal dimer configuration
1175 (left) and for the C-terminal dimer configuration (right), respectively. The black dotted square
1176 indicates the linker region that is imposing a constraint on the orientation of the linkers. In the N-
1177 terminal dimer configuration of the transmembrane helices (left), the linkers from each subunit are
1178 oriented away from one another. In the C-terminal dimer configuration of the transmembrane
1179 helices (right), the linkers are oriented towards each other.

1180 **B**, Enlarged view of the linker located at the junction with the transmembrane helices, zoomed in
1181 from the areas highlighted by dotted black square in **A**, showing well-defined conformations for
1182 this region in both NMR structures.

1183 **C**, Probability distribution of the distances between the Cα atoms of the two Thr 614 residues,
1184 sampled every 250 picoseconds over 2.5 μs of MD simulations of the N-terminal dimer
1185 configuration (purple) and the C-terminal dimer configuration (cyan) of the transmembrane helices.
1186 Initial structures of the transmembrane helices dimers with the linker are shown as cartoons.
1187 Residues Thr 614 located at the end of the extracellular module are shown as spheres. Also shown
1188 as spheres are residues Ile 619 and Ala 623, located at the junction between the linker and the

1189 transmembrane helix, and the interaction between these two residues is maintained over 2.5 μs
1190 MD simulations (see **E** and **F**).
1191 **D**. Free-energy change of the system when the distance between the Thr 614 residues from each
1192 subunit is driven from 27 Å to 8 Å, when the transmembrane helices are in the N-terminal dimer
1193 configuration. Results from three independent umbrella-sampling calculations are shown in green,
1194 orange, and purple. The dotted gray line indicates the value of k_BT at room temperature, 2.479
1195 kJ/mol.
1196 **E** and **F**, Probability distribution of the distances between the Cα atoms of Ile 619 and Ala 623,
1197 sampled every 250 picoseconds over 2.5 μs of MD simulations of the N-terminal dimer
1198 configuration (**E**) and the C-terminal dimer configuration of the transmembrane helices (**F**).
1199

Figure 7



1200

1201

1202

1203 **Figure 7 Comparison between Cryo-EM density maps of the tips-separated and the tips-**

1204 **juxtaposed conformations in different EGFR complexes**

1205 Cryo-EM density maps of the tips-separated (upper row) and the tips-juxtaposed (lower row)

1206 conformations in three different EGFR complexes are shown: EGFR(L834R) bound to EGF (left),

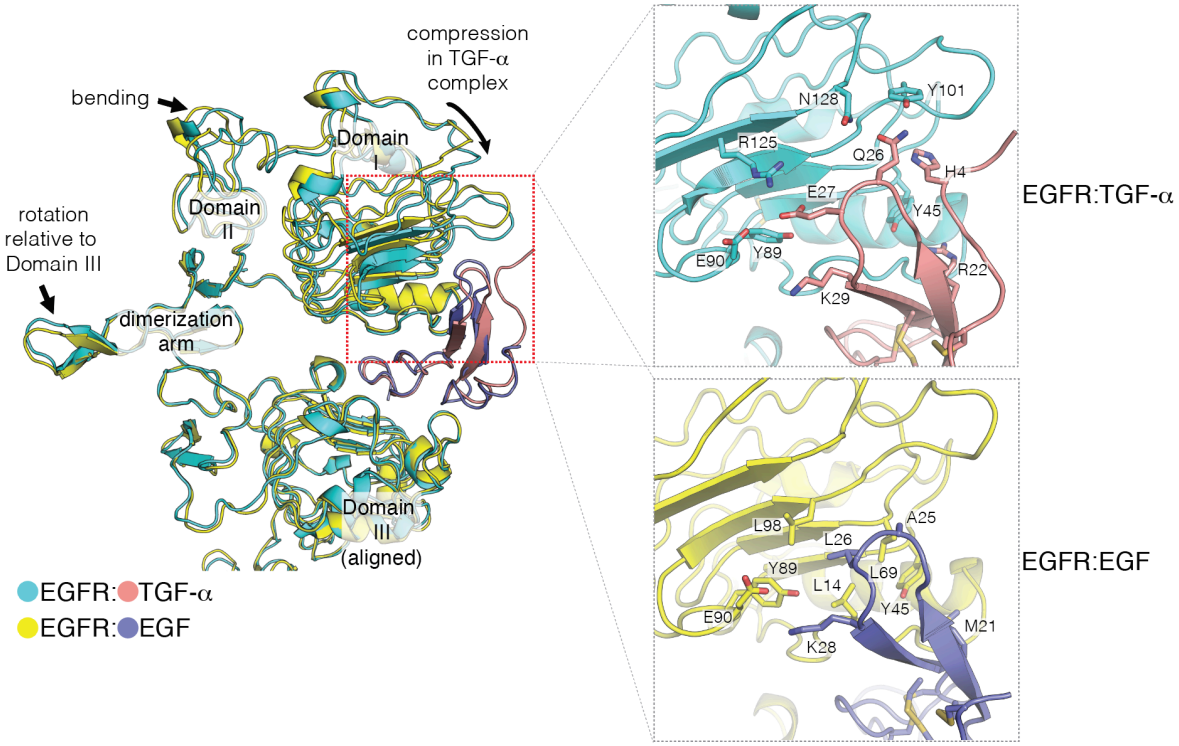
1207 EGFR(WT) bound to EGF (middle), and EGFR(WT) bound to TGF-α (right). The black arrows

1208 highlight the difference in stabilities of the flexible Domain IV legs in these complexes.

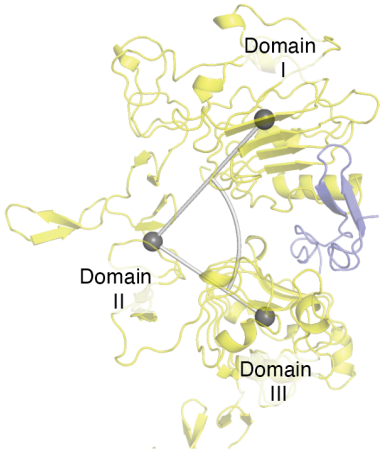
1209

Figure 8

A

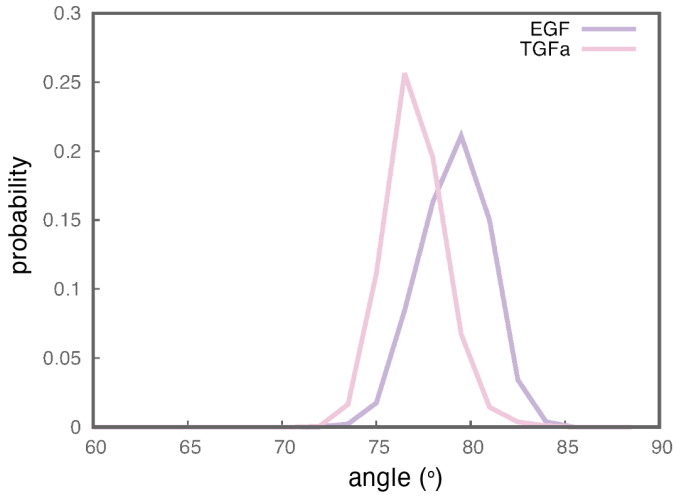


B



angle between centroids of domain I, II, and III

C



1210

1211

1212 **Figure 8 Difference in the ligand-binding heads between EGF- and TGF-α-bound EGFR**
1213 **complexes**

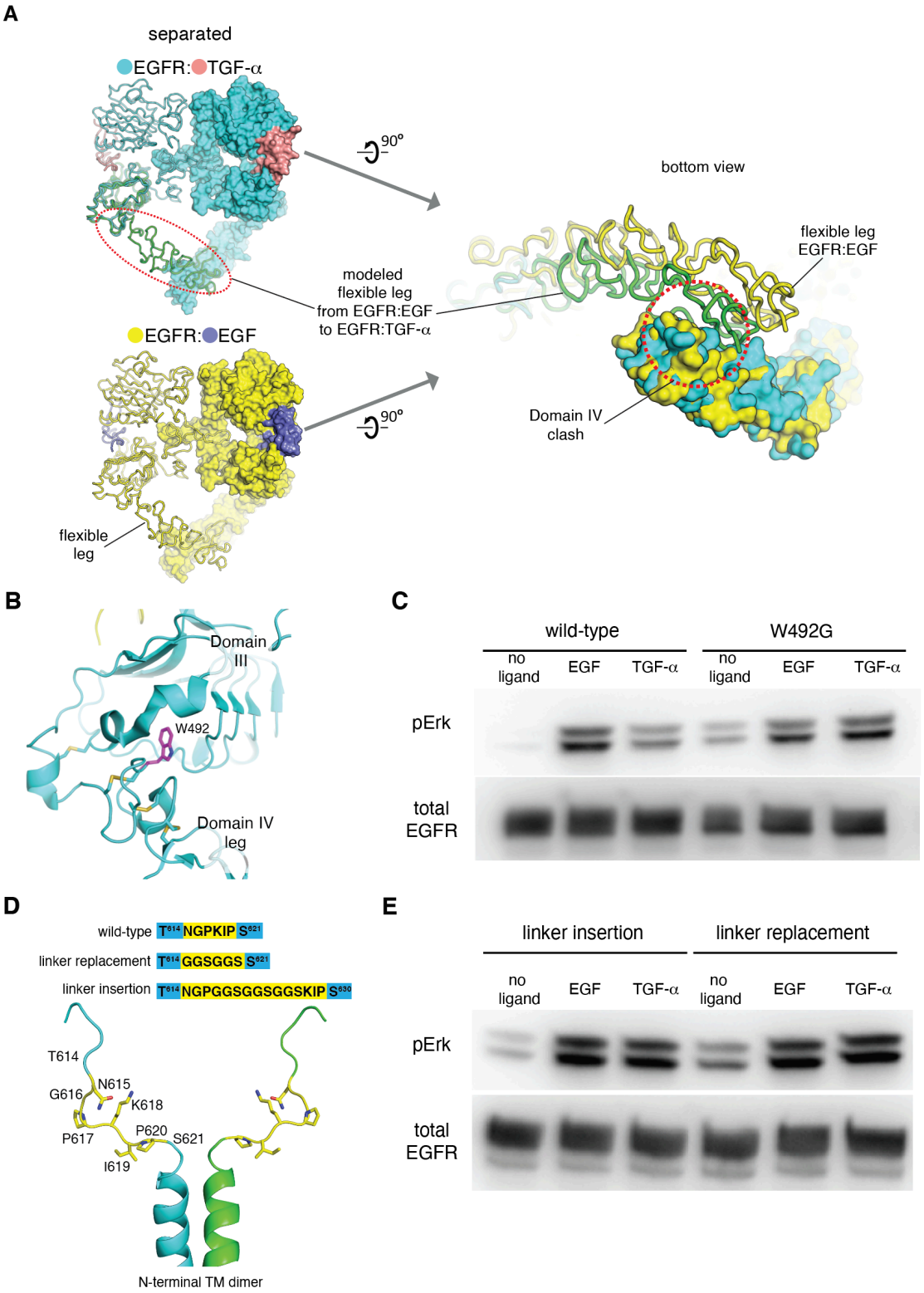
1214 **A**, The differences in structure between the separated conformations of the EGF and TGF-
1215 α complexes of EGFR are illustrated. Structures corresponding to the separated conformations of
1216 the EGFR:EGF complex (yellow:slate) and the EGFR:TGF-α complex (cyan:salmon) are aligned
1217 on Domain III of each structure. Enlarged views of the interface between Domain I of EGFR and
1218 the ligands are shown on the right. The interface between Domain I of EGFR (cyan) and TGF-
1219 α (salmon) features hydrophilic residues that are shown (upper right). The equivalent interface
1220 between Domain I of EGFR (yellow) and EGF (slate) is characterized by hydrophobic residues
1221 (lower right).

1222 **B**, The angle between centroids of Domain I (residues 7-200), part of Domain II (residues 260-
1223 285), and Domain III (317-470) is used to estimate the degree of closure of the ligand-binding site
1224 in a ligand-bound EGFR complex.

1225 **C**, Probability distribution of the angles defined in **B**, sampled every nanosecond over ~1.8 μs and
1226 ~2.0 μs of MD simulations of the EGFR extracellular module bound to EGF (purple) and TGF-
1227 α (pink), respectively.

1228

Figure 9



1229

1230 **Figure 9 EGF and TGF-α differ in their abilities to stabilize the tips-separated conformation**
1231 **and to generate different levels of EGFR activation**

1232 **A**, Destabilization of the flexible Domain IV leg in the separated conformation of the EGFR:TGF-
1233 α complex. The EGFR:TGF-α complex (upper left) is colored in cyan:salmon and the
1234 EGFR:EGF complex (lower left) is colored in yellow:slate. The flexible leg in the EGFR:TGF-
1235 α complex (green, red dotted line) is modeled based on the conformation of Domains III and IV
1236 in the left-hand subunit of the EGFR:EGF complex (lower left). When viewed from the bottom
1237 (right), the two Domain IV legs are accommodated in the EGFR:EGF complex (yellow), but not
1238 in the EGFR:TGF-α complex. There are steric clashes between the legs (green and cyan), within
1239 the dotted red line.

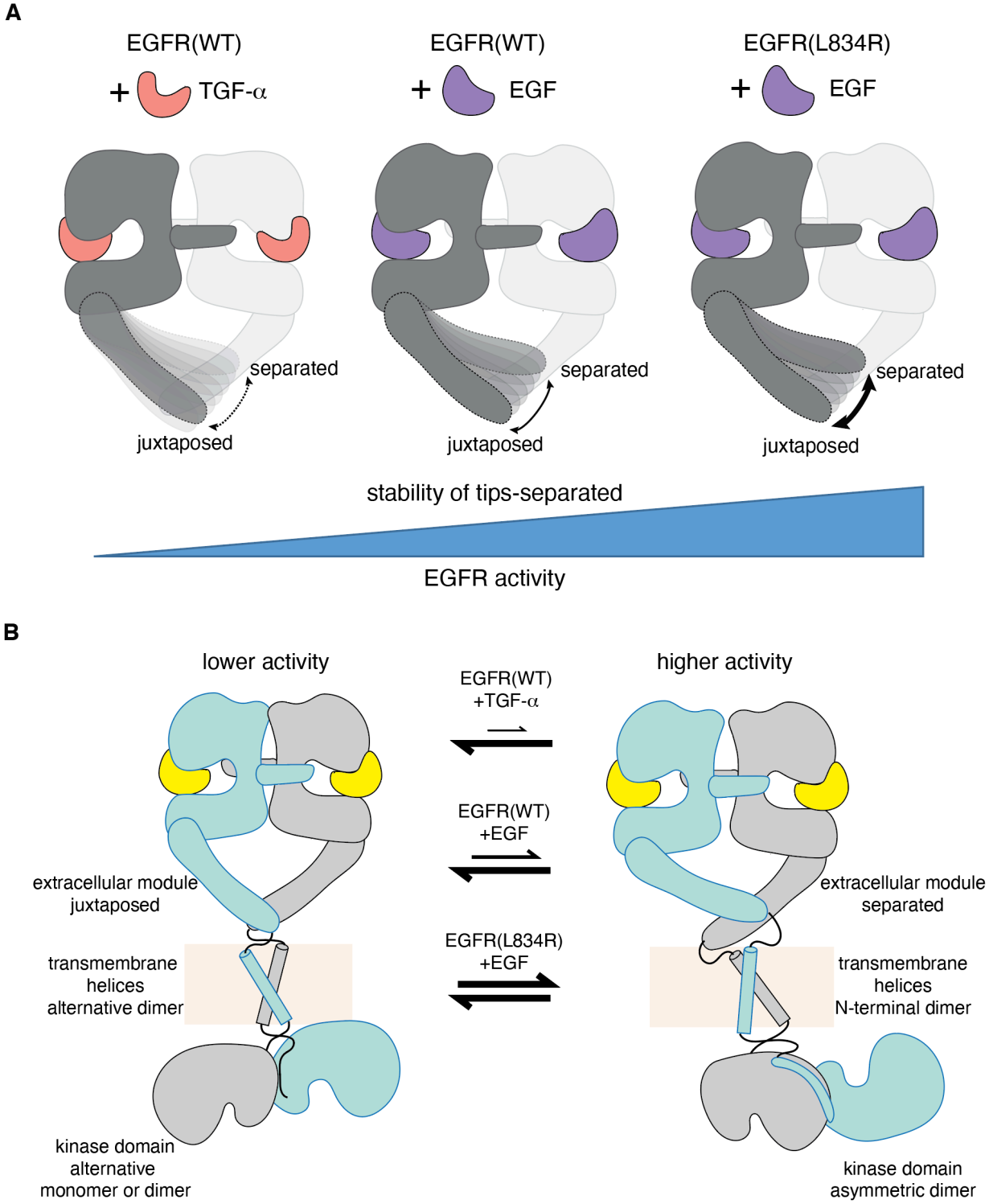
1240 **B**, Location of the Trp 492 (purple), shown in the crystal structure of EGFR bound to EGF (cyan,
1241 PDB code: 3NJP).

1242 **C**, Western-blot analysis of Erk phosphorylation in HEK293T cells transiently transfected with
1243 wild-type EGFR or EGFR variant, W492G. Anti-phospho-Erk antibody is used to detect
1244 phosphorylation of p44 and p42 MAP kinase (Erk1 and Erk2), and anti-FLAG antibody is used to
1245 detect the total amount of EGFR in the cell lysate loaded to each lane.

1246 **D**, Mutations introduced to the linker at the junction between the extracellular module and the
1247 transmembrane helices, shown in the NMR structure of the N-terminal dimer configuration of the
1248 transmembrane helices (PDB code: 5LV6).

1249 **E**, Western-blot analysis of Erk phosphorylation in HEK293T cells transiently transfected with
1250 two EGFR variants, linker insertion (left) and linker replacement (right). In linker insertion variant,
1251 Gly-Gly-Ser-Gly-Gly-Ser-Gly-Gly-Ser is inserted between Pro 617 and Lys 618. In linker
1252 replacement variant, the linker (Asn 615-Pro 620) is replaced by Gly-Gly-Ser-Gly-Gly-Ser.

Figure 10



1253

1254

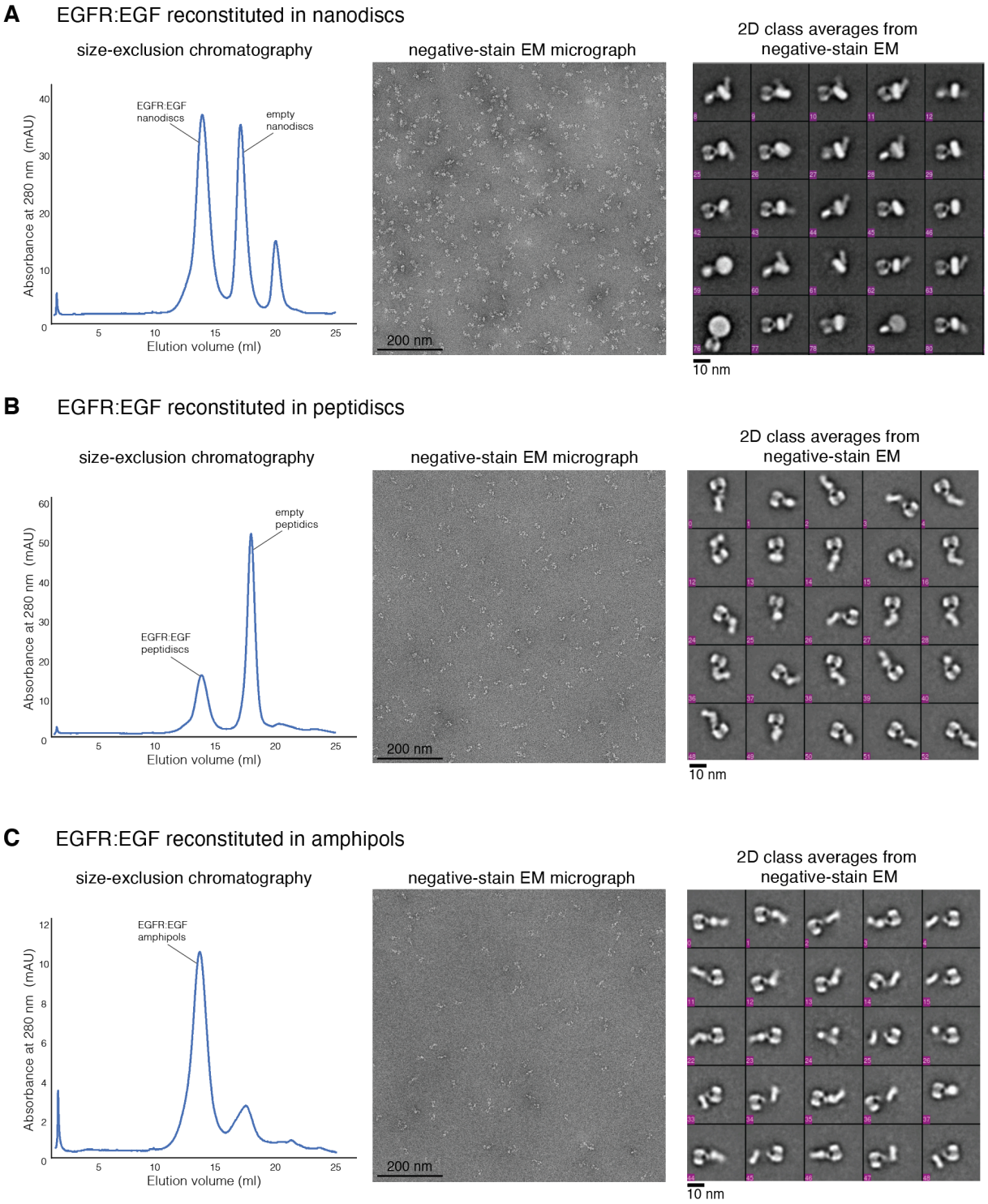
1255 **Figure 10 A mechanism for differential activation of EGFR by EGF and TGF-α**

1256 **A,** Schematic representations of the ligand-bound extracellular module of EGFR illustrate the
1257 differential modulation of the conformation of the Domain IV legs by TGF-α (left), EGF(middle)
1258 and oncogenic L834R mutation (right). Cryo-EM densities for the Domain IV legs in the
1259 juxtaposed conformation is similarly defined in all three complexes (EGFR(WT):TGF-α,
1260 EGFR(WT):EGF, and EGFR(L834R):EGF). In contrast, density for the Domain IV legs in the
1261 separated conformation is poorly defined in the EGFR(WT):TGF-α complex and best defined in
1262 the EGFR(L834R):EGF complex. Increased stability of the tips-separated conformation of the
1263 extracellular module is correlated with higher level of EGFR activity.

1264 **B**, Schematic diagrams illustrate two extreme dimeric conformations of the ligands-bound full-
1265 length EGFR, representing a lower activity state (left) and a higher activity state (right). It shows
1266 how different EGFR ligands and oncogenic mutation could generate different EGFR activity by
1267 altering the balance between these dimeric conformations of the full-length EGFR.

1268

Figure 2-figure supplement 1



1271 **Figure 2-figure supplement 1 Reconstitution of full-length human EGFR:EGF complex.**

1272 **A**, Sample preparation of full-length EGFR:EGF reconstituted into nanodiscs.

1273 **B**, Sample preparation of full-length EGFR:EGF reconstituted into peptidiscs.

1274 **C**, Sample preparation of full-length EGFR:EGF reconstituted into amphipols.

1275 For each sample preparation, a representative profile of size-exclusion chromatography is shown
1276 on the left, a representative negative-stain micrograph is shown in the middle, and the
1277 representative 2D class averages are shown on the right.

1278

1279

1280

1281

1282

1283

1284

1285

1286

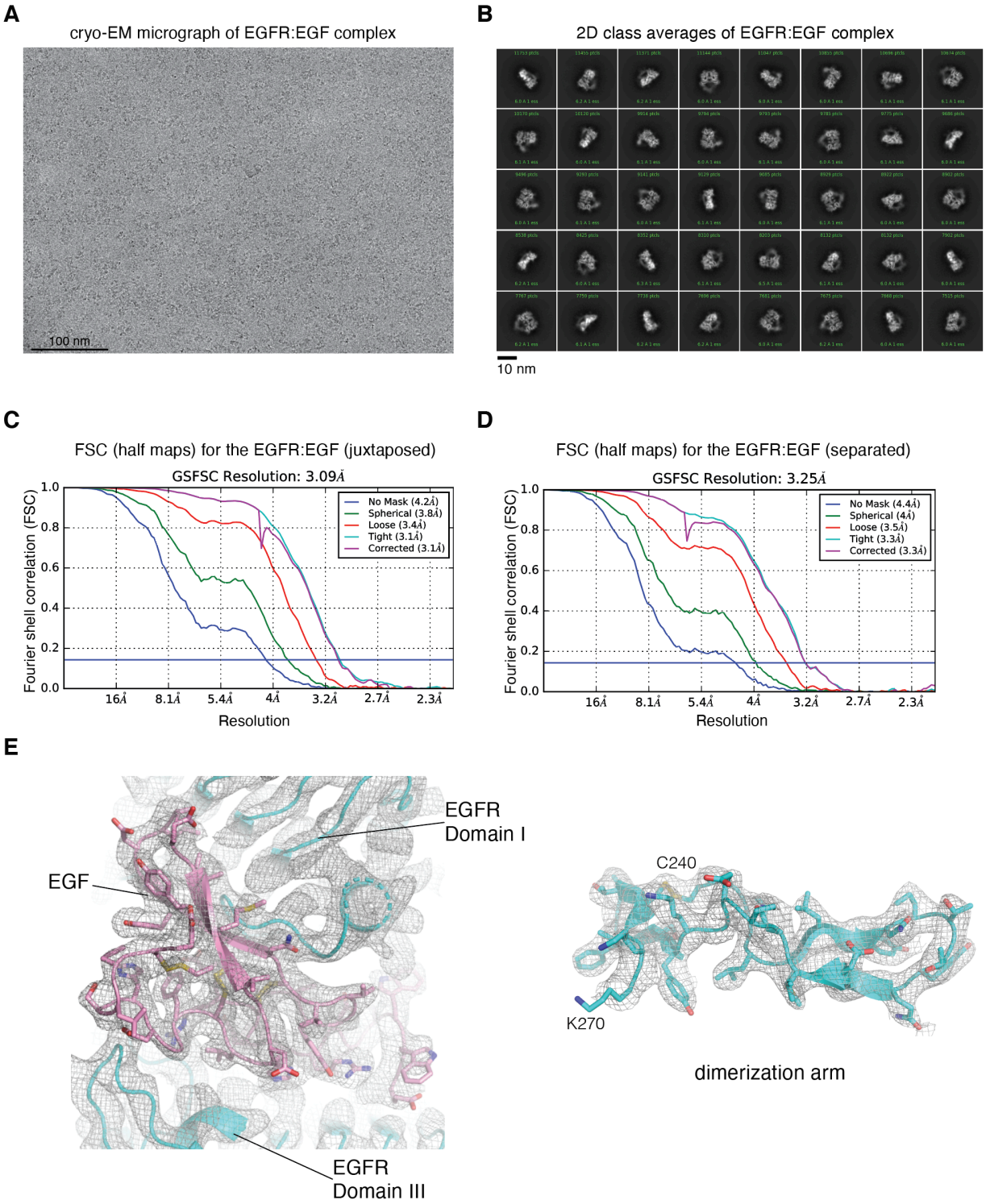
1287

1288

1289

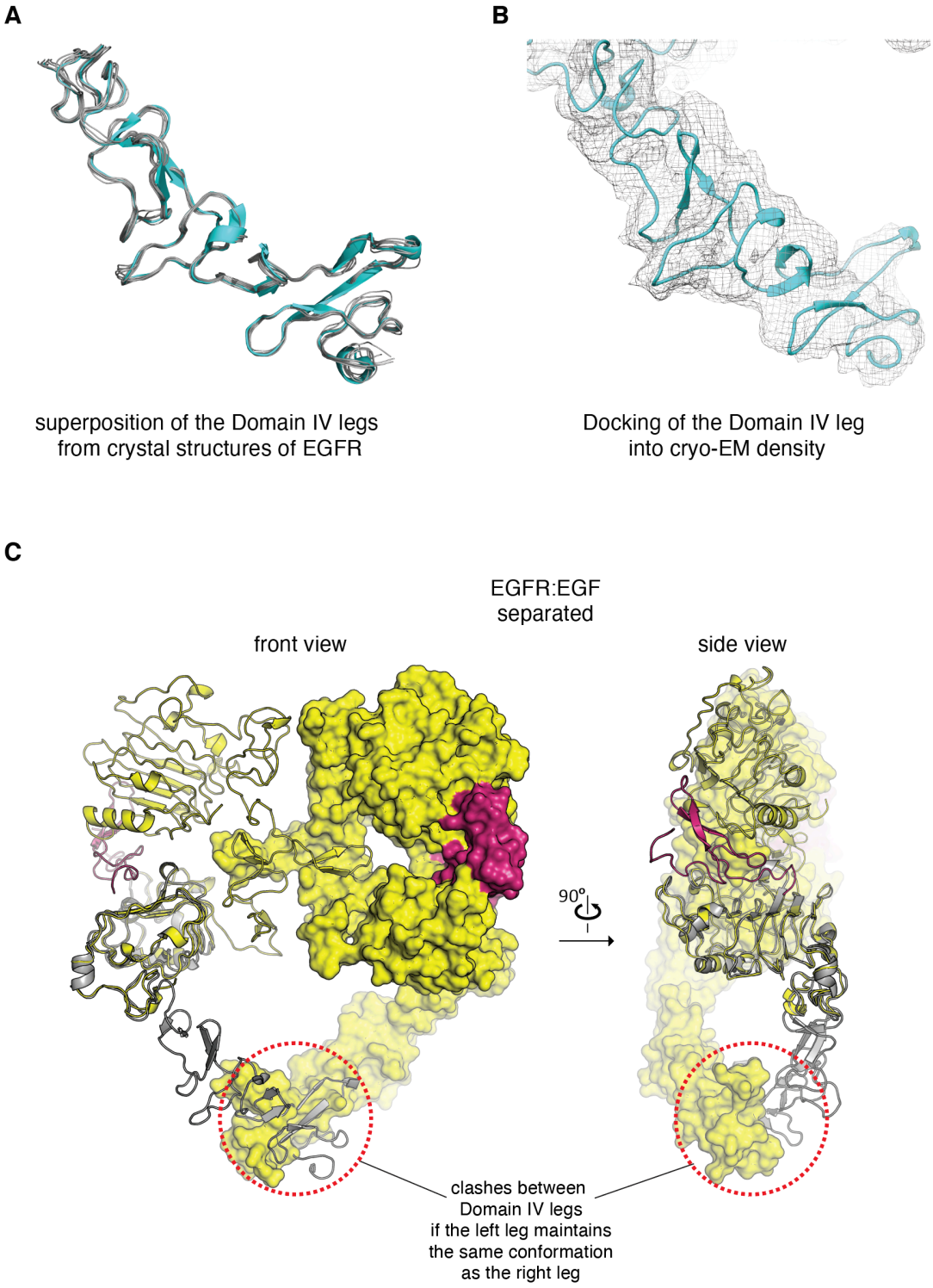
1290

Figure 3-figure supplement 1



1293 **Figure 3-figure supplement 1 Cryo-EM image processing for focused reconstruction on the**
1294 **extracellular module of full-length EGFR:EGF complex reconstituted in peptidiscs.**
1295 **A**, Representative cryo-EM micrograph after motion-correction.
1296 **B**, Representative 2D class averages generated from reference-free 2D classification.
1297 **C**, Fourier shell correlation (FSC) curves output from cryoSPARC v2 Non-uniform refinement
1298 job, for the cryo-EM reconstruction of the juxtaposed conformation of the EGFR:EGF complex
1299 reconstituted in peptidiscs. The resolution is estimated at FSC = 0.143.
1300 **D**, Fourier shell correlation (FSC) curves output from cryoSPARC v2 Non-uniform refinement
1301 job, for the cryo-EM reconstruction of the separated conformation of the EGFR:EGF complex
1302 reconstituted in peptidiscs. The resolution is estimated at FSC = 0.143.
1303 **E**, Representative cryo-EM densites for the juxtaposed conformation of the EGFR:EGF complex.
1304 The structural model and the cryo-EM density corresponding to one of the EGF binding site (left)
1305 and the dimerization arm (right, Cys 240-Lys 270) are shown.
1306

Figure 3-figure supplement 2



1307

1308

1309

1310 **Figure 3-figure supplement 2 Conformation of the Domain IV leg.**

1311 **A**, Superposition of the Domain IV leg from the dimeric crystal structure of the isolated EGFR
1312 extracellular domains bound to EGF (cyan, PDB code: 3NJP), with the Domain IV legs from seven
1313 different crystal structures of monomeric extracellular domains of EGFR (grey, PDB codes: 4UV7,
1314 5XWD, 1NQL, 3QWQ, 1YY9, 6ARU, 4UIP). This comparison shows that the Domain IV leg is
1315 a fairly rigid structural component.

1316 **B**, The Domain IV leg from a crystal structure of the isolated EGFR extracellular domains bound
1317 to EGF (PDB code: 3NJP) is fitted into the cryo-EM density corresponding to the flexible leg in
1318 the cryo-EM map for the juxtaposed conformation of the EGFR:EGF complex.

1319 **C**, The extracellular module of the EGFR:EGF complex in separated conformation is shown in
1320 two orthogonal views. The Domain IV leg shown on the left-hand subunit (grey) is modeled based
1321 on the conformation of Domains III and IV in the right-hand subunit. Maintaining the same
1322 conformation in both legs would lead to steric clashes between the two legs (circled by red dotted
1323 lines), when the extracellular module of EGFR adopts the separated conformation.

Figure 3-table supplement 1
Statistics of cryo-EM data collection and structure determination

	EGFR(WT) :EGF juxtaposed	EGFR(WT) :EGF separated	EGFR(WT) :TGF-α juxtaposed	EGFR(WT) :TGF-α separated	EGFR(L834R) :EGF juxtaposed	EGFR(L834R) :EGF separated
Data collection						
Electron microscope	Titan Krios	Titan Krios	Titan Krios	Titan Krios	Titan Krios	Titan Krios
Camera	Gatan K3	Gatan K3	Gatan K3	Gatan K3	Gatan K3	Gatan K3
Voltage (kV)	300	300	300	300	300	300
Magnification	81,000	81,000	81,000	81,000	81,000	81,000
Pixel sise (Å)	0.54	0.54	0.54	0.54	0.54	0.54
Nominal dose (e^-/Å^2)	50	50	50	50	50	50
Tilt angle (°)	40	40	40	40	40	40
Defocus range (μm)	-0.75 to -1.5	-0.75 to -1.5	-0.75 to -1.5	-0.75 to -1.5	-0.75 to -1.5	-0.75 to -1.5
Image processing						
Micrographs	4,331	4,331	7,406	7,406	5,013	5,013
Final particle number	105,728	89,757	126,471	102,206	123,670	100,615
Map resolution (Å)	3.1	3.3	3.4	3.6	3.3	3.4
Map sharpening B factor	-97.6	-94.7	-108.2	-108.2	-103.5	-92.8
Model refinement						
Rms deviation bonds (Å)	0.007	0.006	0.004	0.007	0.004	0.004
Rms deviation angles (°)	0.691	0.732	0.639	0.820	0.663	0.705
Map CC (mask)	0.82	0.76	0.79	0.74	0.81	0.72
Validation						
Molprobility score	1.96	2.20	2.05	2.45	1.91	2.19
Clash score	7.50	11.14	8.84	17.78	6.41	10.87
Rotamer outliers (%)	0.09	0.19	0.00	0.38	0.00	0.00
Cβ outliers (%)	0.00	0.00	0.00	0.00	0.00	0.00
Ramachandran plot						
Favored (%)	90.08	86.00	88.97	82.75	89.55	86.14
Allowed (%)	9.92	14.00	11.03	17.25	10.45	13.86
Outliers (%)	0.00	0.00	0.00	0.00	0.00	0.00

1324

1325

1326 **Figure 3-table supplement 1 Statistics of cryo-EM data collection and structure**

1327 **determination**

1328 **Figure 3-video supplement 1-3**

1329 3D variability analysis of the EGFR:EGF complex performed in cryoSPARC v2 (Punjani and Fleet,
1330 2020), shown from the front (Figure3_video s1_front), the side (Figure3_video s2_side), and the
1331 top (Figure3_video s3_top). Note the intersubunit rotation between the two heads and the
1332 corresponding movement of the Domain IV legs.

1333

1334

1335

1336

1337

1338

1339

1340

Figure 9-figure supplement 1

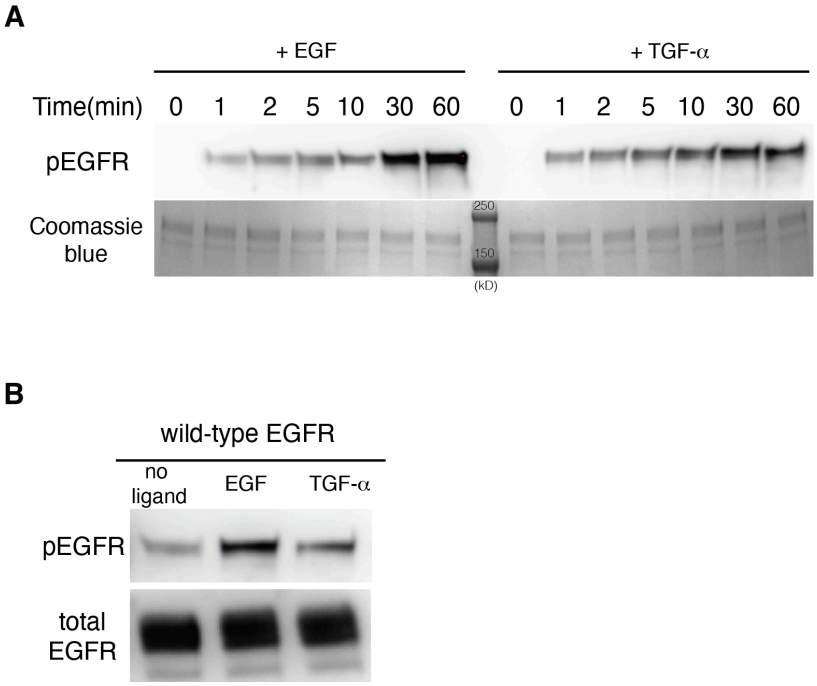


Figure 9-figure supplement 1 EGF and TGF-α generate different levels of EGFR autophosphorylation

A, Western-blot analysis of EGFR autophosphorylation in purified full-length EGFR protein stimulated by EGF and TGF-α. A pan-phosphotyrosine antibody is used to detect total phosphorylation of the receptor at the indicated time points after the reaction was initiated by addition of ATP. Coomassie brilliant blue staining shows the total amount of EGFR protein loaded to each lane on the gel. Two protein standard bands (250 kD and 150 kD) were shown in the middle lane.

B, Western-blot analysis of EGFR autophosphorylation in HEK293T cells transiently transfected with EGFR. A pan-phosphotyrosine antibody is used to detect total phosphorylation of the receptor

1352 in cells that are left unstimulated, stimulated by EGF, and stimulated by TGF-α. An anti-FLAG
1353 antibody is used to detect the total amount of EGFR proteins in the cell lysate loaded to each lane.