EGF signalling in epithelial carcinoma cells utilizes preformed receptor homoclusters, with larger heteroclusters post activation

Charlotte Fournier[1,2,+], Adam J. M. Wollman[3,+], Isabel Llorente-Garcia[1,4,+], Oliver Harriman[1,+], Djamila Ouarat[5], Jenny Wilding[5], Walter Bodmer[5] and Mark C. Leake[3,*]

[1] Department of Physics, Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, United Kingdom.

[2] Current address: Okinawa Institute of Science and Technology Graduate University, 119-1 Tancha, Onna-son, Kunigami-gun, Okinawa, Japan 904-0495.

[3] Biological Physical Sciences Institute (BPSI), Departments of Physics and Biology, University of York, York, United Kingdom.

[4] Current address: Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom

[5] MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, United Kingdom.

[+] These authors contributed jointly to this work

[*] Correspondence should be addressed to M.C.L: mark.leake@york.ac.uk

Abstract: Epidermal growth factor (EGF) signalling regulates cell growth, differentiation and proliferation in epithelium and EGF receptor (EGFR) overexpression has been reported in several carcinoma types. Structural and biochemical evidence suggests EGF binding stimulates EGFR monomer-dimer transitions, activating downstream signalling. However, mechanistic details of ligand binding to functional receptors in live cells remain contentious. We report real time single-molecule TIRF of human epithelial carcinoma cells with negligible native EGFR expression, transfected with GFP-tagged EGFR, before and after receptor activation with TMR-labelled EGF ligand. Fluorescently labelled EGFR and EGF are simultaneously tracked to 40nm precision to explore stoichiometry and spatiotemporal dynamics upon EGF binding. Using inhibitors that block binding to EGFR directly, or indirectly through HER2, our results indicate that pre-activated EGFR consists of preformed homoclusters, while larger heteroclusters including HER2 form upon activation. The relative stoichiometry of EGFR to EGF after binding peaks at 2, indicating negative cooperativity of EGFR activation.

Keywords

Single-molecule, nanoscale, EGF, signal transduction, receptor oligomerization, cancer inhibitors.

Main Text:

The epidermal growth factor receptor (EGFR) is essential for normal growth and development of epithelial tissues and is a key component in several signaling pathways[1]. Aberrant signal transduction is a primary driver of many epithelial cancers, EGFR upregulation implicated in formation and progression of several carcinomas[2]. Human EGFR or ERBB1, (also denoted 'ErB1'or 'HER1') is a 1,186 amino acid (aa) residue 170 kDa molecular weight protein[3] belonging to a family of receptor tyrosine kinase (RTK) receptors with three additional members: ERBB2 ('ErbB2', 'HER2' or 'neu'), ERBB3 ('ErbB3' or 'HER3') and ERBB4 ('ErbB4' or 'HER4') expressed predominantly in the plasma membrane of epithelial cells[4]. EGFR has a 621aa extracellular region, divided into subdomains I-IV[5]. Domains I and III directly participate in ligand binding[6], connected via a 23aa hydrophobic transmembrane α-helix to a 542aa cytoplasmic domain containing a 300aa tyrosine kinase[7].

EGFR activation requires ligand binding, receptor-receptor interactions, and full activation of the tyrosine kinase[8]. At least 11 different ligands bind to the EGFR family, four to EGFR including EGF itself[9]. Prior to ligand binding the tyrosine kinase has low catalytic activity. Ligand binding results in full kinase activation through c-lobe interaction of an 'activator' and n-lobe 'receiver'[10]. Subsequent autophosphorylation of intracellular tyrosine residues[11] initiates intracellular reactions ultimately stimulating cellular growth, differentiation and proliferation[12], terminated by internalization and proteolytic degradation of the receptor-ligand complex[13].

The field has detailed insights concerning extracellular and intracellular interactions that contribute to signal transduction, however, there remains conflicting evidence concerning the *in vivo* composition of EGFR before and after activation and the role of higher order multimeric complexes of EGFR. Small angle X-ray scattering and isothermal titration calorimetry to EGFR's isolated extracellular domain (sEGFR) suggests that EGF binds to an sEGFR monomer and that receptor dimerization involves subsequent association of two monomeric EGF-sEGFR[14]. Molecular weight determination by multi-angle laser light scattering suggests sEGFR is monomeric in solution but dimeric after addition of EGF[15]. Fluorescence anisotropy indicates a 1:1 binding ratio of EGF:sEGFR, with analytical ultracentrifugation suggesting the complex is comprised of 2(EGF-sEGFR)[16]. Structural evidence suggests activation is preceded by ligand binding to a receptor monomer[17–19], and that EGF induces EGFR conformational change by removing interactions that auto-inhibit EGFR dimerization[20]. This model assumes that EGF binding increases the affinity for subsequent EGF to bind to the free EGFR subunit in the dimer (i.e. positively cooperative). However this is in conflict with EGF-EGFR binding studies of the full length receptor indicating that EGF binding reduces the affinity for subsequent EGF binding to the free EGFR subunit in the dimer [21] (i.e. negatively cooperative) mediated through the dimerization arm and intracellular juxta-membrane domain[22]. Recent structural studies of sEGFR in *Drosophila melanogaster* support a negatively cooperative model[23], and it has been shown that EGFR dimers with a single bound EGF can be phosphorylated[24]. A predication from negative cooperativity is that EGFR:EGF bound complexes have a relative stoichiometry of 2:1 [25].

Chemical crosslinking and immunoprecipitation studies of full length receptors support a preformed dimer model[26], suggesting that receptor dimerization is mechanistically decoupled from activation. Similarly, the first single-molecule fluorescence imaging studies on functional cell membranes suggested initial binding of one EGF molecule to a preformed EGFR dimer, rapidly followed by a second

76   EGF to form a 2:2 complex[27]. Förster resonance energy transfer (FRET) studies subsequently reported
77   preformed oligomeric EGFR[28] supported by other live cell microscopy[29], autocorrelation[30], bimolecular
78   fluorescence complementation (BiFC)[31], fluorescence cross-correlation combined with FRET[32], mobility
79   measurements of quantum dot tagged EGFR[33] and pixel brightness analysis of GFP-labeled EGFR[34].
80   Recent single-molecule photobleaching analysis suggests that EGFR forms oligomers prior to EGF
81   binding[35], and that EGFR clustering may be triggered at physiological EGF levels[36], which contradicts live
82   *Xenopous* oocyte studies that report a significant population of monomeric EGFR present before EGF
83   activation[37]. The observed clustering of EGFR is not unique, but a general feature of cell membrane
84   receptors in signal activation[38]. However, the EGFR clustering is nuanced in that it may involve
85   cooperativity not only between monomer subunits of EGFR molecules in a dimer, i.e. an EGFR
86   homodimer, but also between other ErbB receptor monomers of a different class, i.e. heterodimers[31,34].

87   EGFR's oligomeric state before and after activation under physiological conditions remains an
88   open question due to technical limitations in obtaining simultaneous information for the relative
89   stoichiometry of interacting receptors and ligands, the sensitive dependence of EGF expression levels on
90   the EGFR state of oligomerization, the presence of both fluorescently labeled and natively unlabeled
91   EGFR, and species-specific differences of model immortalized cell lines. Previous fluorescence
92   microscopy studies on live cells have used non-epithelial immortalized rodent sources of mouse (BaF/3,
93   B82, NIH/3T3) and hamster (CHO-K1). There have also been studies using human epidermoid carcinoma
94   cells (A431, BT20, A549 and H460). All of these strains have measurable native levels of EGFR
95   expression; in the case of the most commonly used A431 strain a staggering 2-6 million receptors per cell.
96   Similarly, recent single-molecule investigations using transfected GFP-labeled EGFR in *Xenopus* oocytes
97   may still exhibit appreciable expression levels of unlabeled native EGFR since their membrane surface
98   forms microvilli in which EGF receptors localize[37]. Here, instead, we investigate a human epithelial
99   carcinoma cell line, with no detectable native EGFR, to improve our understanding of EGF binding to
100 EGFR in human cancer cells. We overcome previous technical limitations of simultaneous receptor and
101 ligand measurements using single-molecule dual-colour total internal reflection fluorescence (TIRF)
102 microscopy on live human colorectal carcinoma cells into which GFP-labelled EGFR has been stably
103 transfected, coupled to real time nanoscale tracking of the red/orange dye tetramethylrhodamine (TMR)
104 conjugated to EGF (Fig. 1a). We present results in the presence and absence of cetuximab[52] or
105 trastuzumab[41], two popular immunotherapy antibodies which inhibit EGF signalling. We find that EGFR
106 forms oligomeric clusters prior to EGF binding, with a mode peak stoichiometry of 6 EGFR molecules per
107 cluster. After EGF binding, we observe clusters containing both EGFR and HER2. These are consistent
108 with negative cooperativity for EGFR activation by EGF [21], resolving a key question in the field.

109

110 **Results**

111 **Construction of EGFR-GFP carcinoma cells.** Human epithelial cell line SW620 was selected from an
112 extensive colorectal carcinoma library for its undetectable EGFR expression as quantified by DNA
113 microarray[42] (Supplementary Fig. 1) and western blot (Fig. 1b). SW620 was stably transfected with
114 plasmid pEGFR-EGFP-N1 to give SW620-EGFR-GFP (we denote EGFP throughout as simply 'GFP'),
115 GFP tagging the cytoplasmic domain far from the EGF binding site. Confocal microscopy of live cells

116 confirmed membrane localization (Fig. 1c) with immunofluorescence on fixed cells demonstrating
117 colocalization with EGFR (Supplementary Fig. 2a-d).

118
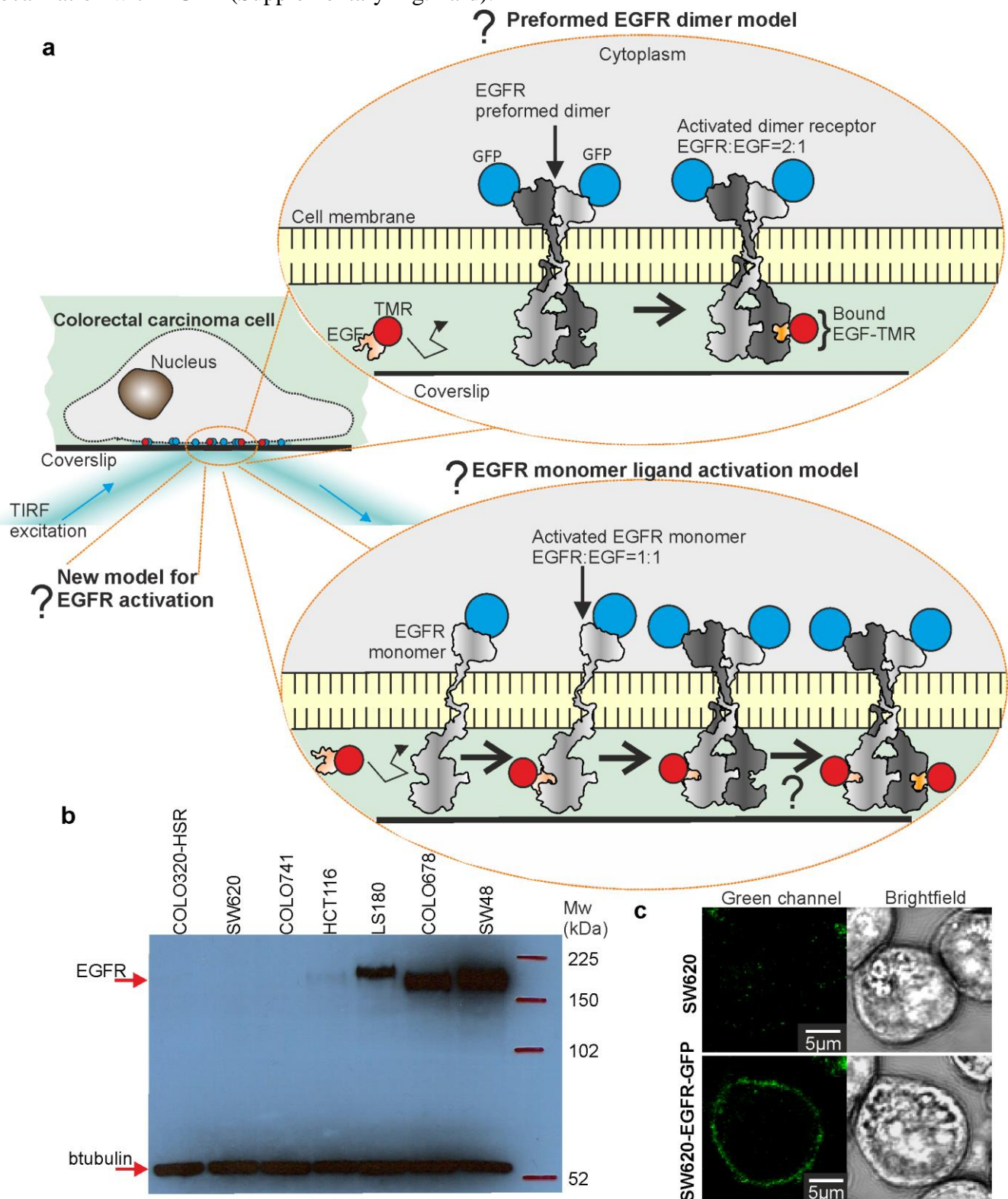
Figure 1. Visualizing functional EGF-EGFR complexes in human carcinoma cells. (**a**) Dual-colour
TIRF applied to EGFR-GFP transfected human colorectal carcinoma cells with and without presence of
fluorescently-labelled EGF-TMR. Several models to explain EGF activation of EGFR have been
postulated, including 'monomer' and 'preformed dimer' models (EGFR structure PDB ID 1egf; EGFR

123 monomer and dimer cartoons have been generated by manually combining separate structures with PDB
124 ID values of 1nql, 1ivo, 2jwa, 1m17and 2gs6). (**b**) SDS-PAGE taken for several candidate colorectal
125 carcinoma cell lines, indicating that SW620 COLO320-HSR (as opposed to COLO320-DM, its duplicate
126 line) and COLO741 (later found to be a melanoma line and so not subsequently used here) have negligible
127 native EGFR expression levels compared to positive controls of HCT116, LS180, COLO678 and SW48,
128 shown to have intermediate EGFR expression levels. Note, there is a difference in apparent molecular
129 weight for EGFR between LS180 and COLO678/SW48, most probably due to glycosylation. (**c**) Parental
130 (non GFP) SW620 carcinoma cells show minimal autofluorescence in the green TIRF channel (left panel),
131 while SW620-EGFR-GFP show membrane localization for EGFR-GFP (right panel).

132

133 **TIRF optimized for single-molecule detection of EGF and EGFR.** We optimized a bespoke dual-colour
134 TIRF microscope (Supplementary Fig. 2e) for single-molecule detection using a fluorophore assay[43] in
135 which either GFP or EGF-TMR are conjugated to a glass coverslip using either IgG antibodies or derived
136 Fab nanobody fragments with binding specificity to GFP or EGF (Supplementary Fig. 3a). We optimized
137 imaging conditions to yield consistent fields of view containing fluorescent foci of GFP or EGF-TMR
138 sampled at a video-rate of 30 ms per frame. Foci had a detectable brightness above background noise and a
139 measured width (defined as half width at half maximum from their pixel intensity profile) in the range 250-
140 300nm (in comparison to the measured point spread function (PSF) width of our microscope of 230nm).
141 After ~1 s of continuous laser illumination foci exhibited irreversible step-wise photobleaching (Fig. 2a),
142 indicative of single molecules of either GFP or EGF-TMR. Each focus had a brightness (summed pixel
143 intensity integrated over each focus) of ~2,000 counts on our detector (Supplementary Fig. 3b). Although
144 each IgG molecule contains two Fab sites, we saw no statistically significant difference in the number of
145 two-step photobleach traces compared to Fab nanobody fragments, suggesting that GFP binding to an IgG
146 Fab site may limit accessibility for a second GFP.

147

148 **EGFR is oligomeric prior to EGF binding.** To explore the architecture and dynamics of functional
149 EGFR we used single-molecule TIRF on live SW620-EGFR-GFP cells. Prior to adding EGF in serum-free
150 medium we observed several fluorescent foci in the GFP detection channel at a low surface density of 0.1-
151 0.4 per μm^2 in the plasma membrane (Fig. 2b and Supplementary Fig. 4). We tracked a mean of 66 ± 28
152 (s.d.) foci per cell and monitored their spatiotemporal dynamics over several seconds to a precision of
153 ~40nm using bespoke software[44,45], indicating a range of mobility (Supplementary Video 1). Foci widths
154 were within ~10% of those observed for single GFP *in vitro,* however, brightness values were far greater.
155 Foci brightness *vs.* time during tracking exhibited steps characteristic of stochastic photobleaching of one
156 or more GFP within a single sampling time window (Fig. 2d), which we used to determine stoichiometry
157 in terms of number of EGFR-GFP molecules present[43]. To estimate stoichiometry, initial foci brightness
158 values were determined by interpolation to the start of each acquisition then divided by the *in vivo*
159 brightness for a single GFP. To determine GFP brightness *in vivo* we quantified the mean foci brightness
160 towards the end of each photobleach, when only one photoactive GFP molecule remained. Our analysis
161 indicates that GFP brightness in a live cell is within 15% of that measured *in vitro* (Supplementary Fig.
162 2b). Previous live cell measurements using the same fluorescent protein indicate that the proportion of
163 immature GFP is less than 15% of the total[56]. We measured a broad range of stoichiometry, both across

164 different cells and within the same cell, of 2-90 EGFR molecules per fluorescent focus, with a peak integer
165 value of 6 and associated mean of 12.8 ± 0.4 molecules ($\pm$s.e.m.) (Fig. 2e).

166   Since our microscope has the sensitivity to detect single GFP, one important conclusion is that
167 there is no significant population of monomeric EGFR before adding EGF. The cell line has no detectable
168 native EGFR expression, so our findings have consistency with a preformed dimer and/or oligomer model
169 for EGFR[46] as opposed to dimer formation being stimulated by EGF binding to monomeric EGFR, or
170 where EGFR dimers are stabilized by two bound EGF[14]. We wondered if the observed stoichiometries
171 could be due to random overlap of diffraction-limited images of individual EGFR-GFP foci. To address
172 this question we modelled foci separation as a Poisson distribution[47] (Methods), and used these to simulate
173 apparent EGFR stoichiometries. We simulated monomeric, dimeric, and mixed oligomeric EGFR
174 (monomers through to tetramers, suggested from a previous single-molecule live cell study[35]), all with
175 poor agreement to the experimental data (Supplementary Fig. 5a, $R^2 < 0$). We then tried a heuristic Monte
176 Carlo overlap model (Methods) that simulated oligomeric EGFR whose stoichiometry was sampled
177 randomly from a Poisson distribution with mean value equal to the peak of 6 that we observed, which
178 resulted in a reasonable fit to the experimental distribution (Supplementary Fig. 5b, $R^2 = 0.4923$).
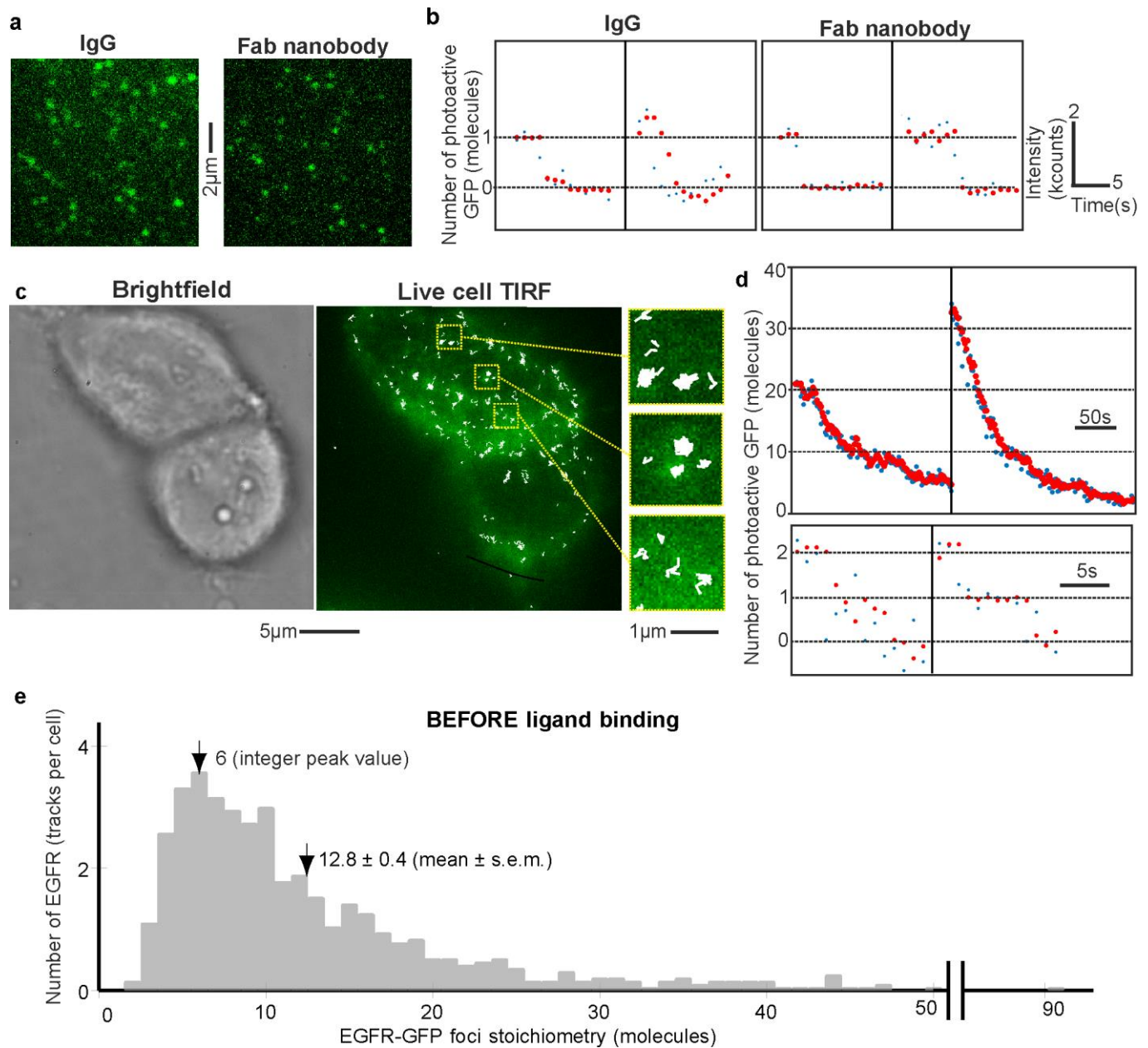
Figure 2. Stoichiometry of EGFR before EGF binding. (**a**) TIRF images of surface-immobilized GFP *in vitro* using IgG and Fab nanobody conjugation. (**b**) Example step-wise photobleach traces show raw (blue) and output data of an edge-preserving Chung-Kennedy filter[48,49], kcounts equivalent to counts on our camera detector x 10^3. (**c**) Example of two nearby SW620-EGFR-GFP cells showing GFP fluorescence (green) and overlaid tracking output (white) with zoom-ins (inset). (**d**) Example photobleach traces from tracked EGFR-GFP foci which have stoichiometries of several tens of EGFR molecules (upper panel), down to an observed minimum of just two molecules (lower panel), raw and overlaid filtered data shown. (**e**) Distribution of EGFR-GFP foci stoichiometry before EGF activation, showing a modal peak at 6 and mean ~12.8 molecules. Data extracted from N=19 cells, detecting N=1,250 foci tracks, corresponding to mean of ~780 EGFR molecules per cell.

191   **EGF binding to EGFR is negatively cooperative.** To determine the effect of EGF binding on EGFR
192   stoichiometry and spatiotemporal dynamics, live SW620-EGFR-GFP cells and non-GFP controls were
193   kept in serum-free media for 12-24 h to minimize binding of any serum-based EGFR ligands. We
194   visualized cells using dual-colour TIRF then added EGF-TMR, enabling simultaneous observation of
195   EGFR and EGF in separate green and red colour channels respectively, before and after EGF activation.
196   Excess EGF-TMR was retained in the sample chamber during imaging enabling observations over
197   incubation times from 3-60 min. We observed a mean of 82 ± 36 EGFR foci tracks per cell across all
198   incubation times, significantly higher than when EGF was absent. Colocalization of EGFR and EGF foci
199   was determined using numerical integration between overlapping green and red channel foci[47].

200   After EGF incubation from as little as a few minutes, colocalization between green and red channel
201   foci was clearly detected (Fig. 3a, Supplementary Video 2 and Supplementary Fig. 6a). We estimated that
202 $40 \pm 18\%$ of foci were colocalized EGFR-EGF when calculated across the full 60 min incubation, ~15 foci
203   per cell or 64% of all EGFR molecules (Fig. 3b,c). EGFR-EGF foci had a statistically higher mean
204   stoichiometry (Student's t-test P<0.0001) of ~31 EGFR molecules compared to isolated receptors whose
205   mean stoichiometry was ~11 EGFR molecules, consistent with measurements made before adding EGF
206   indicating that effects from putative non-EGF ligands in the serum-free media were negligible (Fig. 3d,
207   Table 1). The mean stoichiometry of isolated EGFR clusters remained roughly constant in the range ~8-14
208   molecules during incubation with EGF (Fig. 3e). The mean stoichiometry of EGFR-EGF clusters increased
209   to ~32 EGFR molecules 10-15 min after adding EGF, up to a peak of ~70 EGFR molecules after ~40 min.
210   At higher times EGFR endocytosis is prevalent[50], consistent with observing some brighter EGFR foci in
211   the main body of the cell, which may account for lower mean stoichiometry values of ~20-30 EGFR
212   molecules per focus from ~40 min onwards.

213

214

Biochemical intervention			EGFR foci stoichiometry, uncolocalized		EGFR foci stoichiometry, colocalized with EGF		N cells
E	C	T	Mean ± s.e.m (molecules per EGFR focus)	N foci	Mean ± s.e.m (molecules per EGFR focus)	N foci	
-	-	-	12.8±0.4	770	X	X	19
+	-	-	10.8±0.2	4,741	31.1±1.1	1,969	117
-	+	-	19.9±1.0	531	X	X	10
-	-	+	15.3±0.7	408	X	X	10
+	+	-	18.8±0.5	916	51.0±2.1	303	25
+	-	+	16.8±0.4	1,273	44.2±2.4	334	27

215

216 **Table 1. Mean EGFR foci stoichiometry values.** Number of tracked foci in total (N foci) and individual

217 cells (N cells) in datasets indicated. Biochemical interventions for added EGF (E), cetuximab (C), and

218 trastuzumab (T) shown.

219

220 EGF-TMR quantified *in vitro* using conjugation to glass coverslips exhibited similar step-wise

221 photobleaching as for GFP (Supplementary Fig. 3b). To determine the relative stoichiometry between

222 EGFR and EGF when EGF is bound (i.e. the activated state) we measured red channel stoichiometry

223 simultaneously to the green channel for EGF-EGFR foci. This analysis revealed a clear peak

224 corresponding to a relative stoichiometry for EGFR:EGF of 2:1 (Fig. 3f, which pools data into integer

225 width histogram bins). By using the measured variability in GFP and TMR brightness we estimate the

226 error for the relative stoichiometry is ~0.7, in agreement with the half width at half maximum under the 2:1

227 peak, indicating that the apparent population in the 1:1 peak histogram bin is consistent with measurement

228 error from the 2:1 population. Sub-dividing data by EGF incubation time revealed no significant shift in

229 relative stoichiometry from the 2:1 peak (shown in kernel density estimations of Supplementary Fig. 6b

230 where data has not been pooled into integer histogram bins). Before EGF-TMR was added in control

231 experiments to the parental (non-GFP) strain we detected a small number of autofluorescent foci in red and

232 green channels resulting in pseudo colocalization of ~2-3 tracks per cell (~3% of all colocalized foci).

233 These pseudo colocalized tracks resulted in a small peak for the apparent relative stoichiometry in

234 green:red colour channels equivalent to ~0.5:1 (Supplementary Fig. 6c), thus had a negligible impact on

235 the measurements of the 2:1 peak. Adding EGF-TMR to this strain indicated foci detection levels in the red

236 channel which were statistically indistinguishable (Student's *t*-test P>0.05) to those measured in the

237 absence of EGF-TMR (Fig. 4a).

238 Our findings indicate that the most likely receptor-ligand complex is a singly ligated EGFR dimer,

239 consistent with a negatively cooperative mechanism for EGFR activation (Fig. 1a, upper schematic), i.e. a

240 multiple EGF binding exclusion effect[21]. An alternative model consisting of initial EGF binding to

241 monomeric EGFR to generate an activated state predisposed to form dimeric EGFR[17–19] (Fig. 1a, lower

242 schematic) predicts a significant 1:1 population, contrary to our observations. With this model, the

243 proportion of 1:1 relative to 2:1 states might be expected to increase with longer EGF incubation times

244 since there are putative steps directly dependent on the EGF on-rate, however, we observed no such
245 dependence.
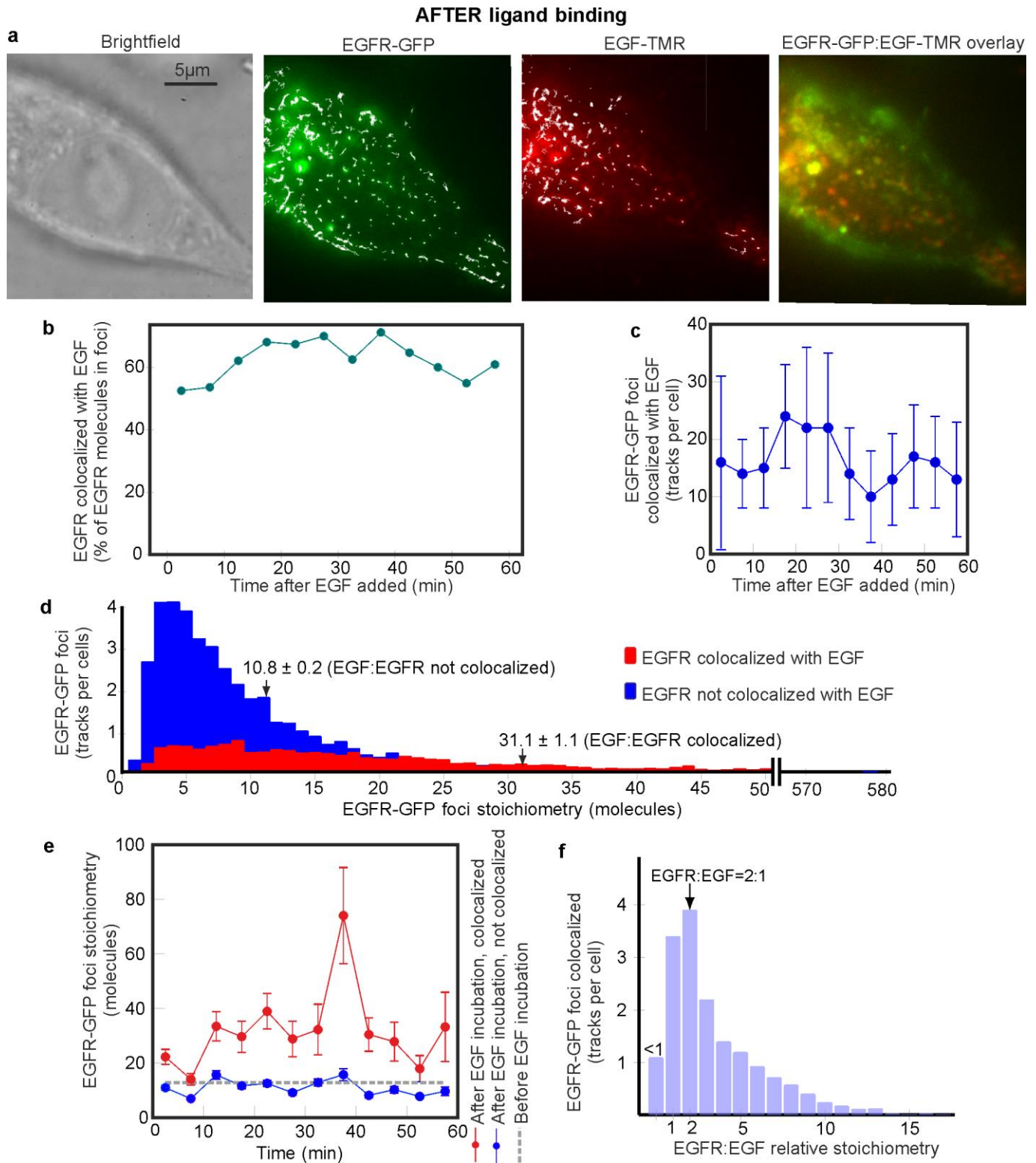
Figure 3. Effect of EGF binding on EGFR stoichiometry. **(a)** Brightfield and TIRF images of SW620-
EGFR-GFP after adding EGF (~10 min incubation time point), GFP (green), TMR (red) and overlay
images shown (yellow indicates high colocalization). **(b)** % of EGFR foci colocalized to EGF, **(c)** number
of EGFR-EGF foci detected per cell (s.d. error bars). **(d)** EGFR-EGF foci stoichiometry (red) and isolated

251   EGFR foci (blue) across all EGF incubation times, mean and s.e.m. indicated (arrows), and **(e)** as a

252   function of incubation time (s.d. error bars). We categorized cells into 6 min interval bins resulting in

253   N = 6-12 cells in each bin. **(f)** Distribution of relative stoichiometry of EGFR:EGF, integer bin widths,

254   peak value at 2:1 indicated (arrow). Data extracted from a total of N = 119 cells.

255

256   **EGFR clustering increases through direct and indirect EGF inhibition.** To further understand the

257   effect of EGF binding on EGFR clustering we performed live cell TIRF in the presence of cetuximab or

258   trastuzumab. Cetuximab is a monoclonal antibody anti-cancer drug commonly used against neck and

259   colorectal cancers in advanced stages to inhibit cell division and growth[51], binding to domain III of the

260   soluble extracellular region of EGFR, and believed to result in partial blockage of the EGF binding region

261   as well inhibiting the receptor from adopting an extended conformation which may be required for EGFR

262   dimerization[52]. Trastuzumab is also a monoclonal antibody anti-cancer drug, commonly used to treat breast

263   cancer[52], with similar effects of inhibiting cell division and growth, however, it does not bind directly to

264   EGFR but instead to domain IV of the extracellular segment of HER2/neu, and its inhibitory action is

265   believed to be related to the association of EGFR and HER2/neu in the plasma membrane[41].

266   Before adding EGF we found that treatment with cetuximab or trastuzumab at concentration levels

267   comparable to those used in cancer treatment resulted in a statistically significant increase in the mean

268   stoichiometry of EGFR-GFP foci by ~25% and ~65% (Student's *t*-test, P<0.0001) respectively (Fig. 4a),

269   but with no effect on the number of detected EGFR-GFP foci per cell. Adding EGF resulted in ~20%

270   fewer EGFR-EGF foci for cetuximab- or trastuzumab-treated cells compared to untreated cells (Fig. 4b).

271

272

273
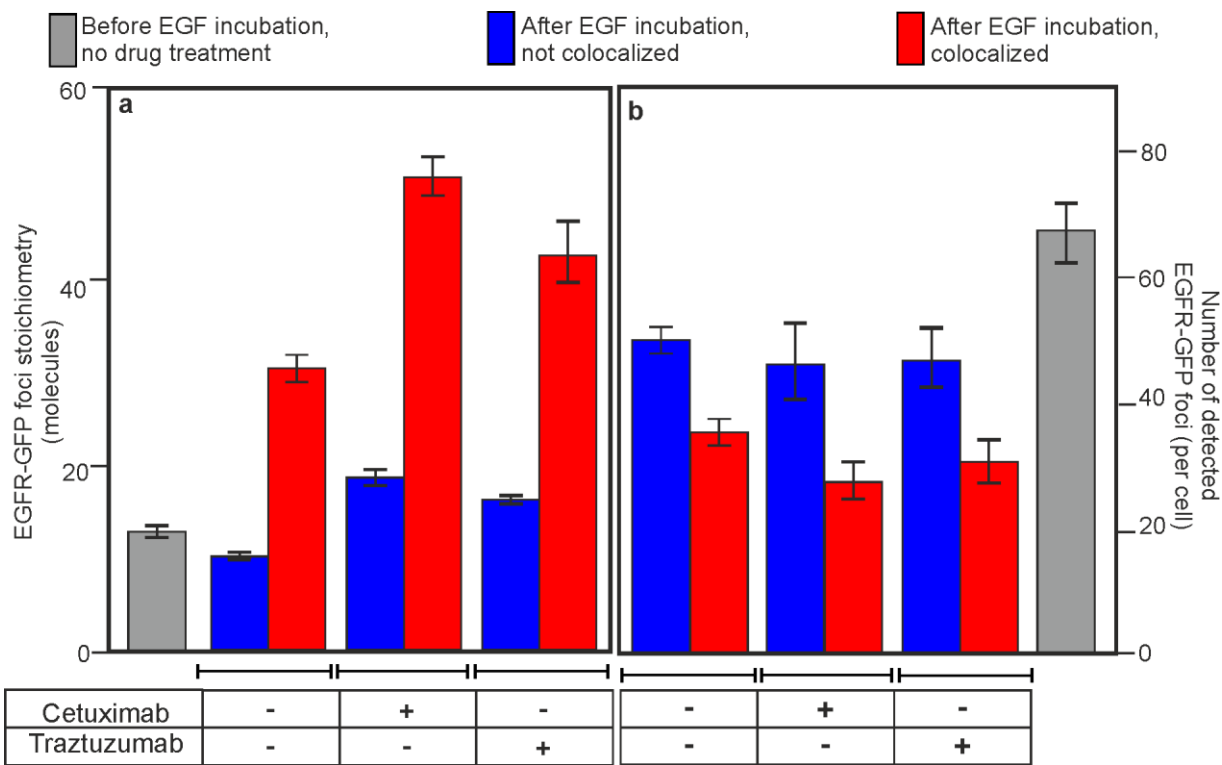
Figure 4. Effect of cetuximab and trastuzumab on EGF binding to EGFR. (a) Variation of mean EGFR-GFP foci stoichiometry, and **(b)** number of EGFR-GFP foci detected per cell. EGFR-EGF (red) and isolated EGFR foci (blue) are indicated for +/- addition of cetuximab and traztuzumab. Errror bars are s.d, number of cells per dataset in the range N $=10-117$.

The mean stoichiometry of EGFR-EGF foci in cetuximab and trastuzumab treatment datasets is 51 ± 2 and 44 ± 2 EGFR molecules per focus respectively, with the upper end having values of several hundred molecules (Fig. 5a, Table 1), consistent with previous qualitative observations that several different EGF pathway inhibitors increase EGRF clustering[53,54]. We also observed a shift to higher EGFR:EGF relative stoichiometry values for both cetuximab and trastuzumab treatments beyond the 2:1 peak observed for untreated cells (Fig. 5b).

284
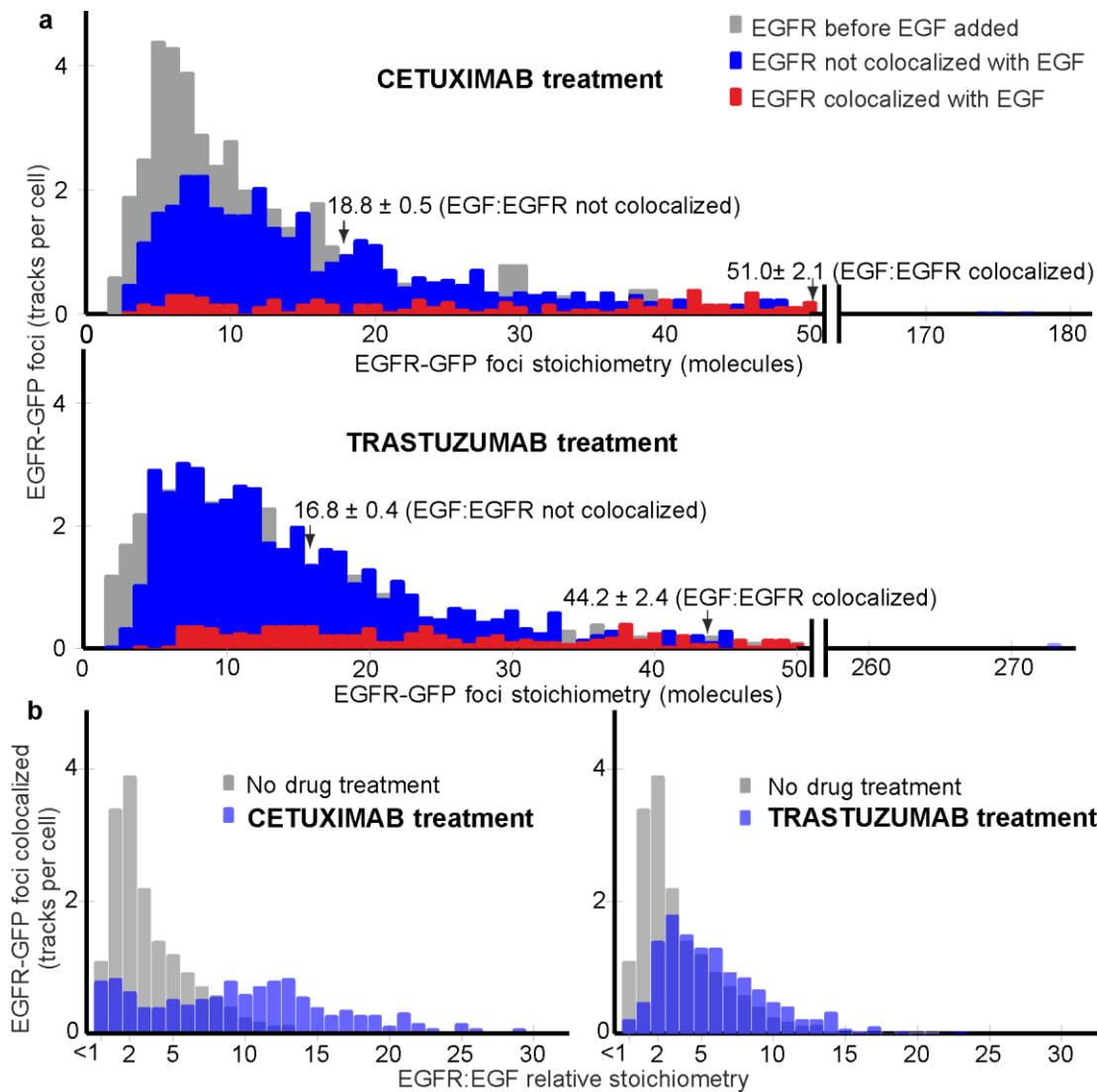
Figure 5. Effect of cetuximab or trastuzumab on EGFR foci stoichiometry. (a) Distribution of EGFR foci stoichiometry for cells treated with cetuximab or trastuzumab, showing pre (grey) and post EGF addition for EGFR-EGF (red) and isolated EGFR (blue) foci, data collated across 60 min EGF incubation time, mean and s.e.m. indicated (arrows). **(b)** EGFR:EGF relative stoichiometry of EGFR-EGF foci for drug-treated cells (blue) contrasted against no drug treatment (grey). Number of cells per dataset in the range N =10 − 117.

EGF can trigger formation of larger EGFR heteroclusters. Tracking of EGFR foci indicated complex mobility in the plasma membrane: Brownian diffusion up to tracking time intervals of ~100 ms (Fig. 6a), transiently confined diffusion into zones of diameter ~400-500nm at time intervals of ~100-600 ms, and Brownian diffusion for time intervals >600 ms (shown indicatively in Supplementary Fig. 7a for the average mean square displacement up to time intervals of several seconds) similar to complicated patterns of diffusion observed previously for membrane proteins interacting with the cytoskeleton[55]. Using the initial gradient of the mean square displacement with respect to tracking time interval for each track we determined the apparent microscopic diffusion coefficient and correlated this against EGFR foci stoichiometry. We used a simple model based on the Stokes-Einstein relation, that the cross-sectional area of an EGFR cluster parallel to the plasma membrane scales linearly with the number of EGFR dimers

303 present. The model assumes that the diffusion coefficient D is given by k_BT/γ where k_B is Boltzmann's
304 constant, T the absolute temperature and γ the frictional drag of the whole EGFR cluster in the membrane.
305 The frictional drag is proportional to the effective radius of the EGFR cluster, which implies that D is
306 proportional to the reciprocal of the square root of the stoichiometry. Our model results in reasonable
307 agreement for data corresponding to pre and post EGF incubation (Fig. 6b).
308
309
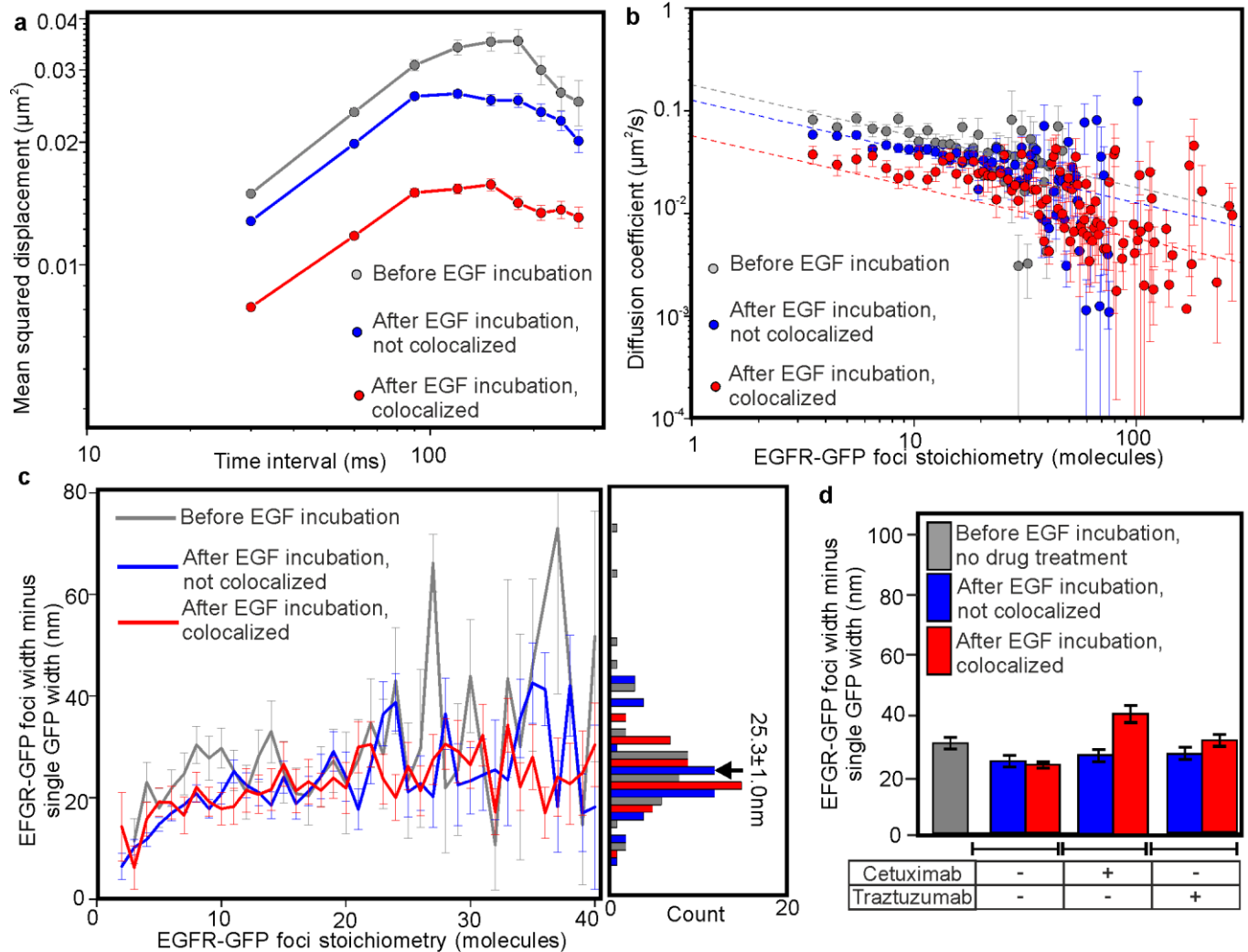
310
311 **Figure 6. EGFR foci mobility depends on stoichiometry and EGF binding. (a)** Log-log plot for
312 average mean squared displacement for time intervals of 300 ms or less, and **(b)** log-log plot for apparent
313 microscopic diffusion coefficient D with EGFR stoichiometry S, fits shown to Stokes-Einstein model
314 assuming $D\sim S^{-1/2}$ (dashed lines). **(c)** EGFR-GFP foci width minus the width of a single GFP *vs.*
315 stoichiometry, and associated histogram, mean and s.e.m. for all datasets combined indicated (arrow).
316 PreEGF incubation (grey, from N=770 foci, taken from N=19 cells) and post EGF incubation for EGFR-
317 EGF (red, from N=1,969 foci, taken from number N=117 cells) and isolated EGFR (blue, from N=1,741
318 foci, taken from N=117 cells) foci shown, s.e.m. error bars. **(d)** Histograms EGFR-GFP mean foci width
319 minus width of a single GFP. Pre EGF incubation for cells untreated with drugs (grey, from N=1,252 foci,
320 taken from N=19 cells); cetuximab-treated cells post EGF incubation for EGFR-EGF (red, from N=151
321 foci, taken from N=10 cells) and isolated EGFR (blue, from N=1,253 foci, taken from N=10 cells) foci
322 shown; trastuzumab-treated cells post EGF incubation for EGFR-EGF (red, from N=263 foci, taken from

323 N=27 cells) and isolated EGFR (blue, from N=1,479 foci, taken from N=27 cells) foci shown; s.e.m. error
324 bars.

325

326 We quantified EGFR-GFP foci widths by performing intensity profile analysis on background-corrected
327 pixel values over each foci image[56], and compared this with measurements obtained from single GFP *in*
328 *vitro*, as a function of foci stoichiometry S (Fig. 6c). In all cases the mean EGFR-GFP foci width was
329 greater than that of single GFP, which increased with the number of EGFR-GFP molecules present,
330 consistent with a spatially extended structure. The dependence of this increase could be modelled with a
331 heuristic power law relation S^a with optimized exponent a of 0.27 ± 0.04 (s.e.m.) showing no dependence
332 with EGF activation (Supplementary Fig. 7b), with a mean for all pooled data of 25.3 ± 1.0nm (s.e.m.). At
333 the low end of S the increase in foci width minus single GFP width was ~11-12nm, while at the high end,
334 corresponding in some cases to several hundred EGFR per focus, the increase in width was 30-40nm. Foci
335 widths indicated no significant differences upon addition of either cetuximab or trastuzumab prior to
336 addition of EGF (P>0.05), however, we observed an increase of ~50% for EGFR-EGF foci for cetuximab-
337 treated cells (P<0.001) (Fig. 6d). Cells treated with cetuximab or trastuzumab exhibited a similar shape for
338 the mean square displacement *vs.* time interval relation to untreated cells (Fig. 7a). Both treatment groups
339 also showed reasonable agreement to a Stokes-Einstein model for diffusion, for before and after addition of
340 EGF (Fig. 7b).

341   We used D to directly estimate the physical diameter of EGFR foci. A full analytical treatment
342 models diffusion of membrane protein complexes as cylinders with their long axis perpendicular to the
343 membrane surface[57] requiring precise knowledge of local membrane thickness, however, here we
344 simplified the analysis by calculating the diameter of the equivalent Stokes sphere to generate indicative
345 values of drag length scale. We approximated the frictional drag by $3\pi\eta d$ where d is the sphere diameter,
346 assuming that drag contributions from the extracellular and cytoplasmic components are negligible since
347 the kinematic viscosity η in the plasma membrane is higher by 2-3 orders of magnitude[58]. Using a
348 consensus value of ~270 cP for the effective plasma membrane viscosity, estimated from human cell lines
349 using high precision nanoscale viscosity probes[59], indicates a mean Stokes diameter of ~40-60nm for
350 isolated EGFR. EGFR-EGF foci had a mean Stokes diameter of closer to ~90nm, reduced back to the level
351 for isolated EGFR to within experimental error upon treatment of cetuximab or trastuzumab (Fig. 7c).

352   We then used Stokes-Einstein fits to determine the Stokes diameter corresponding to the EGFR
353 dimer (i.e. a stoichiometry of precisely 2), which indicated values in the range ~7-10nm for isolated EGFR
354 across the cetuximab, trastuzumab and untreated cell datasets (Fig. 7d), broadly consistent with
355 expectations from the crystal structure of dimeric EGFR[17,18]. EGF-EGFR foci corresponding to an EGFR
356 stoichiometry of 2 had Stokes diameters of ~30nm, which were unaffected by cetuximab but reduced by a
357 factor of ~2 almost to the level of isolated EGFR dimers by trastuzumab (Fig. 7d).
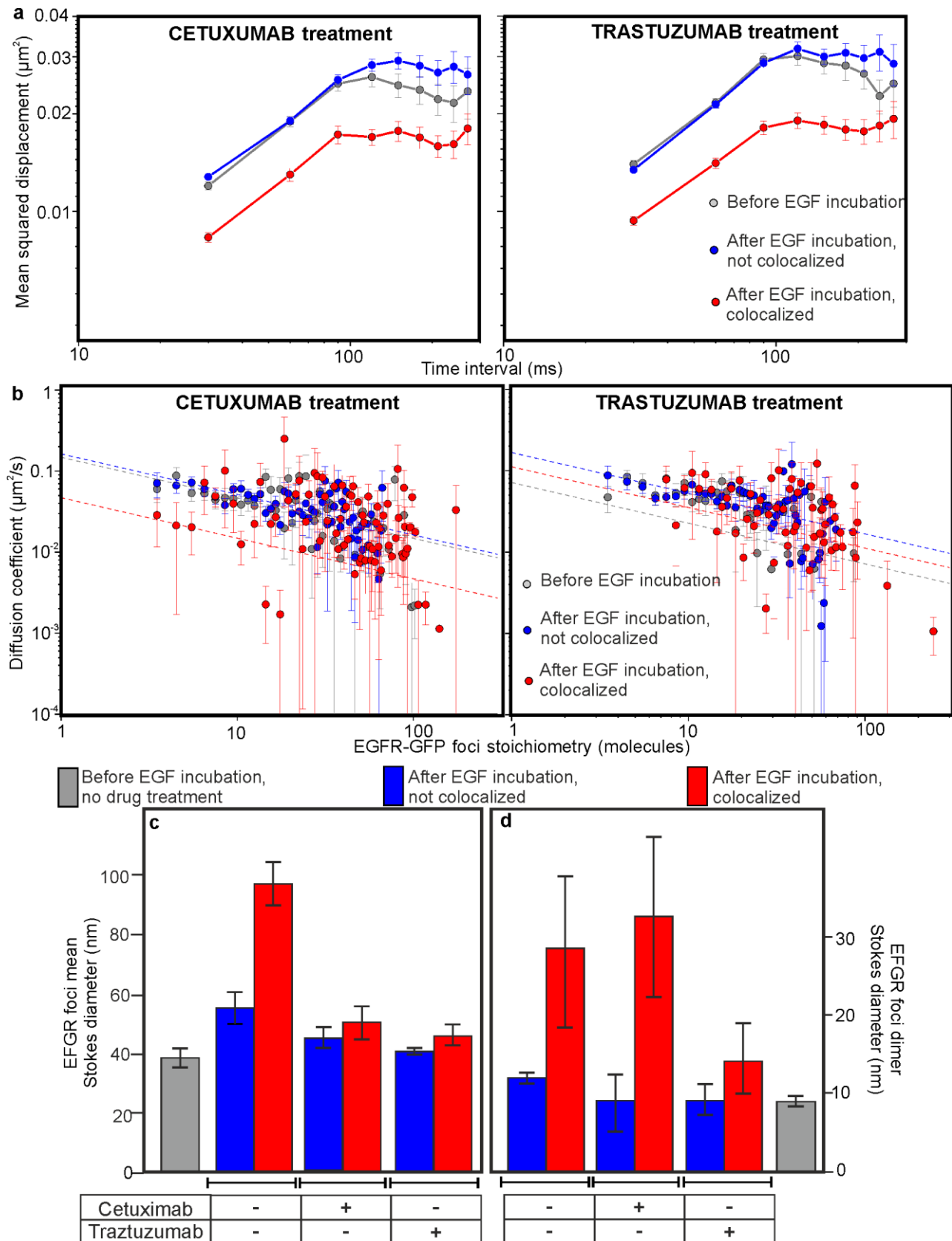
Figure 7. EGFR mobility can be affected by EGF inhibitors. (a) Log-log plots for average mean squared displacement for time intervals of 300 ms or less, and **(b)** log-log plots for variation of apparent microscopic diffusion coefficient D with EGFR stoichiometry S, fits shown to Stokes-Einstein model

362   assuming $D \sim S^{-1/2}$ (dashed lines) for cetuximab- and trastuzumab-treated cells. **(c)** Histogram of mean
363   Stokes diameter, and **(d)** equivalent diameter values extrapolated for EGFR dimeric foci using same
364   datasets as for Fig. 6d, s.e.m. error bars.

365

366   The Stokes diameter for an EGFR cluster is a measure of visible EGFR-GFP content plus any
367   unlabelled associated protein contributing to overall frictional drag. Here, the proportion of non-fluorescent
368   EGFR is low. However, other studies have suggested that EGFR forms heterocomplexes with other RTK
369   receptors[31–33,60]. Here, we observe that treatment with the HER2-binder trastuzumab results in a similar
370   measured Stokes diameter for EGF-EGFR foci to that for isolated EGFR dimers, suggesting that HER2
371   may form heterocomplexes with EGFR following EGF binding (Fig. 8, left panel). Also, since the mean
372   Stokes diameter of EGFR-EFR foci of ~90nm corresponds to a stoichiometry ~32 EGFR molecules, i.e. 16
373   EGFR dimers, then the average diameter associated with a single EGFR dimer which can account for the
374   same cluster area is ~22nm, greater than the measured diameter of an EGFR dimer from crystal
375   structures[17,18] by a factor of ~2. In other words, the observed Stokes diameter could be explained if EGFR-
376   GFP dimers associate in a 1:1 relative stoichiometry with unlabelled HER2 dimers of similar same size and
377   structure.
378   An additional phenomenon to consider is plasma membrane invagination as EGFR clusters grow,
379   ultimately culminating in a clathrin-coated vesicle inside the cytoplasm. Since the visible focus that we
380   detect in TIRF corresponds to the GFP localization pattern in the invaginated membrane projected laterally
381   onto our camera detector then its apparent visible diameter might appear to approach an asymptotic plateau
382   with respect to EGFR-GFP stoichiometry (Fig. 8, right panel), which is broadly what we observed
383   (Fig. 6c).

384

Preformed oligomeric homocluster/activated heterocluster model for EGF receptor platform growth
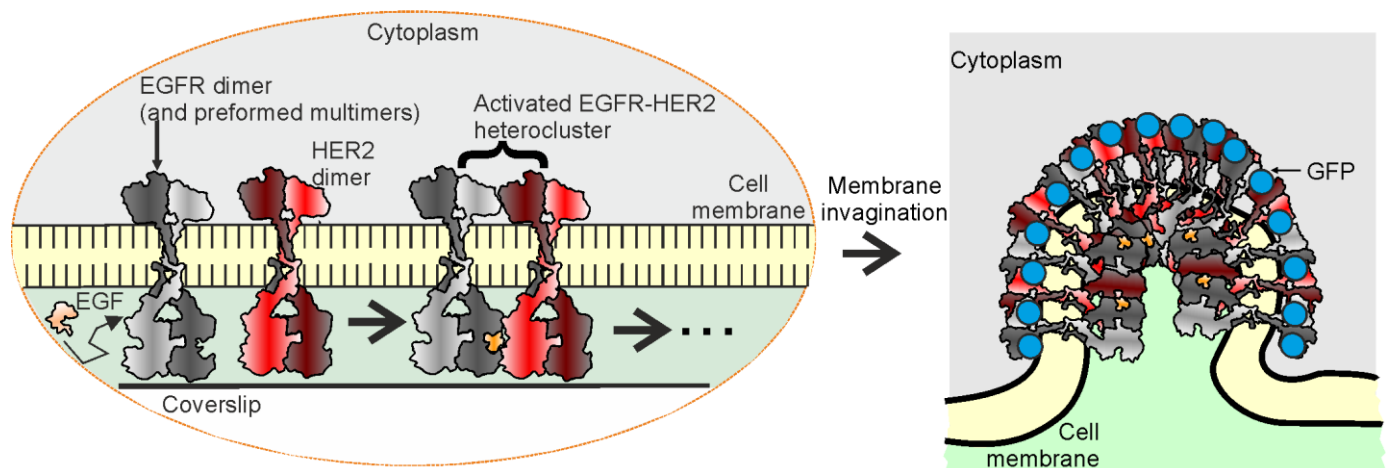
Figure 8. Activated EGFR nanoclusters grow in platforms containing heteroclusters of EGFR and HER2. Schematic illustrating how HER2 and EGFR dimers may be associated following EGFR activation by EGF (left panel) and how further oligomerization may result in local membrane invagination to form hetero receptor 'platforms' of several tens ofnm diameter.

Discussion

In summary, data acquired using genetics, cell biology, biochemical and biophysical techniques, in particular single-molecule imaging with tracking to 40nm precision and quantitative molecular and mobility analysis on live bowel carcinoma cells, suggest that preformed homo-oligomeric EGFR is present in the plasma membrane prior to EGF activation, comprising predominantly clusters of EGFR dimers. These observations are consistent with negative cooperativity of EGF binding to EGFR. Using a GFP probe on EGFR in combination with the spectrally distinct TMR tagged to EGF enabled unparalleled insight into the molecular stoichiometry, mobility and kinetics of single functional EGFR clusters in their pre and post activated states. Our observations indicate that the most prevalent EGFR complex in the absence of bound EGF is a hexamer, though with higher order oligomers also present extending to clusters containing up to ~90 molecules. We find that activation by EGF results in a shift to higher cluster stoichiometry, contrary to earlier speculation from experiments in which just EGF was fluorescently labelled suggesting tetrameric EGFR is the most likely multimeric state[36]. We observe that the action of cetuximab and trastuzumab, commonly used anti-cancer drugs, results in increases in the mean EGFR content of receptor clusters by a factor of 3-5. In addition, our findings suggest that EGF activation generates hetero clusters of EGFR and HER2, a response which results in the formation of super-clusters whose effective diameter is up to ~90nm.

Our findings clearly indicate that EGFR is clustered both before and after EGF activation, consistent with observations from earlier AFM imaging experiments which probed the surface morphology of the human lung adenocarcinoma cell line A549, known to have high levels of EGFR expression in the cell membrane[61]. This study suggested that half of the EGFR clusters quantified had a diameter in the range 20-70nm in the pre-activated state, and 35-105nm post activation, comparable with our light microscopy measurements. However, we find important differences with respect to recent single-molecule studies of EGFR activation[27,34–37]. We observe no significant monomeric population of EGFR before or

after EGFR activation, despite having the sensitivity to detect single GFP under our imaging conditions, though we do observe the presence of single EGF-TMR molecules associated with multimeric EGFR clusters. Two key improvements in our study are that: (i) we specifically selected a human carcinoma cell strain with negligible native EGFR expression, whereas earlier single-molecule studies utilised cell lines likely to have much higher EGFR expression; (ii) unlike earlier studies we have definitive spatial information concerning the localization of EGFR and EGF simultaneously and so have a high level of confidence concerning the effects of EGF binding on the stoichiometry of specific individual EGFR foci. In single-molecule experiments in which EGFR is labelled with a fluorescent protein reporter for which there is some expression of native EGFR even if low[34,37] then apparently monomeric EGFR foci may inevitably be detected even if a functional cluster has a higher stoichiometry, due to mixing of unlabelled and labelled EGFR molecules. In single-molecule experiments in which labelled EGF is not imaged simultaneously with labelled EGFR[27,35,36] then no direct inference can be made as to the relative stoichiometry of associated clusters.

The lack of evidence in our experiments for a monomeric EGFR population coupled to a distinct peak of 2:1 for the EGFR:EGF relative stoichiometry as determined on a unique cluster-by-cluster basis provides clear evidence in support of negatively cooperative binding of EGF to an EGFR dimer. The peak value of 6 EGFR molecules per focus before EGF activation that we observe cannot be explained by a model as proposed previously[36] which suggested that face-to-face dimers associate with the EGFR dimer interface between back-to-back dimers to generate higher order oligomeric complexes; analysis of the steady state solution for this model predicts a most likely stoichiometry of 4 EGFR per focus. Instead, a more likely state of 6 molecules (and higher after activation) suggests a more complex mechanism in which additional EGFR molecules result in greater stability for the overall cluster. This begs a question of what is the driving force behind cluster formation, which we do not directly address here. However, there is evidence from other studies that forces associated with molecular crowding in the membrane may result in oligomerization of integrated membrane proteins and the appearance of complex cytoskeletal and clathrin pit morphologies[34,55,62–64]. Ionic protein-lipid interactions[65] and direct protein-protein interactions[33] have also been implicated as contributory factors towards EGFR cluster generation.

Earlier work on heterocomplex formation in the Erb receptor family has suggested that EGFR and HER2 associate[31,34], however, there are discrepancies in the interpretations of experimental data as to whether this association is before or after EGF activation. Our observations suggest that heterocomplex formation is most likely following EGF activation of EGFR. The physiological role of heterocomplex formation is unclear. HER2 is known to act as coreceptor but has no known direct ligand. However, upon transactivation (i.e. following activation of EGFR by EGF) it exhibits the highest of all kinase activities across the ErbB family[66], thereby augmenting signalling efficiency. The mobility of heterocomplexes may potentially enable a spread of the signal across the surface of the cell, especially if HER2 molecules were to turn over between different EGFR complexes, however, this hypothesis remains to be tested. One consequence for having HER association post EGF binding is that the signal response at the level of the whole cell is more likely to be distinctly binary (i.e. highly biphasic) since the augmentation of the response due to HER2 association after activation results in a very high and rapid signal response. Our findings of post activation heterocomplex formation may suggest potential new strategies for anti-cancer drug design. For example, although there are anti-cancer drugs already established which bind specifically to HER2, one new strategy could be to target the specific interaction interfaces between HER2 and EGFR

457 directly. Alternatively, it may also be valuable to explore new strategies to disrupt the oligomeric nature of
458 the EGFR receptors before EGF activation.

459

460 **Methods**

461

462 **Strain construction.** We screened all ~100 colorectal cancer cell lines from the Cancer and
463 Immunogenetics Laboratory (Weatherall Institute of Molecular Medicine, Oxford University, U.K.)
464 for EGFR mRNA using available microarray data[42] (Supplementary Fig. 1) and selected three
465 preliminary lines (SW620, COLO320HSR and COLO741) on the basis of negligible native EGFR
466 expression levels prior focussing on SW620 (COLO741 was found to be a melanoma line and
467 COLO320HSR exhibited transfection instability). Total protein levels were estimated from cell lysates
468 prepared from pellets using a radio immunoprecipitation assay lysis buffer supplemented with Roche
469 cOmplete Mini ethylenediaminetetraacetic acid free protease inhibitor cocktail and Roche PhosSTOP
470 phosphatase inhibitor cocktail. Total protein concentration was estimated using Thermo Scientific™
471 Pierce™ bicinchoninic acid Protein Assay Kit referenced against known concentrations of BSA. EGFR
472 protein quantification was performed with western blotting, including cell lines with intermediate levels of
473 EGFR expression as positive controls, probing nitrocellulose membranes with anti-EGFR mouse
474 monoclonal antibody (1:1000, clone 1F4, Cell Signalling Technology®) and anti-β-tubulin mouse
475 monoclonal antibody (1:1000, Sigma-Aldrich®) prepared in TBS-T, 5% milk and incubated overnight at
476 4ºC. After the washing, membranes were incubated with secondary antibody (1h, room temperature) using
477 a polyclonal Rabbit anti-Mouse antibody conjugated to horseradish peroxidase (Dako) diluted at 1:10,000
478 and 1:100,000 for respectively EGFR and β-tubulin detection prior to Amersham Biosciences enhanced
479 chemiluminescence (ECL) exposure.

480 Plasmid perbB1-EGFP-N1 (donated by Philippe Bastiaens, Max Planck Institute of Molecular
481 Physiology, Dortmund, Germany) was used for transformations, which comprised an insertion of the
482 human *EGFR* gene into the enhanced GFP Clontech backbone, pEGFPN1, plus selectable kanamycin
483 (*kan*) resistance markers for bacterial/eukaryotic vectors. Competent *E. coli* cells were transformed with
484 pEGFR-EGFP following the Invitrogen manufacturer's protocol and plasmid DNA purified using the
485 QIAGEN Plasmid Mini Kit. The concentration of purified DNA was determined using a Thermo Scientific
486 NanoDrop™ 1000 Spectrophotometer at 230nm wavelength. SW620 cells were transfected with pEGFR-
487 EGFP using Invitrogen's cationic lipid transfection formulation, Lipofectamine® LTX and Plus™ reagent.
488 1 day pre transfection 200,000 cells in 1ml growth medium were seeded into each well of a 12-well plate;
489 the following day DNA-lipid complexes were prepared according to the manufacturer's instructions. For
490 each well we added 2μg plasmid DNA, 200μl Invitrogen Opti-MEM® I Reduced Serum Medium, 1μl Plus
491 Reagent and 6μl of Lipofectamine LTX. DNA-lipid complexes were added dropwise to the cells then
492 placed in a 5% CO_2 37 ºC incubator and the media changed after 5h to the usual cell media. The following
493 day cells were trypsinized by trypsinization and reseeded onto 15cm plates in their usual growth media
494 Gibco® Dulbecco's modified eagle medium (DMEM) supplemented with 4.5g/l glucose, pyruvate, L-
495 glutamine and phenol red plus 2μg/ml Gibco™ Geneticin® (G418 Sulphate) selection antibiotic. Once
496 colonies were visible by naked eye, these were isolated using a silicon Cloning cylinder (Corning®),
497 harvested using trypsin and transferred in a 12-well plate. Transgene expression was confirmed by three

498 different methods of imaging live cells directly with confocal microscopy, imaging immunofluorescently
499 stained fixed cells with confocal microscopy, and western blotting.

500 **Nanobody preparation.** IgG antibodies to EGF and anti-EGF rabbit anti-mouse polyclonal IgG
501 (Molecular Probes) were digested by papain, confirmed by migration of 28-30kDa and 25kDa molecular
502 weight proteins under reducing conditions, corresponding to reduced Fc and Fab, respectively. Fab
503 nanobodies were purified from the digest using protein A immobilized within a spin column. The
504 completeness of IgG digestion and Fab purification were evaluated by measuring absorbance at 280nm
505 wavelength using a Thermo Scientific NanoDrop spectrophotometer. Following purification of the digest
506 a protein band at 25kDa only was seen in the protein A flow-through under denaturing and reducing
507 conditions for both antibodies, consistent with reduced Fab.

508 **Fluorescence microscopy.** For confocal microscopy we used a Zeiss inverted Axio Observer Z1
509 microscope with LSM 510 META scanning module and Plan-Aprochromat 63x 1.40NA oil immersion
510 DIC M27 objective lens, enabling simultaneous imaging of green and red colour channels: excitation path
511 used 488nm wavelength argon ion laser; first detection channel contained a 565nm wavelength dichroic
512 beamsplitter and 505nm longpass emission filter for GFP, second channel collected transmitted light to
513 produce a DIC image. Cells were grown in Corning 75 cm^2 treated plastic cell culture flasks in a
514 humidified incubator at 37 ºC with 5% carbon dioxide. Once cells were 70-100% confluent they were
515 subcultured by enzymatic disaggregation with trypsin. 2-7 days prior to imaging, 150,000-300,000
516 SW620-EGFR-GFP cells were seeded onto Ibidi μ-dish 35mm, high glass bottom using their normal
517 culture medium. SW620-EGFR-GFP cells were either seeded in DMEM containing phenol red, then
518 changed to DMEM with addition of 4.5g/l glucose, L-glutamine, HEPES, without phenol red, and
519 supplemented with 10% FBS, 100 units/ml of penicillin and 100 μg/ml of streptomycin, or SW620:EGFR-
520 GFP cells were seeded directly into DMEM without phenol red. Prior to imaging the media was changed to
521 Molecular Probes® Live Cell Imaging Solution. All media used for SW620-EGFR-GFP cells were
522 supplemented with 1.5mg/ml of G418 sulfate.

523   For immunofluorescent characterization we harvested SW620-EGFR-GFP cells 48h prior to
524 fixation at density of ~50,000 per well seeded into Ibidi μ-Slide VI0.4, cultured in DMEM without phenol
525 red, supplemented with 4.5g/l glucose, L-glutamine, HEPES 10% FBS and 100 units/ml of penicillin and
526 100μg/ml streptomycin, 1.5mg/ml of G418. Cells were fixed with 4% formaldehyde at room temperature
527 for 10min and washed. Non-specific antibody adsorption was blocked with 10% FBS in PBS for
528 10-20min. Primary antibodies were EGFR (D38B1) XP rabbit monoclonal 4267P (Cell Signaling
529 Technology, 1:50 dilution) and anti-GFP chicken IgY (H+L) (Cell Signaling Technology, 1:400 dilution)
530 diluted in PBS with 10% FBS and 0.1% saponin overnight at 4 ºC. Each well was washed with 10% FBS
531 and incubated with secondary antibodies, DyLight 633 goat anti-rabbit immunoglobulin G (IgG) highly
532 cross adsorbed (PN35563, Thermoscientific), dilution 1:200, and Alexa Fluor 633 goat anti-chicken IgG
533 (H+L) 2 mg/ml (Invitrogen) diluted in PBS with 10% FBS and 0.1% saponin. Channels were washed with
534 PBS and Sigma Aldrich Mowiol 4-88 was added to solidify overnight. GFP, DyLight 633 or Alexa Fluor
535 633 and 4',6-diamidino-2-phenylindole (DAPI) were individually illuminated and scanned. Transmitted
536 light images were scanned simultaneously with GFP. GFP was excited as for live cell imaging, while
537 DyLight 633 and Alexa Fluor 633 were excited by a 633nm HeNe laser and detection beam path contained
538 a 565nm secondary dichroic beamsplitter and 650nm longpass filter.

539 For TIRF microscopy, a bespoke dual-colour single-molecule microscope was modified from a previous
540 design[47,67,68] equipped with nanostage (Mad City Labs), samples imaged at 37 ºC in a humidified stage top
541 incubator supplemented with 5% CO_2 (INUB-LPS, Tokai Hit). We used excitation sources of an Elforlight
542 B4-40 473nm 40mW diode laser and Oxxius SLIM 561nm 200mW diode-pumped solid-state laser,
543 independently attenutated and recombined into a common optical path prior to polarization circularization
544 using an achromatic λ/4 plate before entering a Nikon Eclipse-Ti inverted microscope body. An
545 achromatic doublet lens mounted onto a translation stage controlled the angle of incidence into the
546 objective lens to generate TIRF via a Semrock 488/561nm BrightLine® dual-edge laser-flat dichroic beam
547 splitter into a Nikon CFI Apo TIRF 100x NA1.49 oil immersion objective lens. Our imaging system
548 enabled simultaneous GFP and TMR detection across a laser excitation field of full width at half maximum
549 20μm laterally, intensity $1kW/cm^2$ and set depth of penetration ~100nm. Continuous fluorescence
550 emissions were sampled at 30ms per frame and split into green/red channels via a 488/561nm dual-pass
551 dichroic mirror (Semrock) and imaged onto two 512x512 pixel array EMCCD cameras (Andor, iXon+
552 DU-897 and iXon DU-887 for green and red respectively, piezoelectrically cooled to -70ºC) at
553 ~50nm/pixel magnification, via Semrock 561nm StopLine® single notch and Chroma 473nm notch filters.
554 Cells were seeded and grown in culture medium onto glass-bottomed Petri dishes at 37 ºC in humidified
555 5% CO_2, prior to imaging exchanging to Molecular Probes® Live Cell Imaging Solution supplemented
556 with G418 sulfate. When appropriate, EGF-TMR was added to stimulate activation, in addition to EGF
557 inhibitors of cetuximab or trastuzumab (Molecular Probes) at a final concentration of 100 ng/ml.
558 SW620:EGFR-GFP, and native SW620 cells with negligible endogenous EGFR as negative control, were
559 imaged on plasma cleaned glass coverslips (25mm×75mm No. 1.5 D263M Schott) covered by a sterile
560 Ibidi sticky-Slide VI0.4 in a laminar flow hood. 48h prior to imaging cells at a density of ~100,000 in a
561 50μl volume of DMEM without phenol red supplemented with 10% FBS, 100 units/ml Penicillin and
562 100μg/ml Streptomycin, were seeded into each channel of the slide. Prior to imaging the media was
563 changed to DMEM without phenol red supplemented with 100 units/ml Penicillin and 100μg/ml
564 Streptomycin but without the addition of FBS, supplemented with 1.5mg/ml G418 sulfate, serum starving
565 the cells for ~12-24h prior to imaging to remove serum EGF. Although we cannot entirely exclude residual
566 amounts of non-EGF EGFR ligands we checked the SW620 cell line for secretion of the most common
567 ligands, indicating: EGF: not expressed; TGFA: low level expression; HBEGF: low level expression;
568 AREG: not expressed; BTC: not expressed; EREG: not expressed; EPGN: no data available. Fluorescence
569 image sequences and a brightfield image were acquired immediately after adding EGF-TMR (Molecular
570 Probes) where appropriate to a final concentration of 100ng/ml, acquiring images at 5min intervals up to
571 60min.
572   For single-molecule *in vitro* TIRF we used surface-immobilized GFP or EGF-TMR via anti-GFP or
573 anti-EGF antibodies (Molecular Probes) respectively followed by BSA to passivate the surface prior to
574 washing off[69]. Whole IgG has in principle two binding sites and to test if two fluorophores may be seen in
575 the same diffraction-limited fluorescent spot we also prepared an antigen binding fragment (Fab) with only
576 one binding site. In brief, slides were constructed from Ibidi sticky-Slides VI0.4 and 25mm×75mm No. 1.5
577 D263M Schott glass coverslip. The coverslip was plasma-cleaned prior the 50μl of whole IgG or Fab
578 applied to a single channel and incubated at room temperature for 5min. Channel then washed three times
579 with 120μl of PBS and the remaining surface blocked with 50μl 1mg/ml of BSA for 60min. The channel
580 was again washed three times with 120μl of PBS and then incubated with 50μl GFP for 7.5min or EGF-
581 TMR for 4min. Finally the channel was washed five times with 120μl of PBS before applying 50μl
582 1:10000, 200nm diameter, 4% w/v, Invitrogen Molecular Probes carboxyl latex beads. These beads could
583 be visualised in brightfield illumination for focussing to avoid using the GFP or TMR itself to focus on
584 which would result in photobleaching. The slides were left 1-12h to allow latex beads to settle. Automated
585 detection of fluorescent foci indicated no significant difference between brightnesses (Supplementary Fig.
586 3b) for the whole IgG or nanobody conjugation methods. We estimated the mean Gaussian sigma width for

587 single GFP fluorescent foci images to be 230nm for GFP, a value which we interpret as the point spread
588 function width of our imaging system corresponding to a peak emission wavelength of ~500nm.

589 **Foci tracking.** Bespoke code written in MATLAB (Mathworks)[44,47] was used to track single fluorescent
590 foci in green and red channels to determine spatial localization and calculate integrated foci pixel
591 intensities and microscopic diffusion coefficients. The centroid of each fluorescent focus is determined
592 using iterative Gaussian masking to a sub-pixel precision of ~40nm. The focus brightness is calculated as
593 the sum of the pixel intensities inside a 5-pixel-radius region centred on the centroid, after subtraction of
594 local background intensity. The signal-to-noise ratio (SNR) for a fluorescent focus is defined as the total
595 focus intensity per pixel divided by the standard deviation of the background intensity per pixel. When the
596 SNR for a focus is >0.3, the focus is accepted and fitted with a 2D radial Gaussian function to determine
597 its Gaussian sigma width. We decided on an SNR threshold level of 0.3 as a compromise between a high
598 probability for true positive detection but a low likelihood for false positive detection at single-molecule
599 fluorophore intensity levels. We simulated fluorescent foci as 2D Gaussian functions in bespoke code
600 written in MATLAB with comparable integrated pixel intensity values and widths as for those measured
601 experimentally for single GFP/TMR used in the surface-immobilization assays, adding similar levels of
602 Poisson-distributed noise, and ran these synthetic data through the same foci detection algorithms as for
603 real experimental data, but exploring a range of SNR detection thresholds. We found that a threshold of 0.3
604 gave a true position detection probability of approaching 50% over a signal range corresponding to 1-10
605 fluorophores per focus, but with a false positive detection probability an order of magnitude less.

606 Foci detected in the tracking algorithm in consecutive image frames separated by 5 pixels or less
607 (approximately the point spread function width of our imaging system), and which are not different in
608 brightness or sigma width by more than a factor of two, are linked into the same track.

609 **Stoichiometry analysis.** Stoichiometry per fluorescent focus was estimated in bespoke code written in
610 MATLAB using integrated intensities and step-wise fluorophore photobleaching with Fourier spectral
611 analysis to determine the brightness of either GFP or TMR during live cell imaging[69]. The brightness of a
612 single GFP or TMR in our microscope was determined from *in vivo* data and corroborated using *in vitro*
613 immobilised protein assays. The brightness of tracked foci in live cells followed an approximately
614 exponential photobleach decay function of intensity with respect to time. Every fluorescent foci as it
615 photobleaches to zero intensity will emit the characteristic single GFP brightness value, I_{GFP}, in the case of
616 EGFR-GFP, and I_{TMR} in the case of EGF-TMR, given in our case by the modal value of all spot intensities
617 over time, and can bleach in integer steps of this value at each sampling time point. Estimates for I_{GFP} and
618 I_{TMR} were further verified by Fourier spectral analysis of the pairwise distance distribution[69] of all spot
619 intensities which yielded the same value to within measurement error. The initial intensity I_0 was estimated
620 by interpolation of the first 3 measured data points in each focus track. Stoichiometries were obtained by
621 dividing I_0 by of a given focus track by the appropriate single-molecule fluorophore brightness.
622 Stoichiometry distributions were rendered as Gaussian kernel density estimations[69] using standard
623 MATLAB routines.

624 **Mobility analysis** For each accepted focus track, the 2D mean square displacement (MSD) was calculated
625 in bespoke code written in MATLAB from the fitted focus centroid at time t, $(x(t),y(t))$, assuming a track
626 of N consecutive image frames, and a time interval $\tau = n\Delta t$, where n is a positive integer and Δt the frame
627 integration time[70]:

628
$$MSD(\tau) = MSD(n\Delta t) = \frac{1}{N-1-n}\sum_{i=1}^{N-1-n}\left\{[x(i\Delta t + n\Delta t) - x(i\Delta t)]^2 + [y(i\Delta t + n\Delta t) - y(i\Delta t)]^2\right\}$$

$$= 4D\tau + 4\sigma^2$$

629   The localisation precision from our tracking algorithm (i.e. on the x-y image plane) is given by σ, which
630   we estimate as 40 ± 20nm. The apparent diffusion coefficient D is estimated from a linear fit to the first
631   three data points in the MSD curve as a function of τ (i.e. $1 \leq n \leq 3$, corresponding to the linear region of
632   the average MSD $vs.$ τ plot) for each accepted track, the fit constrained to pass through a point $4\sigma^2$ on the
633   vertical axis corresponding to $\tau = 0$, allowing σ to vary in the range 20 - 60nm in line with the measured
634   range during the fitting optimisation.

635   **Colocalization analysis** The extent of colocalization was quantified using foci overlap integration between
636   green and red channels[47] determined by calculating the overlap integral between each green/red pair in
637   bespoke code written in MATLAB, whose centroids were within 5 pixels of each other (i.e. green/red pairs
638   in close proximity). In brief, assuming two normalized, 2D Gaussian intensity distributions $g_1(x, y)$ and
639    $g_2(x, y)$, centered around (x_1, y_1) with width σ_1, and centred around (x_2, y_2) with width σ_2 for green and red
640   foci respectively, the overlap integral v can be analytically calculated as[47]:

641    $$v = \exp\left(-\frac{\Delta r^2}{2(\sigma_1^2 + \sigma_2^2)}\right)$$

642   where:

643    $$\Delta r^2 = (x_1 - x_2)^2 + (y_1 - y_2)^2.$$

644   Our simulations indicate that a green/red foci pair that have identical centroid coordinates can have a
645   measured overlap integral as low as ~0.75 due to the finite localization precision of 40nm. Therefore, we
646   used an overlap integral threshold of ≥ 0.75 as a criterion for colocalization for the experimental data.

647   **Modelling the overlap probability of EGFR-GFP foci images.** The probability that two or more
648   fluorescent foci are within the diffraction limit of our microscope was determined in bespoke code written
649   in MATLAB using a previously reported model[47] at foci surface density values observed here. Such
650   overlapping foci are detected as higher apparent stoichiometry foci. The stoichiometry distribution from
651   overlapping foci was modelled by convolving a Poisson distribution generated from the probability of
652   overlap with the expected intensity distribution of an isolated multimer. The latter is obtained by scaling
653   the width of the single fluorophore intensity distribution (Supplementary Fig. 3) by $S^{1/2}$ where S is the
654   model stoichiometry. This model stoichiometry was fixed for those shown in Supplementary Fig. 5. For
655   the Monte Carlo model, the model stoichiometry was generated from a population distribution of
656   oligomeric EGFR whose stoichiometry was sampled from a random Poisson distribution with mean value
657   equal to the mode peak value of 6 that we observed. This prediction resulted in a reasonable fit to the
658   experimental distribution with goodness-of-fit $R^2 = 0.4923$.

659   **Software access.** All our bespoke code written in MATLAB is available from file EGFRanalyser at
660   https://sourceforge.net/projects/york-biophysics/.

661   **Statistical tests and replicates.** Two-tailed Student t-tests were performing for comparisons between
662   pairs of datasets to test the null hypothesis that data in each was sampled from the same statistical

663 distribution. We assume (n_1+n_2-2) degrees of freedom where n_1 and n_2 are the number of independent data
664 points in each distribution and by convention that *t* statistic values which have a probability of confidence
665 P>0.05 are statistically not significant. For TIRF each cell was defined as a biological replicate sampled
666 from the cell population. We chose sample sizes of 10-117 cells per experimental condition which
667 generated reasonable estimates for the stoichiometry distributions. Technical replicates are not possible
668 with irreversible photobleaching, nevertheless, the noise in all light microscopy experiments has been
669 independently characterized for the imaging system used previously.
670

675

676 **References**

677 1.   Lemmon, M. A. & Schlessinger, J. Cell Signaling by Receptor Tyrosine Kinases. *Cell* **141,** 1117–
678   1134 (2010).

679 2.   Roskoski, R. The ErbB/HER family of protein-tyrosine kinases and cancer. *Pharmacol. Res.* **79,** 34–
680   74 (2014).

681 3.   Gertz, J., Siggia, E. D. & Cohen, B. A. Analysis of combinatorial cis-regulation in synthetic and
682   genomic promoters. *Nature* **457,** 215–8 (2009).

683 4.   Jorissen, R. N. *et al.* Epidermal growth factor receptor: mechanisms of activation and signalling.
684   *Exp. Cell Res.* **284,** 31–53 (2003).

685 5.   Lax, I. *et al.* Localization of a major receptor-binding domain for epidermal growth factor by
686   affinity labeling. *Mol. Cell. Biol.* **8,** 1831–1834 (1988).

687 6.   Lax, I. *et al.* Functional analysis of the ligand binding site of EGF-receptor utilizing chimeric
688   chicken/human receptor molecules. *EMBO J.* **8,** 421–427 (1989).

689 7.   Roskoski, R. ErbB/HER protein-tyrosine kinases: Structures and small molecule inhibitors.
690   *Pharmacol. Res.* **87,** 42–59 (2014).

691 8.   Yarden, Y. & Schlessinger, J. Self-phosphorylation of epidermal growth factor receptor: evidence
692   for a model of intermolecular allosteric activation. *Biochemistry* **26,** 1434–1442 (1987).

693 9.   Schneider, M. R. & Wolf, E. The epidermal growth factor receptor ligands at a glance. *J. Cell.*
694   *Physiol.* **218,** 460–466 (2009).

695 10.   Zhang, X., Gureasko, J., Shen, K., Cole, P. A. & Kuriyan, J. An Allosteric Mechanism for
696   Activation of the Kinase Domain of Epidermal Growth Factor Receptor. *Cell* **125,** 1137–1149
697   (2006).

698 11.   Downward, J., Parker, P. & Waterfield, M. D. Autophosphorylation sites on the epidermal growth
699   factor receptor. *Nature* **311,** 483–5

700 12.   Oda, K., Matsuoka, Y., Funahashi, A. & Kitano, H. A comprehensive pathway map of epidermal

growth factor receptor signaling. *Mol. Syst. Biol.* **1,** 2005.0010 (2005).

13. Cohen, S. & Fava, R. A. Internalization of functional epidermal growth factor:receptor/kinase complexes in A-431 cells. *J. Biol. Chem.* **260,** 12351–12358 (1985).

14. Lemmon, M. A. *et al.* Two EGF molecules contribute additively to stabilization of the EGFR dimer. **16,** 281–294 (1997).

15. Odaka, M., Kohda, D., Lax, I., Schlessinger, J. & Inagaki, F. Ligand-binding enhances the affinity of dimerization of the extracellular domain of the epidermal growth factor receptor. *J. Biochem.* **122,** 116–21 (1997).

16. Domagala, T. *et al.* Stoichiometry, kinetic and binding analysis of the interaction between epidermal growth factor (EGF) and the extracellular domain of the EGF receptor. *Growth Factors* **18,** 11–29 (2000).

17. Ogiso, H. *et al.* Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110,** 775–787 (2002).

18. Garrett, T. P. J. *et al.* Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor alpha. *Cell* **110,** 763–773 (2002).

19. Cho, H.-S. & Leahy, D. J. Structure of the extracellular region of HER3 reveals an interdomain tether. *Science* **297,** 1330–3 (2002).

20. Ferguson, K. M. *et al.* EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. *Mol. Cell* **11,** 507–517 (2003).

21. Macdonald, J. L. & Pike, L. J. Heterogeneity in EGF-binding affinities arises from negative cooperativity in an aggregating system. *Proc. Natl. Acad. Sci.* **105,** 112–117 (2008).

22. Macdonald-Obermann, J. L. & Pike, L. J. The Intracellular Juxtamembrane Domain of the Epidermal Growth Factor (EGF) Receptor Is Responsible for the Allosteric Regulation of EGF Binding. *J. Biol. Chem.* **284,** 13570–13576 (2009).

23. Alvarado, D., Klein, D. E. & Lemmon, M. A. Structural basis for negative cooperativity in growth factor binding to an EGF receptor. *Cell* **142,** 568–579 (2010).

24. Liu, P. *et al.* A single ligand is sufficient to activate EGFR dimers. *Proc. Natl. Acad. Sci.* **109,** 10861–10866 (2012).

25. Lemmon, M. A. Ligand-induced ErbB receptor dimerization. *Exp. Cell Res.* **315,** 638–648 (2009).

26. Moriki, T., Maruyama, H. & Maruyama, I. N. Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. *J. Mol. Biol.* **311,** 1011–1026 (2001).

27. Sako, Y., Minoghchi, S. & Yanagida, T. Single-molecule imaging of EGFR signalling on the surface of living cells. *Nat. Cell Biol.* **2,** 168–172 (2000).

28. Martin-Fernandez, M., Clarke, D. T., Tobin, M. J., Jones, S. V & Jones, G. R. Preformed oligomeric epidermal growth factor receptors undergo an ectodomain structure change during signaling. *Biophys. J.* **82,** 2415–2427 (2002).

29. Valley, C. C., Lidke, K. A. & Lidke, D. S. The spatiotemporal organization of ErbB receptors: insights from microscopy. *Cold Spring Harb. Perspect. Biol.* **6,** (2014).

30. Clayton, A. H. A. *et al.* Ligand-induced dimer-tetramer transition during the activation of the cell surface epidermal growth factor receptor-A multidimensional microscopy analysis. *J. Biol. Chem.* **280,** 30392–30399 (2005).

31. Tao, R.-H. & Maruyama, I. N. All EGF(ErbB) receptors have preformed homo- and heterodimeric structures in living cells. *J. Cell Sci.* **121,** 3207–3217 (2008).

32. Liu, P. *et al.* Investigation of the dimerization of proteins from the epidermal growth factor receptor family by single wavelength fluorescence cross-correlation spectroscopy. *Biophys. J.* **93,** 684–698 (2007).

33. Chung, I. *et al.* Spatial control of EGF receptor activation by reversible dimerization on living cells. *Nature* **464,** 783–787 (2010).

34. Nagy, P., Claus, J., Jovin, T. M. & Arndt-Jovin, D. J. Distribution of resting and ligand-bound ErbB1 and ErbB2 receptor tyrosine kinases in living cells using number and brightness analysis. *Proc. Natl. Acad. Sci.* **107,** 16524–16529 (2010).

35. Needham, S. R. *et al.* Measuring EGFR separations on cells with ~10nm resolution via fluorophore localization imaging with photobleaching. *PLoS One* **8,** e62331 (2013).

36. Needham, S. R. *et al.* EGFR oligomerization organizes kinase-active dimers into competent signalling platforms. *Nat. Commun.* **7,** 13307 (2016).

37. Huang, Y. *et al.* Molecular basis for multimerization in the activation of the epidermal growth factor receptor. *Elife* **5,** e14107 (2016).

38. Metzger, H. Transmembrane signaling: the joy of aggregation. *J. Immunol.* **149,** 1477–87 (1992).

39. Chung, C. H. *et al.* Cetuximab-induced anaphylaxis and IgE specific for galactose-alpha-1,3-galactose. *N. Engl. J. Med.* **358,** 1109–17 (2008).

40. Li, S. *et al.* Structural basis for inhibition of the epidermal growth factor receptor by cetuximab. *Cancer Cell* **7,** 301–311 (2005).

41. Cho, H.-S. *et al.* Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Nature* **421,** 756–760 (2003).

42. Wilding, J. L., McGowan, S., Liu, Y. & Bodmer, W. F. Replication error deficient and proficient colorectal cancer gene expression differences caused by 3'UTR polyT sequence deletions. *Proc. Natl. Acad. Sci.* **107,** 21058–21063 (2010).

43. Leake, M. C. *et al.* Stoichiometry and turnover in single, functioning membrane protein complexes. *Nature* **443,** 355–8 (2006).

44. Miller, H., Zhou, Z., Wollman, A. J. M. & Leake, M. C. Superresolution imaging of single DNA molecules using stochastic photoblinking of minor groove and intercalating dyes. *Methods* **88,** 81–8 (2015).

45. Wollman, A. J. M. & Leake, M. C. Millisecond single-molecule localization microscopy combined with convolution analysis and automated image segmentation to determine protein concentrations in complexly structured, functional cells, one cell at a time. *Faraday Discuss.* **184,** 401–24 (2015).

46. Yu, X., Sharma, K. D., Takahashi, T., Iwamoto, R. & Mekada, E. Ligand-independent dimer

777   formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced
778   EGFR signaling. *Mol. Biol. Cell* **13,** 2547–2557 (2002).

779   47.   Llorente-Garcia, I. *et al.* Single-molecule in vivo imaging of bacterial respiratory complexes
780   indicates delocalized oxidative phosphorylation. *Biochim. Biophys. Acta* **1837,** 811–24 (2014).

781   48.   Leake, M. C., Wilson, D., Bullard, B. & Simmons, R. M. The elasticity of single kettin molecules
782   using a two-bead laser-tweezers assay. *FEBS Lett.* **535,** 55–60 (2003).

783   49.   Leake, M. C., Wilson, D., Gautel, M. & Simmons, R. M. The elasticity of single titin molecules
784   using a two-bead optical tweezers assay. *Biophys. J.* **87,** 1112–35 (2004).

785   50.   Sorkin, A. & Duex, J. E. Quantitative analysis of endocytosis and turnover of epidermal growth
786   factor (EGF) and EGF receptor. *Curr. Protoc. cell Biol.* **Chapter 15,** Unit 15.14 (2010).

787   51.   Kirkpatrick, P., Graham, J. & Muhsin, M. Fresh from the pipeline: Cetuximab. *Nat. Rev. Drug
788   Discov.* **3,** 549–550 (2004).

789   52.   Garnock-Jones, K. P., Keating, G. M. & Scott, L. J. Trastuzumab. *Drugs* **70,** 215–239 (2010).

790   53.   Abu-Ali, S., Fotovati, A. & Shirasuna, K. Tyrosine-kinase inhibition results in EGFR clustering at
791   focal adhesions and consequent exocytosis in uPAR down-regulated cells of Head and Neck
792   cancers. *Mol. Cancer* **7,** 47 (2008).

793   54.   Abulrob, A. *et al.* Nanoscale imaging of epidermal growth factor receptor clustering: effects of
794   inhibitors. *J. Biol. Chem.* **285,** 3145–56 (2010).

795   55.   Bag, N., Huang, S. & Wohland, T. Plasma Membrane Organization of Epidermal Growth Factor
796   Receptor in Resting and Ligand-Bound States. *Biophys. J.* **109,** 1925–36 (2015).

797   56.   Wollman, A. J. *et al.* Transcription factor clusters regulate genes in eukaryotic cells. *Elife* **6,** e27451
798   (2017).

799   57.   Hughes, B. D., Pailthorpe, B. A. & White, L. R. The translational and rotational drag on a cylinder
800   moving in a membrane. *J. Fluid Mech.* **110,** 349 (1981).

801   58.   Fushimi, K. & Verkman, A. S. Low viscosity in the aqueous domain of cell cytoplasm measured by
802   picosecond polarization microfluorimetry. *J. Cell Biol.* **112,** 719–25 (1991).

803   59.   López-Duarte, I., Vu, T. T., Izquierdo, M. A., Bull, J. A. & Kuimova, M. K. A molecular rotor for
804   measuring viscosity in plasma membranes of live cells. *Chem. Commun.* **50,** 5282–5284 (2014).

805   60.   Yarden, Y. & Sliwkowski, M. X. Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.*
806   **2,** 127–137 (2001).

807   61.   Zhao, W. *et al.* Mapping the resting and stimulated EGFR in cell membranes with topography and
808   recognition imaging. *Anal. Methods* **6,** 7689–7694 (2014).

809   62.   Ibach, J. *et al.* Single Particle Tracking Reveals that EGFR Signaling Activity Is Amplified in
810   Clathrin-Coated Pits. *PLoS One* **10,** e0143162 (2015).

811   63.   Boggara, M., Athmakuri, K., Srivastava, S., Cole, R. & Kane, R. S. Characterization of the diffusion
812   of epidermal growth factor receptor clusters by single particle tracking. *Biochim. Biophys. Acta -
813   Biomembr.* **1828,** 419–426 (2013).

64. Ariotti, N. *et al.* Epidermal growth factor receptor activation remodels the plasma membrane lipid environment to induce nanocluster formation. *Mol. Cell. Biol.* **30,** 3795–804 (2010).

65. Wang, Y. *et al.* Regulation of EGFR nanocluster formation by ionic protein-lipid interaction. *Cell Res.* **24,** 959–976 (2014).

66. Klapper, L. N. *et al.* The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4995–5000 (1999).

67. Plank, M., Wadhams, G. H. & Leake, M. C. Millisecond timescale slimfield imaging and automated quantification of single fluorescent protein molecules for use in probing complex biological processes. *Integr. Biol. (Camb).* **1,** 602–12 (2009).

68. Reyes-Lamothe, R., Sherratt, D. J. & Leake, M. C. Stoichiometry and architecture of active DNA replication machinery in Escherichia coli. *Science* **328,** 498–501 (2010).

69. Leake, M. C. *et al.* Stoichiometry and turnover in single, functioning membrane protein complexes. **443,** 355–358 (2006).

70. Michalet, X. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Phys. Rev. E* **82,** 41914 (2010).

Supplementary Information
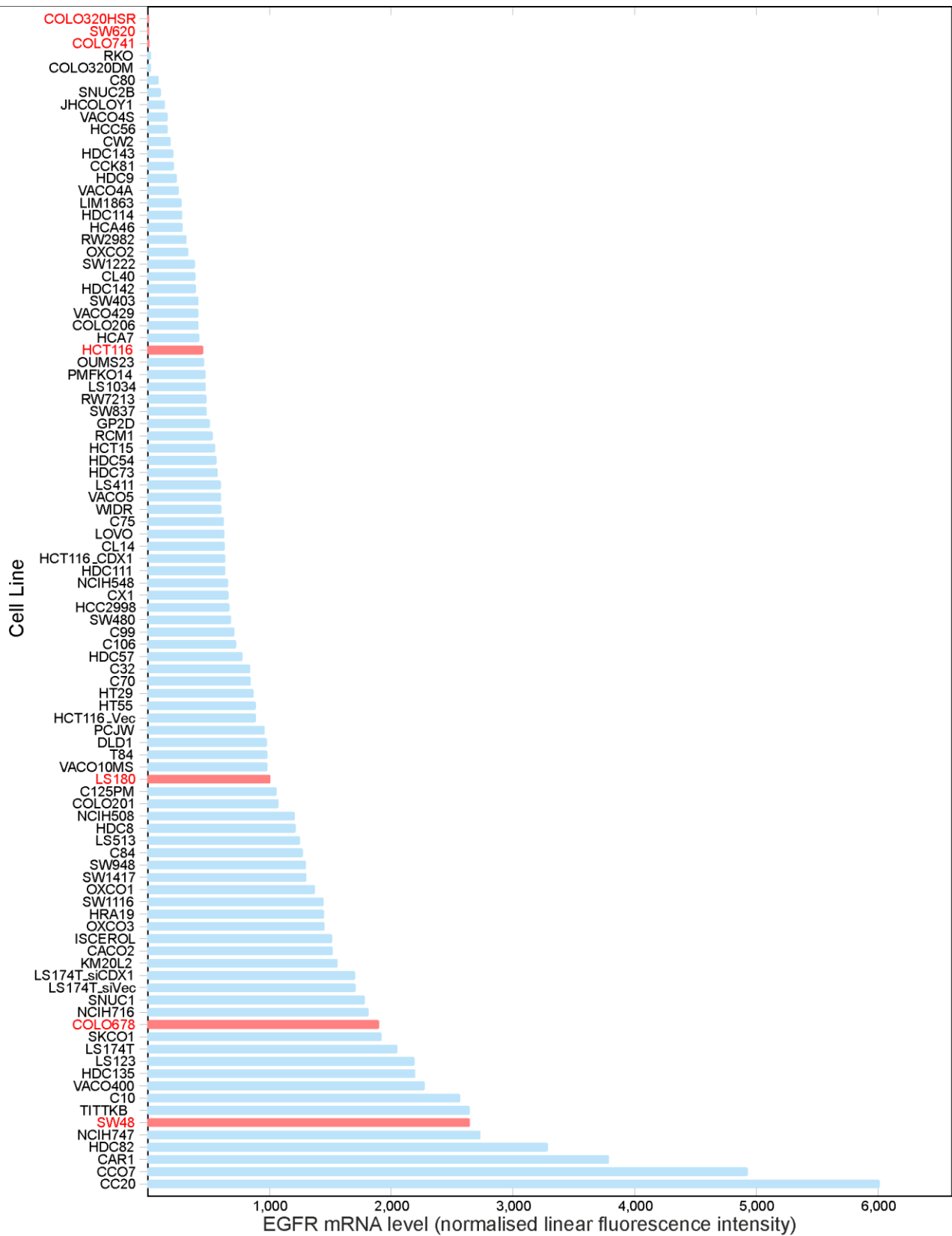

Supplementary Figure 1 EGFR mRNA expression levels. Expression levels were quantified for a colorectal cancer cell line panel using Affymetrix U133+2 mRNA microarray data. Measurements

837   indicated three candidate cell lines, SW620, COLO320HSR and COLO741 (labelled in red, top of

838   panel), as having very low levels of native EGFR expression, as tested in subsequent western blot

839   analysis in comparison to EGFR-expressing cell lines as positive controls (indicated as red columns,

840   middle and bottom of panel). Three candidate cell lines with very low or absent levels of EGFR

841   mRNA (SW620, COLO320HSR, COLO741; Y axis text label in red, top of panel) and a further four

842   positive controls with medium to high levels (HCT116, LS180, COLO678; indicated as red columns,

843   middle and bottom of panel), were selected and protein levels confirmed by western blot (Figure 1b).
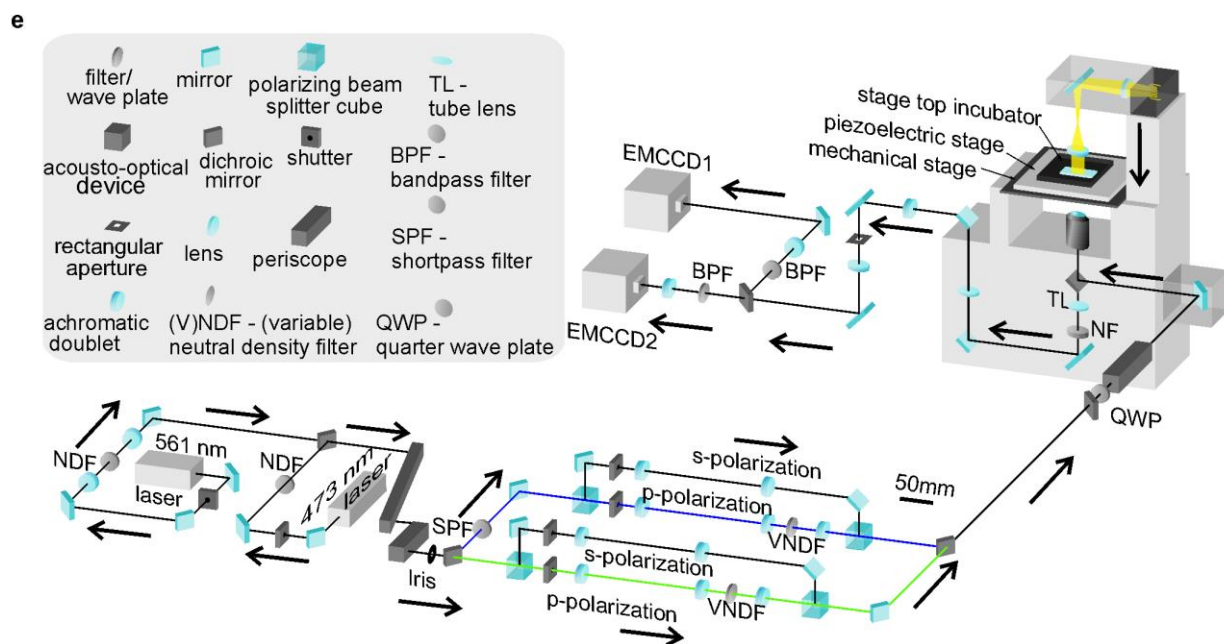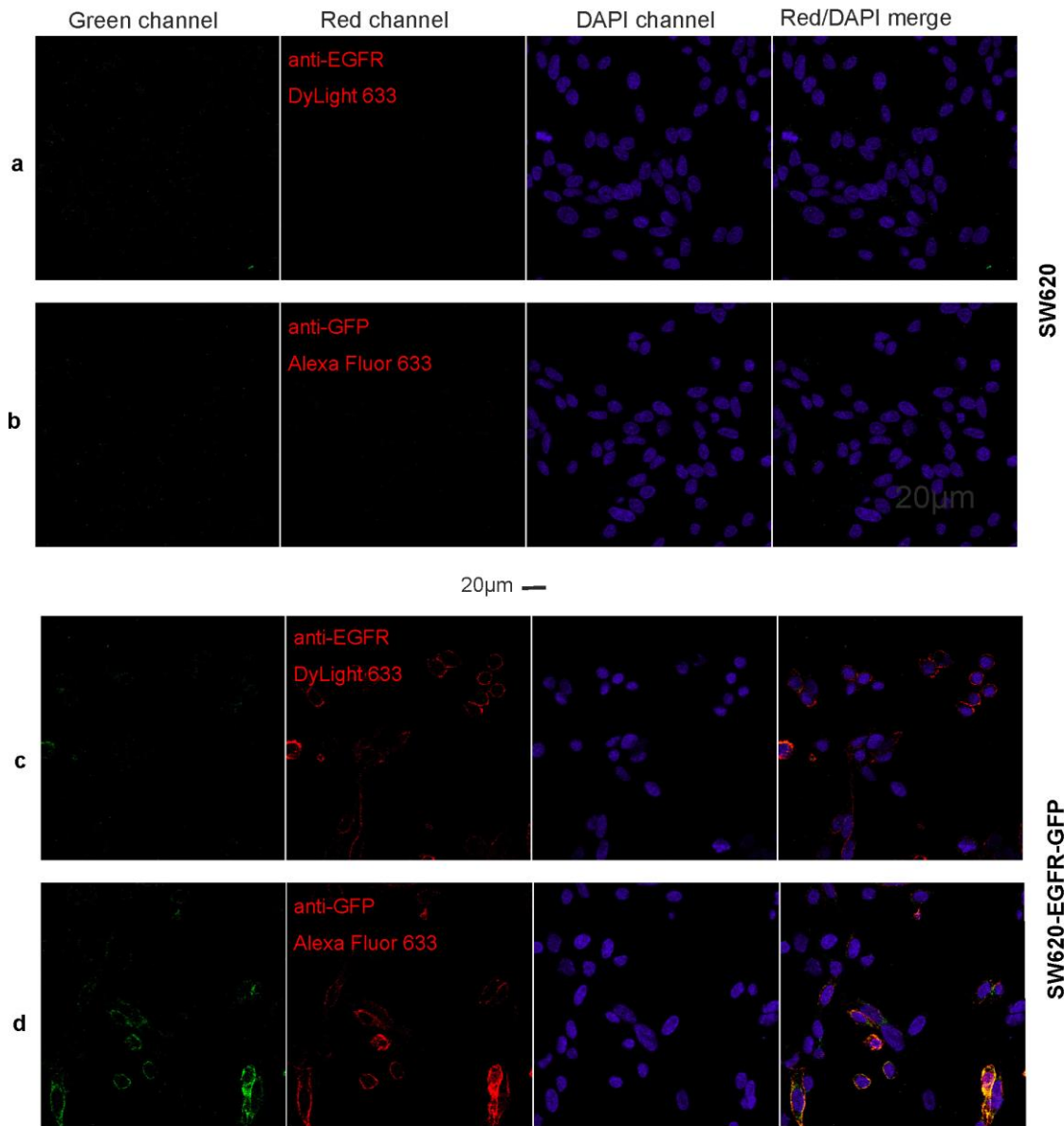
845 **Supplementary Figure 2. Confocal and TIRF characterization.** Confocal microscopy images of fixed
846 cells using GFP, anti-GFP immunofluorescence, and DAPI staining: **(a,b)** non-GFP background cell line
847 SW620; **(c,d)** SW620-EGFR-GFP; **(e)** optical path diagram of bespoke single-molecule TIRF microscope.

848

849
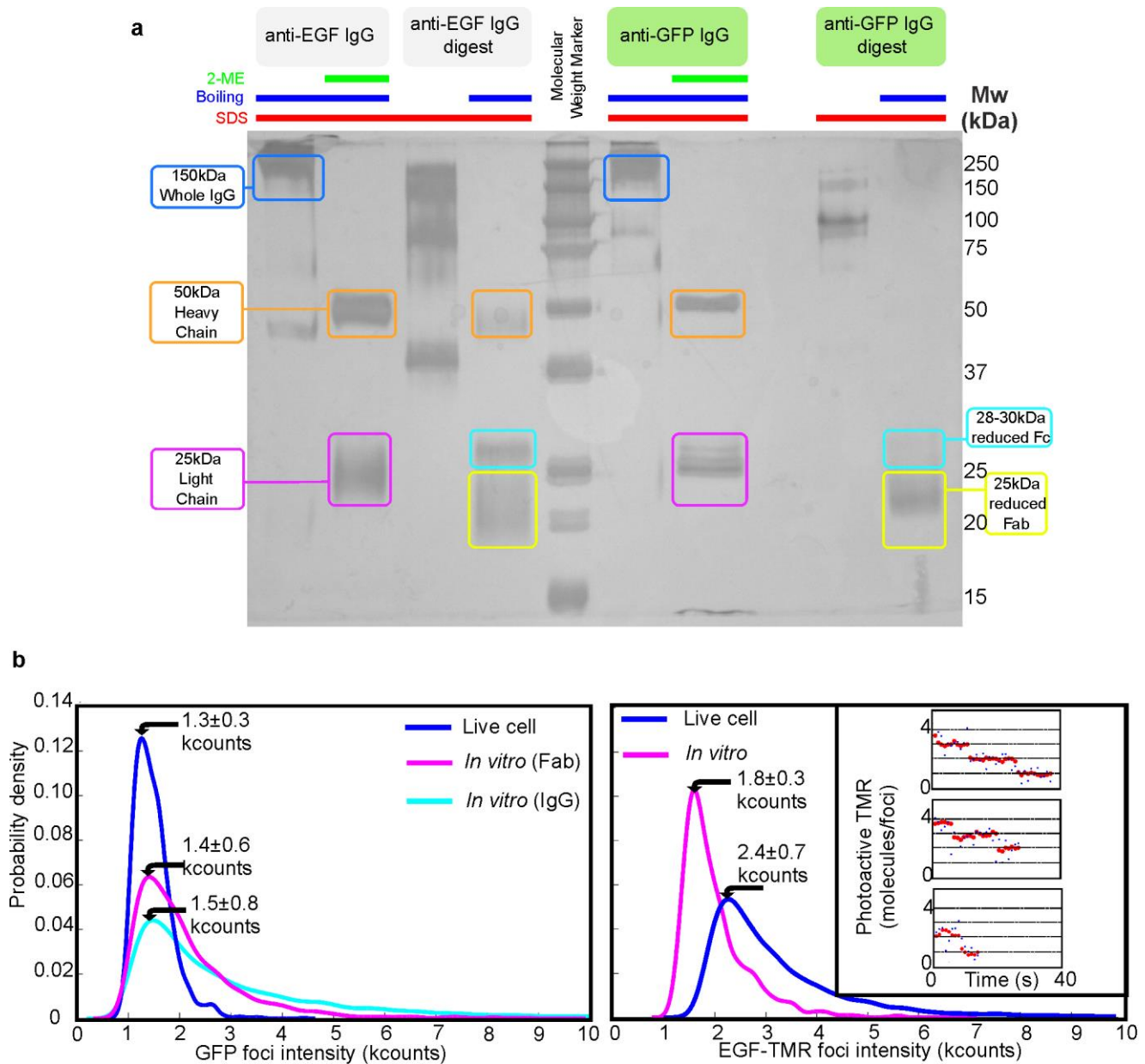
850

851 **Supplementary Figure 3. Characterization of unitary fluorophore brightness values. (a)** SDS-PAGE
852 gel indicating generation of Fab nanobody fragments (yellow) from anti-EGF and anti-GFP IgG antibodies
853 (blue), heaving (orange) and light chains (magenta) indicated with reduced Fc (cyan). **(b)** Kernel density
854 estimation[69] distributions of fluorescent foci intensity values measured in kcounts (i.e. counts x 10^3) for
855 single GFP (left panel) for live cell, at the end of the photobleach, before EGF is added compared with *in*
856 *vitro* Fab and whole IgG data. TMR molecule data for *in vitro* EGF-TMR and live cell, at the end of the
857 photobleach, post EGF binding data taken from colocalized EGF-EGFR foci is shown (right panel); inset

858   shows live cell EGF-TMR photobleach steps after EGF has been added, taken from colocalized EGF-

859   EGFR foci, with raw (blue) and Chung-Kennedy filter[48,49] (red) traces, mean and s.e.m. indicates (arrows).
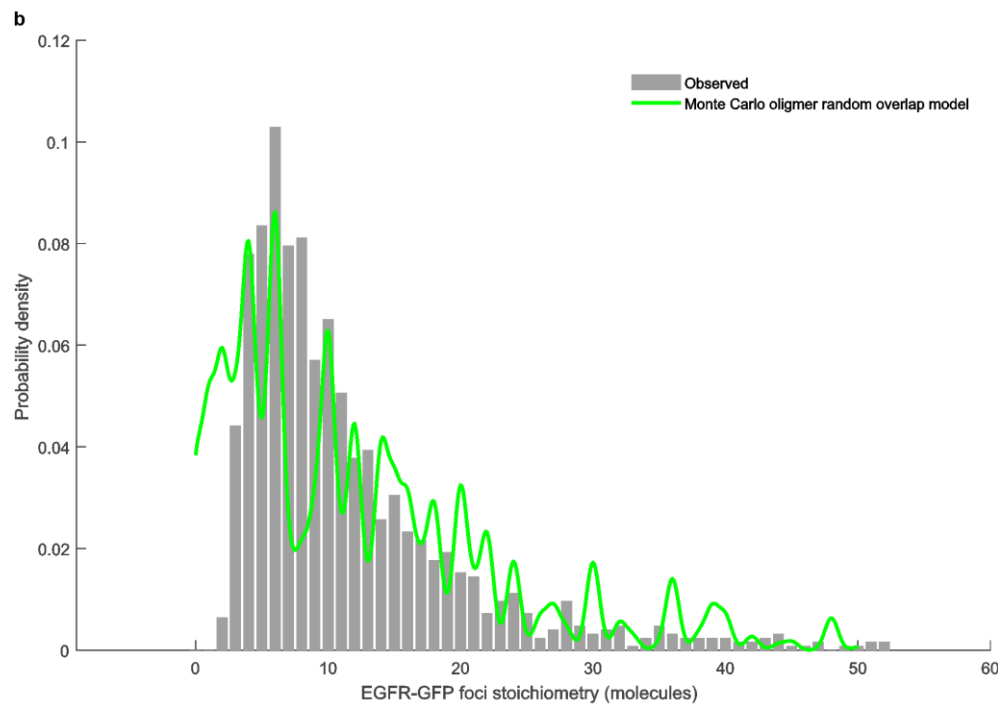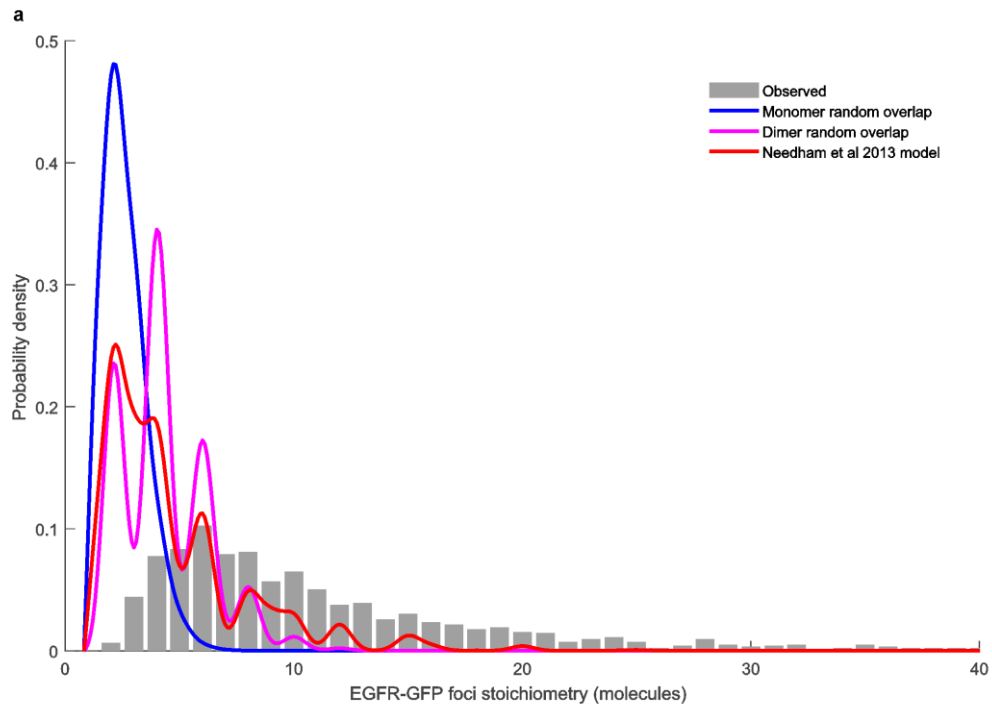
860

BEFORE ligand binding

861
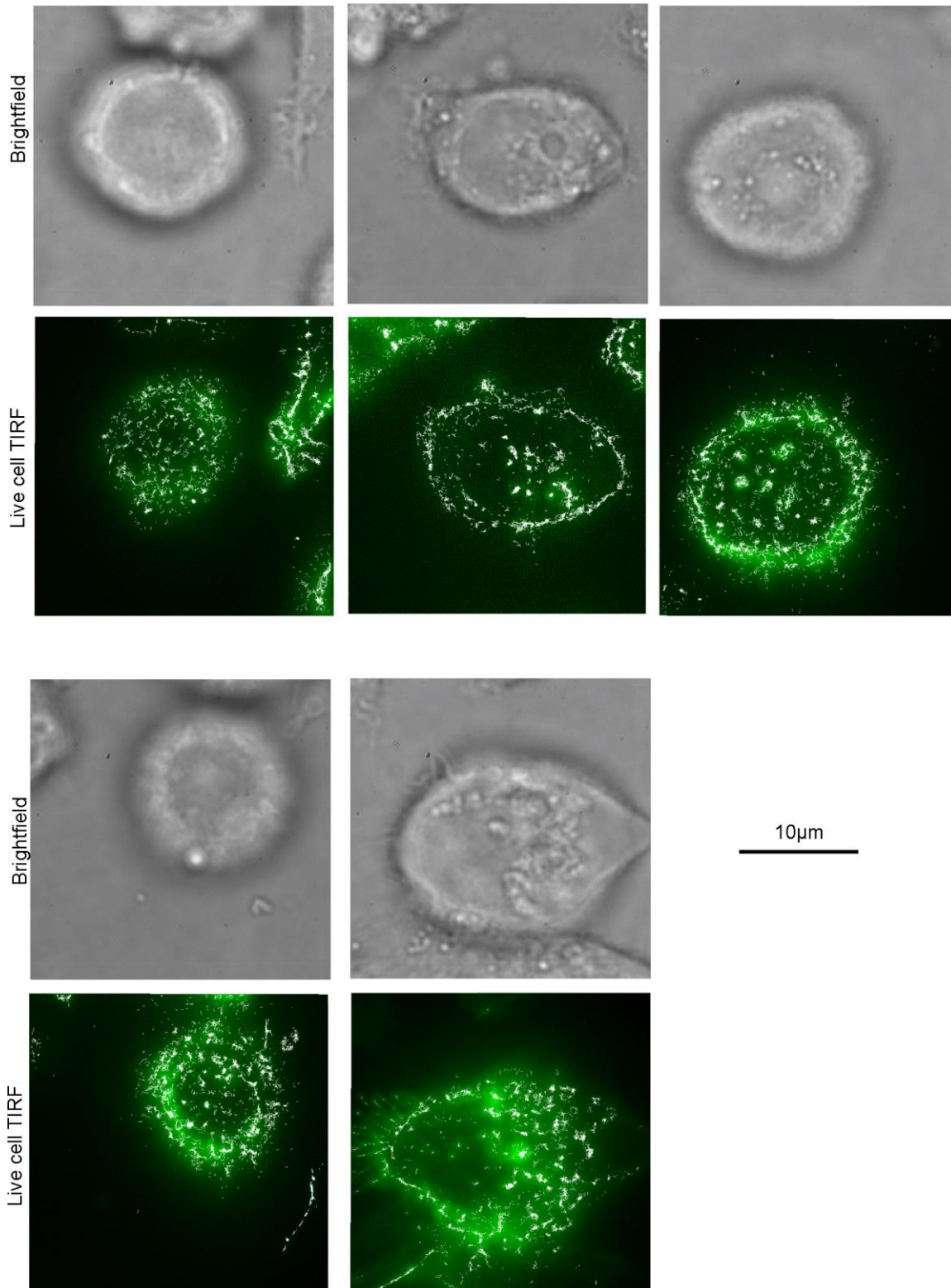
Supplementary Figure 4. More examples of cells before addition of EGF ligand. Brightfield images (grey) and TIRF (green) shown with overlaid foci tracking output (white).

864

865

866
867 **Supplementary Figure 5. Random foci overlap model.** Random overlapping foci predictions for **(a)**
868 monomeric (blue) and dimeric EGFR (magenta), and a mixed model oligomer model suggested from a
869 previous single-molecule study (red)[35], all showing poor agreement ($R^2<0$) to our experimental
870 observations for stoichiometry distribution (grey). **(b)** Monte Carlo Poisson model using an expected
871 average value of 6 molecules for EGFR foci stoichiometry (green) showing reasonable fit ($R^2=0.4923$) to
872 experimental data (grey).
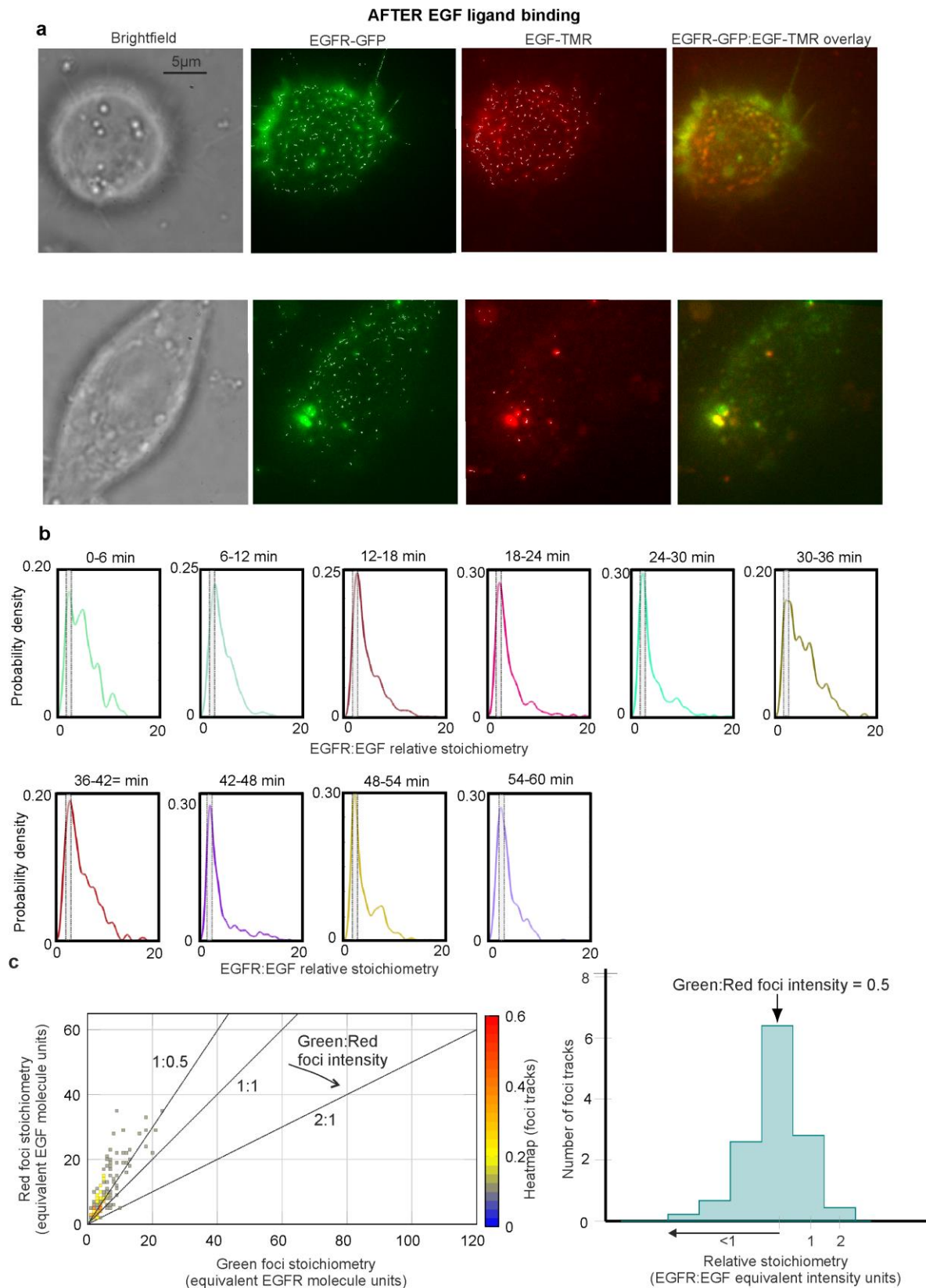
**Supplementary Figure 6. Charactering EGFR and EGF foci stoichiometry after addition of EGF. (a)** Two examples of cells taken ~10 min after the addition of EGF: brightfield images (grey), green channel showing EGFR-GFP localization (green), red channel showing EGF-TMR localization (red), and the overlay of green and red channels together (right panels, with yellow indicating regions of high

878 colocalization) are shown here. **(b)** Variation of the EGFR:EGF relative stoichiometry, rendered as kernel
879 density estimations, as a function of incubation time with EGF (shown in 6 min bins). The region
880 corresponding to 2.0 ± 0.5 relative stoichiometry is indicated as a grey rectangle. **(c)** Heatmap (left panel)
881 and histogram (right panel) characterizing 'false' colocalization due to cellular autofluorescence in green
882 and red channels.

883
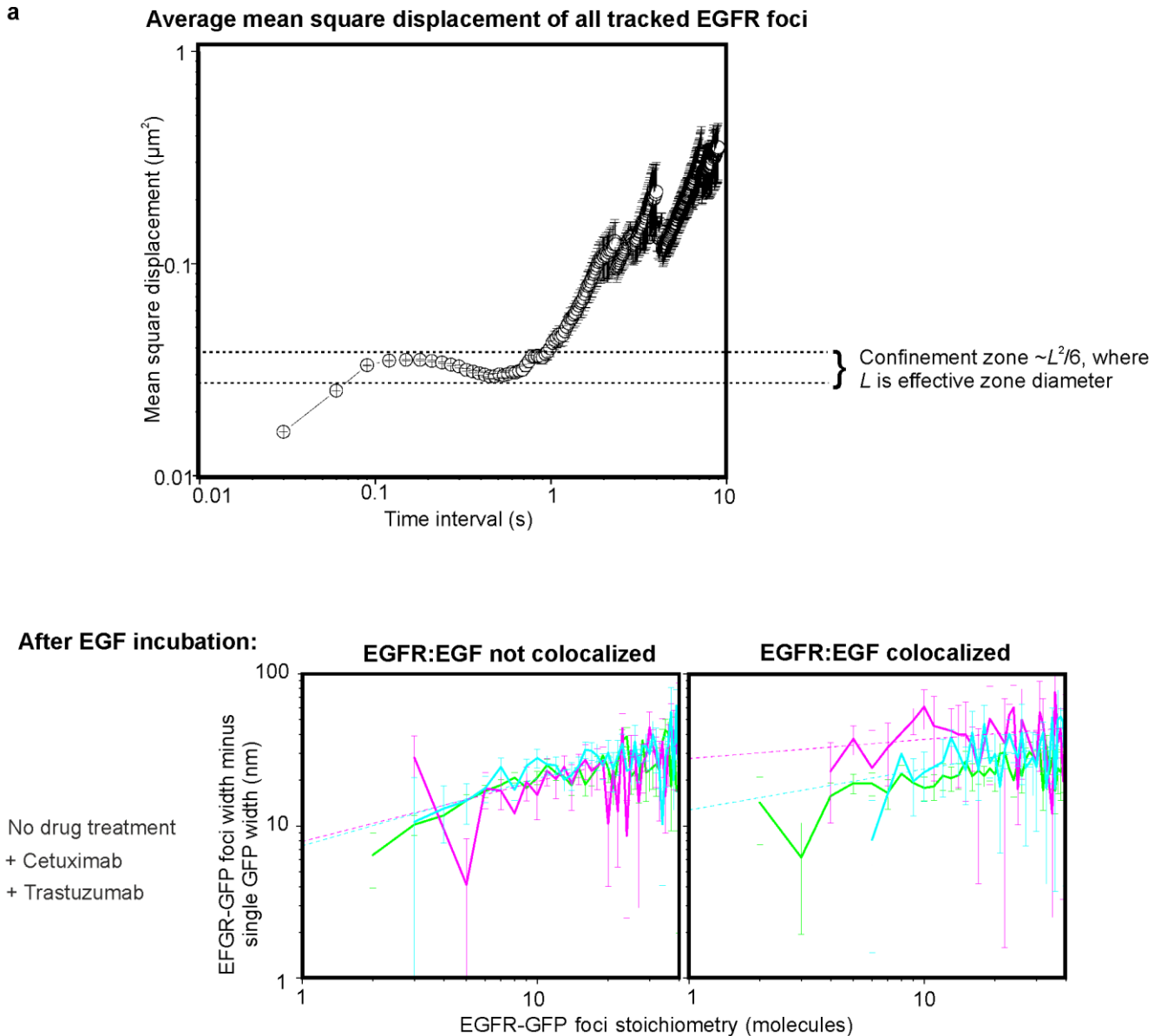
884

885 **Supplementary Figure 7 EGFR foci diffusion.** (a) Log-log plot for average mean square displacement
886 *vs.* time interval for all collated EGFR-GFP foci tracks before addition of EGF, putative confinement zone
887 indicated (dashed lines), from number of foci N=770, acquired from number of cells N=19. (b) Log-log
888 plots of EGFR-GFP foci diameters minus the width for a single GFP molecule *vs.* stoichiometry for not
889 colocalized (left panel) and colocalized foci (right panel), showing cells with no cetuximab or trastuzumab
890 treatment (green, N=6,710 foci, N=117 cells), those treated with cetuximab (magenta, N=1,219 foci, N=25
891 cells), and those treated with trastuzumab (cyan, N=1,607 foci, N=27 cells), with heuristic power law fit
892 (dash lines), s.e.m. error bars.