Electronic Supporting Information (ESI)

Single Particle Tracking-Based Reaction Progress Kinetic Analysis Reveals a Series of Molecular Mechanisms of Cetuximab-Induced EGFR Processes in a Single Living Cell

Do-Hyeon Kim,^{‡a} Dong-Kyun Kim,^{‡b} Kai Zhou,^{‡a} Soyeon Park,^a Yonghoon Kwon,^a

Min Gyu Jeong,^c Nam Ki Lee*bd and Sung Ho Ryu*ab

^a.Department of Life Sciences, Pohang University of Science and Technology, Pohang, 790-784, Republic of Korea.

^b.School of Interdisciplinary Bioscience and Bioengineering, Pohang University of Science and Technology, Pohang, 790-784, Republic of Korea.

^c.Integrative Biosciences and Biotechnology, Pohang University of Science and Technology, Pohang, 790-784, Republic of Korea.

^d.Department of Chemistry, Seoul National University, Seoul, 08826, Republic of Korea.

Corresponding author: <u>sungho@postech.ac.kr</u> for Sung Ho Ryu Corresponding author: <u>namkilee@snu.ac.kr</u> for Nam Ki Lee Antibodies and reagents. Cetuximab was obtained from Merck Serono, and Pitstop 2 was purchased from Abcam (ab120687). The Fab fragment of the anti-EGFR antibody was generated from the monoclonal antibody clone 199.12 (Life Technologies, AHR5072) using a Fab preparation kit (Pierce, 44685). Subsequently, the Fab fragment was conjugated to the Alexa Fluor 647 fluorescent dye using the Alexa Fluor® 647 Antibody Labeling Kit (Life Technologies, A-20186). Protocatechuate-3,4-dioxygenase (PCD) and β -mercaptoethylamine (MEA) were purchased from Sigma Aldrich (P8279 and 30070, respectively). Protocatechuic acid (PCA) was purchased from Santa Cruz Biotech (sc-205818).

Cell culture and transfection. COS7 cells were obtained from the American Type Culture Collection (ATCC) and cultured in Dulbecco's modified Eagle's medium (DMEM, Lonza) supplemented with 10% fetal bovine serum (FBS, Gibco) at 37° C in 5% CO₂ and 95% humidity. The cells were transfected using Lipofectamine 2000 (Invitrogen) according to the manufacturer's instructions.

Sample preparation. The transfected cells were sorted 24 h after transfection to collect only cells exhibiting a low level of protein expression. The coverslips were sonicated in acetone at 42°C for 30 min, followed by sonication in 1% hydrofluoric acid at 42°C for 10 min. The coverslips were then rinsed 15 to 20 times in distilled water to fully remove the hydrofluoric acid. The coverslips were washed three times with PBS before cell seeding. The coverslips were maintained in phenol red-free DMEM (Thermo Scientific).

Microscopy and image acquisition. sptPALM was performed on a home-built cis-total internal reflection fluorescence microscope (TIRFM) based on Olympus IX-71. Cells on glass were maintained at 37° C supplied with 5% CO₂ in a live-cell chamber (Chamlide TC-A, Live

Cell Instrument) mounted on an automated stage (MS-2000, Applied Scientific Instrumentation). A 561-nm laser (for a red form of mEos3.2 excitation, YLK 6150T, Lasos) and a 642-nm laser (for Alexa 647 excitation, VFL-P-1000-642, MPB communications) were collimated and focused with an oil-immersion objective lens (APON 100XOTIRF/1.49, Olympus). Emission light was separated with a dichroic mirror (Z405/561 RPC, Chroma) and an emission filter (BLP01-561R, Semrock). Shutter-controlled 405-nm diode-pumped laser (DL-405-120, Crystal Laser) was separately introduced for fluorophore activation. Images were recorded with an EM-CCD (iXon3 897, Andor Technology).

To track sufficient numbers of single molecules of mEos3.2, cells were repeatedly applied with activation-imaging cycles. Activation time and power of 405-nm activating power varied due to local mEos3.2 density, typically ranging 1-1,000 ms and 0.2-0.5 W/cm² (measured at the back focal plane of objective). Imaging of mEos3.2 was acquired with 561-nm laser with an intensity of 20-30 W/cm² at a frame rate of 19.2 Hz. Typically, 4-10 cycles were applied to achieve enough trajectories. The first 15 frames immediately after the photoactivation were excluded from further analysis because of the high density of mEos3.2 particles. All instrument operations and data acquisition were controlled using MetaMorph (Molecular Devices) and custom plug-ins written in MATLAB (The MathWorks).

Analysis of image data. Multiple particle tracking was performed as previously described.¹

The diffusion coefficient from a single trajectory was calculated from the MSD with the following two-dimensional free diffusion model equation:

 $MSD(\Delta t) = 4D\Delta t + 4e^2$

where $MSD(\Delta t) = E((x_{t+\Delta t} - x_t)^2 + (y_{t+\Delta t} - y_t)^2)$ and (x_t, y_t) are the Cartesian coordinates of a particle at the tth point of its trajectory; D is a diffusion coefficient and e is an error. Trajectories with a duration longer than eight frames were used to calculate the diffusion coefficient using the four time lags of MSD.

Simulation of sptRPKA. Trajectories of Brownian motion were randomly produced by an Euler simulation method. The trajectories were realistically generated with a trajectory length (< 1 s, > 8 frames), signal-to-noise ratio (> 4), localization uncertainty (> 50 nm), and range of diffusion coefficients (0.004 to 0.2 μ m²/s) that were similar to those of trajectories typically observed in SPT experiments with various membrane proteins and in our experiments. For multiple complex states, the mean diffusion coefficient of each state was assigned as the ratios between the two nearest diffusion coefficients of each state are equal.

The density of trajectories in 1 frame was determined as the probability of an overlap between

two diffusional particles is less than 5% using the equation $P(R,\Delta t) = 1 - exp^{[m]}(-\frac{R^2}{4D\Delta t})$, which indicates the probability that a particle with diffusion coefficient *D* has diffused a distance *R* or less in time Δt .² Because Δt is equal to the trajectory duration, the density of trajectories decreased as longer trajectories were utilized.

The exact subpopulation ratio of each state at a certain time was obtained by solving a differential equation for a first-order sequential reaction with a given rate constant. Because of the stochastic nature of photoactivation of mEos3.2, trajectories belonging to each

subpopulation were randomly selected using the exact solutions as probabilities. Thus, the variation in subpopulation ratios was large if the accumulated number of trajectories was low.

Images were constructed frame-by-frame from the simulated trajectories, and were subjected to multiple particle tracking. Thus, particle detection and tracking errors were included to examine the performance of sptRPKA using the simulation. Particles were produced with an average local signal-to-noise ratio of 4 and an average localization uncertainty of 50 nm.

To estimate the mean diffusion coefficient of each subpopulation from the acquired diffusion coefficient distribution appeared as the sum of the distributions of multiple subpopulations, a global solution was obtained by an optimization covering the entire time series of the diffusion coefficient distributions. To reduce the complexity of the search for a global solution, the initial values were simply estimated at the initial and late periods of the reaction for the first and last states, respectively.

Diffusional kinetic analysis of EGFR subpopulations induced by cetuximab for sptRPKA.

During the imaging of a single COS7 cell, the cells were treated with mock (0.1% dimethylsulfoxide, DMSO) or 24 μ M Pitstop 2 (in 0.1% DMSO) for 30 min. Subsequently, the mock-pretreated cells were treated with 20 μ g/ml IgG or cetuximab, and 20 μ g/ml cetuximab was added to the Pitstop 2-pretreated cells. The diffusion coefficient distribution was monitored for sptRPKA at time points -32, -28, -2, 3, 7, 11, 15, 26, 37, 50, 60, 75, 90, 105, and 120 min, where 0 min indicates the time point of treatment with IgG or cetuximab and -30 min indicates the time point of pretreatment. The time points of the measurements were

determined considering the kinetics of the EGFR subpopulation variations. The probability density functions of the converged diffusion coefficient distributions were produced by normalizing each distribution. To generate the temporal maps of the converged EGFR diffusion coefficient distributions, the probability density functions of two variables (diffusion coefficient and time) were constructed using two-dimensional linear interpolation.

To resolve the kinetics of the decrease in the diffusion coefficient of the mobile subpopulation that was induced by cetuximab following Pitstop 2 pretreatment, we acquired images continuously for more than 5 min before and after cetuximab treatment; cetuximab was added in the middle of the imaging. The half-life was obtained using a peak-based analysis. The normalized mean diffusion coefficient was subjected to the exponential decay fit. A peak-based analysis provides a simpler way to compare the overall characteristics of the changes in timeseries distributions than a state-based analysis because the number of states does not need to be determined with the assumption of a Gaussian mixture model prior to the analysis. If the two states involved, both the results from peak- and state-based analyses produced the same halflives, according to the derivation:

$$\hat{D} \sim b * e^{-\lambda t} + c \rightarrow t_1 = \frac{\ln 2}{\lambda}$$

Ď

$$=\frac{E(a_{A}X_{A}+a_{B}X_{B})}{E(X_{A})}=\frac{a_{A}E(X_{A})+(1-a_{A})E(X_{B})}{E(X_{A})}=a_{A}\left(1-\frac{E(X_{B})}{E(X_{A})}\right)+\frac{E(X_{B})}{E(X_{A})}$$

$$a_A \sim \frac{b}{1-c} * e^{-\lambda t} \to t_1 = \frac{\ln 2}{\lambda}$$

where, \hat{D} is a normalized mean diffusion coefficient, a_A is a subpopulation ratio of state A, and X_A is a random variable for a diffusion coefficient of state A.

To examine the subpopulations responsible for the intermediate and trapped-in-a-CCP states was a result of the bivalency of cetuximab, $20 \mu g/ml$ of a Fab fragment of cetuximab was treated on COS7 cells in a similar way mentioned above.

Determination of the number of states in the time-series diffusion coefficient distributions for sptRPKA. Considering that there exist other pre-existing states of EGFR that are not involved in cetuximab-induced EGFR molecular processes (or any processes, in general) while they still appear in the diffusion-coefficient distribution, the determined number of states from the Bayesian information criterion (BIC) analysis does not guarantee the exact number of states involved in the reaction. To solve this problem, we utilized the subtracted density distribution representing the changes only made by the reaction between the two time points. However, the subtraction of two probability density distributions does not yield the probability distribution of which sum should be always one, which does not allow to assess the goodness of distribution fits such as BIC. Thus, we applied an absolute value function to the subtracted distributions and normalized them to be the sum becoming one, which makes them the probability distributions allowing to calculate the BIC. Although the original characteristics of the changes made by the reaction is still well represented in the absolute-valued subtracted probability density, the determined number of states were validated by a curve fit (a coefficient of determination) to

the subtracted distribution.

The sum of the subtracted distribution is zero, which makes even noise signals like the case of the diffusion coefficient distribution at -30 min subtracted by that at 0 min can be determined as a Gaussian mixture. Thus, the effective number of states was determined if the weight of the Gaussian mixture for each state is larger than 3%, which is further validated by fitting the determined number of states to the subtracted distribution. Although BIC was minimum when two states are assumed in the subtracted distributions between -30 min and 0 min, the weight of these two states were less than 1%, indicating that these states should be regarded as noise signals.

Determination of the subpopulation ratios of multiple complex states from singlemolecule diffusion coefficient distributions. The time-series diffusion coefficient distributions were simultaneously fitted using a four-component Gaussian mixture model, i.e., a global fit was applied with independent (weights of Gaussian mixtures) and shared parameters (mean and variance of Gaussian mixtures) to the time-series diffusion-coefficient distribution. For this global fit, the initial value for the mean diffusion coefficient of each state was estimated: the basal free-diffusive EGFR state was determined from the peak of the mobile subpopulation before cetuximab treatment; the cetuximab-bound EGFR state was determined from the peak of the mobile subpopulation 3 min after cetuximab treatment; and CCP-trapped EGFR state was determined from the peak of the slow subpopulation 120 min after cetuximab treatment. The initial value for the mean diffusion coefficient of the unknown intermediate state was set as the median of the diffusion coefficients of the cetuximab-bound EGFR and CCP- trapped EGFR states due to insufficient information about this state. Although the initial values for the standard deviation of each Gaussian component were not given tightly, they were well approximated as long as the mean of each Gaussian component was optimized.

To analyze the cetuximab-induced changes in the subpopulation ratios, we subtracted the diffusion-coefficient distribution of EGFR before cetuximab treatment to those after cetuximab treatment. This procedure is very critical because there exist different EGFR states at the basal status (before cetuximab treatment) and we could not determine which basal states are involved in this reaction. We determined that the amount of state A equals to 1 at the initial time point because the subtracted distribution immediately after cetuximab treatment clearly showed exactly two states A and B (the upper second panel in Figure 3c); the state A existed originally before cetuximab treatment (shown as minus in the subtracted distribution) and the state B newly appeared by cetuximab treatment (shown as plus in the subtracted distribution), indicating that there exist other basal (pre-existing) states of EGFR that are not involved in cetuximab-induced EGFR molecular processes, while they still appear in the diffusioncoefficient distribution. Thus, we subtracted the portions of these reaction-non-involved (cetuximab-insensitive) basal states in the distribution before cetuximab treatment to the distributions after cetuximab treatment, which can be regarded as a concept of baseline correction to analyze only the subpopulations involved in the reaction.

The same results were obtained when we applied a global fit to the subtracted distributions, which provides relative changes in the subpopulation ratios.

Reaction progress kinetic analysis of cetuximab-induced EGFR endocytosis. The rate laws and constants were obtained using a numerical differential method. We built a set of ordinary differential equations (ODEs) to account for the reaction progress of the cetuximab-induced EGFR endocytosis.

The basic equation for a series of first-order consecutive reactions involving four EGFR complex states (Figure 4B) is $(A \rightarrow B \rightarrow C \rightarrow D)$:

$$\frac{dA}{dt} = -k_1[A],$$

$$\frac{dB}{dt} = k_1[A] - k_2[B],$$

$$\frac{dC}{dt} = k_2[B] - k_3[C],$$

$$\frac{dD}{dt} = k_3[C],$$

where A is the basal free-diffusive EGFR state, B is the cetuximab-bound EGFR state, C is the unknown intermediate state, and D is the CCP-trapped EGFR state. To estimate the parameters for these multiple ODEs, we used an iterative prediction-error minimization method by generating a nonlinear grey-box model for these rate equations.³ We evaluated the different plausible models by mapping the simulated integrated laws from a set of the ODEs onto the reaction progress data.

Our proposed model (Figure 4C) is a first-order consecutive reaction with cooperativity in the middle of the reaction $(A \rightarrow B \xrightarrow{n} C \rightarrow D)$:

In sptRPKA, although the state C is composed by *n* molecules of the state B, the probability of

randomly observing mEos3.2 by stochastic photoactivation for the state B and C must be equal because the state C has *n* molecules of mEos3.2 and the state B has one mEos3 (or proportional to the reaction order, in general). This property makes the modifications in the differential equations as follows:

$$\frac{dA}{dt} = -k_1[A],$$
$$\frac{dB}{dt} = k_1[A] - k_2 n[B]^n,$$
$$\frac{dC^*}{dt} = n\frac{dC}{dt} = k_2 n[B]^n - k_3 n[C],$$
$$\frac{dD^*}{dt} = n\frac{dD}{dt} = k_3 n[C],$$

 dX^*

where dt indicates the effective (visualized) amount of a state X and *n* is a cooperativity factor for a high-order rate law.

To evaluate the competitive model (Figure 4D) for the unknown intermediate state and the CCP-trapped EGFR state, branched reaction equations were used $(A \rightarrow B \rightarrow C \text{ and } A \rightarrow B \rightarrow D)$:

$$\frac{dA}{dt} = -k_1[A],$$
$$\frac{dB}{dt} = k_1[A] - k_2 n[B]^n - k_3[B],$$
$$\frac{dC^*}{dt} = n \frac{dC}{dt} = k_2 n[B]^n,$$
$$\frac{dD}{dt} = k_3[B],$$

The equations for reversible reaction models (Figure S14) were used $(A \rightleftharpoons B \rightarrow C \rightarrow D)$:

$$\frac{dA}{dt} = -k_1[A] + k_2[B],$$

$$\frac{dB}{dt} = k_1[A] - k_2[B] - k_3[B],$$

$$\frac{dC}{dt} = k_3[B] - k_4[C],$$

$$\frac{dD}{dt} = k_4[C],$$

The differential equations for other plausible models were established as above.

Cluster analysis of PALM images. COS7 cells expressing a low level of EGFR-mEos3.2 were subjected to a cluster analysis. Cells were treated with mock or cetuximab ($20 \mu g/ml$) for 2 h followed by fixation in methanol at -20 °C for 10 min. The fixed cells were subjected to PALM as previously described.⁴ To assess the cluster patterns of EGFR, Ripley's K-function was applied using the following equation:

$$K(r) = \frac{A}{N^2} \sum_{i} \sum_{j \neq i} I_{ij}$$

where A is the examined area and N is the number of points within the area. R is the radius of clusters to be assessed, and I_{ij} is an indicator function that equals 1 when the distance between *i* and *j* is less than *r*; otherwise, I_{ij} equals 0. Edge correction was achieved by introducing a buffer zone outside the examined area with a width of the maximum examined cluster radius. Ripley's K-function was then linearly transformed:

$$L(r) - r = \sqrt{\frac{K(r)}{\pi}} - r$$

L(r)-r is zero when subjects are distributed randomly, and a positive value indicates that subjects are clustered. The radius of clusters was determined when L(r)-r reached its maximum

value.

To relatively determine the degree of clusters and the ratio of the clustered EGFR (a trappedin-a-CCP state) in trajectory data obtained in a live cell, a k-nearest neighbors (KNN) algorithm was applied. For each localized point, the nearest neighbors and corresponding distances were calculated. The radius of the clusters determined from Ripley's K-function analysis from a fixed cell was used as a threshold for the distance criteria for the KNN analysis. The clustering ratio was determined by the number of EGFR that have over 5 neighbors within a cluster radius over total number of EGFR.

Cetuximab-induced EGFR internalization assay The internalized cetuximab amount in normal and cancer cell lines with different surface EGFR expression levels (MCF7, HeLa, COS7, MDA-MB-231, and A549) was quantified using flow cytometry with Atto 488-labeled cetuximab. Cetuximab was conjugated with Atto 488 with Lightning-Link Rapid 488 kit (Innva Biosciences, 350-0010). Cells were labeled with Atto 488-labeled cetuximab (10 µg/ml) for 10 min on ice to block non-specific internalization. After washing the cells with cold PBS for 3 times, cells were transferred to 37 °C to initiate the cetuximab-induced internalization. After 2 h, cells were rinsed with cold PBS and kept on ice for 5 min to stop internalization and detached with 1 mM of EDTA. Detached cells were fixed with 4% paraformaldehyde. The non-internalized cetuximab bound on cell surface were quenched with membrane non-permeable 0.4% Trypan Blue (ThermoFisher, 15250-061) to specifically measure the internalized amount of cetuximab.⁵ The total internalized amount of cetuximab was quantified by normalizing the mean fluorescent intensity of the sample with the quencher by that of the sample without the

quencher.



Figure S1 Balance between acquisition time and narrowness of diffusion coefficient distribution controlled by trajectory duration. The acquisition time and the full width at half maximum (FWHM) of the logarithm of the diffusion coefficient distribution (the narrowness of the distribution) were inversely proportional (a red solid line) and could be interconverted by controlling the threshold of the trajectory duration (dots). The insets show the diffusion-coefficient distributions generated by the accumulated trajectories with the specified trajectory duration and acquisition time for an equal amount of three substrates of which diffusion coefficients are distinct (a red dashed line). The degree of the inverse relationship may be dependent on other imaging parameters, such as the brightness and photostability of the fluorescent molecule, the readout speed and sensitivity of the detector, and the particle density for SPT.



Figure S2 Performance of sptRPKA to estimate the rate constants for a series of first-order reactions involving five complex states. The reaction progress data of sptRPKA (dots) for a sequential reaction (line) with rate constants of 0.02 s^{-1} for all five states and diffusion coefficients of 0.2, 0.075, 0.028, 0.011, and 0.004 μ m²/s are shown. Trajectories with a duration of 15 frames were used.



Figure S3 (a) More than 200,000 particles of EGFR-mEos3.2 (left) were tracked in a living COS7 cell. The 7 trajectories were overlaid on one frame of a raw image (middle). 30,000 EGFR trajectories were shown in the trajectory map (right). Scale bars, 2 μ m (white) and 5 μ m (black). (b) Differential interference contrast (DIC) images were obtained before (indicated as 0 min) and after imaging for sptPALM in the same cell at 5, 10, 15, 30, and 60 min. Scale bar, 5 μ m. (c) There was no sign of retraction of the edges of the plasma membrane, indicating that negligible photo damage occurred to the cell.



Figure S4 Determination of EGFR diffusion model and time lags for diffusion coefficient calculation. (a) The mean squared displacement (MSD) curves of EGFR trajectories longer than eight frames were analyzed. Two hundred randomly chosen trajectories between the 25th and 75th percentiles of MSD at the time lag of five (gray) and a curve for the mean of the MSDs of all the trajectories are shown (red). Approximately 86% of the trajectories displayed free diffusion (MSD=4Dt^{α}, 0.7< α <1.3). We used a free diffusion model to account for the EGFR trajectories in a single COS7 cell. (b) Next, we determined the optimal number of steps in the random walks to calculate the diffusion coefficient from the MSD curves of each EGFR trajectory. The mean diffusion coefficients at each step size of random walks were determined by averaging the diffusion coefficients calculated from the trajectories longer than eight frames by applying a free diffusion model with (red) or without (black) error incorporation $(MSD=4Dt+4e^2)$. The error bars represent the standard error of the mean at the single-molecule level (n>10,000). The number of steps did not significantly affect the mean diffusion coefficient because of the linearity of the diffusion model, except for one step. (c) The standard deviation of the diffusion coefficient distribution of EGFR using different step sizes for the random walks with (red) or without (black) error incorporation decreased as the number of steps in the random walks decreased, even below the time lag of four that is typically used for SPT analysis using

fluorescence. Thus, we used the four time lags of MSD ($0 < \Delta \tau < 160 \text{ ms}$) with error incorporation to utilize the EGFR trajectories with durations longer than eight frames.



Figure S5 (a) A histogram showing the diffusion coefficient distribution of EGFR-mEos3.2 trajectories in (Figure S3a). A two-component Gaussian mixture model was used to separate the EGFR subpopulations (blue: mobile subpopulation; red: immobile subpopulation) for quantitative subpopulation analysis. (b and c) The profile of the mean diffusion coefficient of the EGFR subpopulations at the single-cell level with respect to the acquiring time for the mobile (b) and immobile (c) subpopulations. The error bars represent the 95% confidence interval. The mean diffusion coefficients of the mobile and immobile subpopulations converge from ~ 7.3 s and 8.4 s, respectively. (d-f) Single-cell diffusion coefficient Histograms at the

specified times are presented with two reference lines (dashed lines) indicating the mean diffusion coefficients of each subpopulation at 0 min (d). The mean diffusion coefficients (e) and relative ratios (f) of the subpopulations are shown. The error bars (e) represent the standard deviation of each subpopulation.



Figure S6 Rapid determination of diffusion-coefficient distributions of EGFR in various cell lines. COS7 (a),HEK293 (b), HeLa (c) and CHO-K1 (d) cells were transfected with EGFR-mEos3.2, followed by sptPALM to measure the diffusion coefficient distributions of EGFR in these three cell lines. A histogram showing the diffusion coefficient distribution of EGFR-mEos3.2 in each cell is displayed. The insets provide the 95% confidence interval of the mean diffusion coefficient with respect to the accumulated number of trajectories for determining the diffusion-coefficient distributions. Below the 1% confidence interval was achieved by trajectories accumulated within 10 s in all three cell lines.



Figure S7 Variation in the number of EGFR complex states induced by cetuximab over the reaction course. The number of states was estimated by the Bayesian information criterion (BIC) with respect to the time of cetuximab addition with (red) or without (black) Pitstop2. The number of states that exhibited the lowest BIC was chosen for each time point.



Figure S8 Singular value decomposition analysis of the time-series diffusion coefficient distribution of EGFR treated with cetuximab subtracted by the diffusion coefficient distribution at 0 min. (a) the subtracted time-series diffusion coefficient density map. (b) Singular values sorted in a descending order. (c-f) rank k-approximation of the subtracted map for rank 1 (c), 2 (d), 3 (e), and 4 (f).



Figure S9 Anomalous diffusion analysis of cetuximab-induced EGFR subpopulations. (a) Two-dimensional distribution of EGFR for a diffusion coefficient (D) and an anomalous exponent (α) was obtained by analyzing EGFR trajectories with a step size of up to seven ($\Delta t=7$) after cetuximab treatment for 60 min using MSD(Δt)=4D_a Δt^{a} +4e². (b) The effect of step sizes on the mean anomalous exponent for state D (the trapped-in-a-CCP state, D_a < ~10^{-2.8} µm²/s corresponding to D < ~10^{-1.9} µm²/s calculated from a free diffusion equation) was examined. The anomalous exponents were consistent throughout the step size from five to ten. The mean anomalous exponent for state D is 0.299, indicating the state is sub-diffusive.



Figure S10 Quantitative measurement of cetuximab binding to EGFR with Pitstop 2 pretreatment through a dose-dependent diffusion-coefficient shift. The dissociation constant and cooperativity of cetuximab binding to EGFR on a COS7 cell membrane were analyzed from the dose-dependent diffusion-coefficient shift by cetuximab. The cells were pretreated with Pitstop 2 for 30 min prior to adding cetuximab to observe the effects of cetuximab binding alone. The inset shows a Scatchard plot for evaluating binding cooperativity. The error bar represents the standard error of the mean of a single-molecule population (n > 5,000).



Figure S11 Subpopulation-specific kinetics analysis of EGFR diffusion coefficient distributions altered by cetuximab with or without Pitstop 2. Histograms of the diffusion coefficient of EGFR before and 2 h after cetuximab treatment with (a) or without (b) Pitstop 2 pretreatment are presented for the subpopulation-specific comparisons. The kinetics of the decreased mean diffusion coefficient of the fast-diffusive EGFR subpopulation caused by cetuximab are shown in the insets. The half-lives of each subpopulation's kinetics ($t_{1/2}$) were obtained by fitting the kinetics using exponential decay.



Figure S12 Cetuximab-induced reaction progress of four EGFR complex states. The normalized subpopulation ratios (dots) from the diffusion coefficient distribution of EGFR over 2 h after cetuximab treatment are shown (*left panel*). The early reaction course is displayed (*right panel*) to demonstrate the sub-minute resolution of sptRPKA.



Figure S13 A total expression level of eGFP-EGFR before and after cetuximab treatment for 2 h. (a-c) TIRF images for EGFR conjugated with eGFP before cetxuimab (20 μ g/ml) treatment (a) and after cetuximab treatment for 60 min (b) and 120 min (c). Scale bars, 5 μ m. The average fluorescence intensity for corresponding images were quantified in (d). (e) A time-dependent change in the total level of EGFR by cetuximab was examined across multiple cells (n > 10). EGF treatment (10 nM) serves as a positive control for EGFR internalization.



Figure S14 Examination of reversible reaction models for cetuximab-induced EGFR endocytosis. A reversible reaction at each step of the first-order sequential reaction model was examined to determine which model best accounted for the kinetics of the four EGFR subpopulations induced by cetuximab. Reversible reaction models could not explain the reaction progress data (a, c, d), except for (b). However, considering that EGFR should be clustered (Figure 5) after the reversible reaction step in (b), the model in which its reverse rate is higher than the forward rate is implausible. The RSS of each state for all the reversible reaction models was greater than that for the sequential reaction with cooperativity model shown in Figure 4C.



Figure S15 Residual analysis of the sequential and parallel reaction models for the CCPtrapped state of EGFR. The pattern of the residues of the CCP-trapped state resulting from optimal regression by the sequential (a, corresponding to Figure 4C) or parallel (b, corresponding to Figure 4D) reaction with cooperativity model to the kinetics of the four EGFR complex states induced by cetuximab was analyzed.



Figure S16 EGFR-level dependent diffusion dynamics of EGFR subpopulations induced by cetuximab. Time-series diffusional probability density maps of EGFR with cetuximab in low (~53 molecules/ μ m²) (a), middle (~294 molecules/ μ m²) (b), and high (~1685 molecules/ μ m²) (c) EGFR-expressing cells. (d, e) The subpopulation ratios of states C (a previously unknown intermediate state, d) and D (a trapped-in-a-CCP state, e) were determined from individual COS7 cells variously expressing EGFR at 60 min after cetuximab treatment (left panels). The data are fitted using a Hill's equation (black solid line) displayed with 95% confidence interval

(black dashed lines). Rate constants for k2 and k3 were determined using sptRPKA from the individual COS7 cells (right panels) using the sequential reaction model with cooperativity in Figure 4c. The cooperativities (n) and rate constants (k) of both states C and D are consistent with the result in Figure 6a.



Figure S17 Temporal diffusion dynamics of EGFR subpopulations induced by a Fab fragment of cetuximab. (a) a time-series diffusional probability density map of EGFR with a Fab of cetuximab. (b) diffusion coefficient dstributions of EGFR before and after a Fab of cetuximab treatment for 60 min. Slow diffusive (the intermediate unknown state) or immobilized EGFR (trapped-in-a-CCP state) was not induced. The peak shift (~10% of mean diffusion coefficient) of free diffusive EGFR was due to the direct binding of a Fab of cetuximab, which was previously reported.¹



Figure S18 sptRPKA for endogenous EGFR in a COS7 cell. (a) Endogenous EGFR was labeled with a Fab fragment of the anti-EGFR antibody conjugated to Alexa Fluor 647 (A647). The single-molecule fluorescent image of A647-EGFR in one frame and 10,000 randomly chosen trajectories are shown. Scale bar, 5 μm. (b) The profile of the mean diffusion coefficient of the two EGFR subpopulations with respect to the number of accumulated trajectories. The error bars represent the 95% confidence interval. The mean diffusion coefficients of the first and second subpopulations converge beginning at approximately 10.05 s and 5.7 s, respectively. (c-e) The same experimental scheme used in Figure 3 was applied for cetuximab-induced endogenous EGFR instead of heterologously expressed EGFR-mEos3.2. The characteristic dynamics of the endogenous EGFR subpopulations after treatment with IgG (c) and cetuximab without (d) or with (e) Pitstop 2 are presented as histograms (white arrows). The half-lives of

each of the subpopulation-specific kinetics are shown. The three distinct characteristic kinetics and the half-lives exhibited by cetuximab-induced endogenous EGFR endocytosis were similar to those of EGFR-mEos3.2.

References

- D. H. Kim, K. Zhou, D. K. Kim, S. Park, J. Noh, Y. Kwon, D. Kim, N. W. Song, J. B. Lee,
 P. G. Suh, N. K. Lee and S. H. Ryu, *Angew. Chem. Int. Ed. Engl.*, 2015, 54, 7028-7032.
- 2. M. J. Saxton, *Biophys. J.*, 1997, **72**, 1744-1753.
- R. Romijn, L. Ozkan, S. Weiland, J. Ludlage and W. Marquardt, *J. Process. Contr.*, 2008, 18, 906-914.
- E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M.
 W. Davidson, J. Lippincott-Schwartz and H. F. Hess, *Science*, 2006, **313**, 1642-1645.
- 5. T. Patino, J. Soriano, L. Barrios, E. Ibanez and C. Nogues, *Sci. Rep.*, 2015, **5**, 11371.