Supplemental Text

Reconstitution of an EGFR construct containing the transmembrane and intracellular modules into lipid bilayers results in membrane stacking

We visualized membrane preparations of the TM-ICM construct by using negative-stain electron microscopy. Normal preparations of small unilamellar vesicles, with no protein, have the appearance of flattened disks in the electron microscope, with an average cross-sectional diameter corresponding to the expected value of 100 nm (Sup. Fig. 8A). Attachment of the EGFR kinase domain to small unilamellar vesicles (SUVs) by a hexahistidine tag, as done in our original analysis of asymmetric dimer formation (Zhang, Gureasko et al. 2006), leads to vesicles that also look normal in the electron micrographs (Fig. 11A). The electron micrographs are consistent with the results of light scattering analysis, which show particles with a diameter of 100 nm (Fig. 11D).

Quite different results are obtained with the TM-ICM construct, for which negative-stain EM reveals highly elongated caterpillar-like structures with a uniform banding pattern of light and dark staining (Sup. Fig. 8A). The band-to-band distance within each "caterpillar" is ~7 nm and the widths of the caterpillar-like structures is a nearly uniform ~40 nm.

Elongated and caterpillar-like structures for vesicle preparations have been described before (Waggoner, Last et al. 2001, Vedhachalam, Duong et al. 2007). Such structures are seen in negative-stain electron micrographs of vesicles containing components that have membrane-embedded parts that are attached to components that can self-associate. For example, membrane preparations that contain lipids with metal-chelating head groups have been found to form caterpillar-like structures that bear a

striking resemblance to the structures we see for the EGFR TM-ICM construct (Waggoner, Last et al. 2001). The formation of the caterpillar-like structures has been ascribed to the ability of the extra-membrane self-associating components to bring membranes together, and to flatten the apposed membrane surfaces so as to the increase the number of interactions between these groups. In addition, the dehydrating conditions used for the negative stain preparations has been found to lead to the caterpillar-like appearance of the stacked membranes (Zhang, Song et al. 2011).

We also analyzed the TM-ICM membrane preparations by electron cryomicroscopy, which avoids the harsh conditions necessary for the negative stain procedure. These show onion-like multilamellar vesicles when the TM-ICM is present, in which multiple membranes are stacked together (Sup. Fig. 8A; Sup. Fig. 9A). In contrast, empty vesicles show a single membrane in the cryomicroscopic images (Sup. Fig. 8A; Sup. Fig. 9B).

Based on our observations, as well as the previous reports of the caterpillar-like structures, we surmise that the stacked membrane structures in our preparations of the TM-ICM construct arise from the interactions between kinase domains when they form the asymmetric dimer (Sup. Fig. 8C). We confirmed that this is the case by preparing the V924R variant of the TM-ICM construct and reconstituting it in lipid membranes. This construct yields normal SUVs, as visualized by negative-stain EM (Sup. Fig. 8A). This difference is confirmed by light-scattering measurements, which show large-scale aggregation of vesicle preparations of the TM-ICM construct, but monodisperse -100 nm particles for the TM-ICM variant with the V924R mutation (data not shown).

Reconstructing asymmetric dimers using V924R/I682Q combinations of the TM-ICM construct also leads to vesicle stacking, indicating that the stacks arise as a consequence of individual asymmetric dimers, rather than chains of them. There is a

rotation of ~120° between kinases in an asymmetric dimer, which might favor the formation of asymmetric dimers in which each subunit is presented by apposed membranes rather than the same one. As the measured dimensions of the stacks (5-7 nm thickness) can accommodate EGFR kinase asymmetric dimers, and the V924R mutation disrupts stacking, we assume that it is the asymmetric dimer interaction that acts as a molecular "glue" between the vesicles. Such an extensive membrane remodeling brought forth by the formation of the asymmetric dimer points out the strength of this interaction.

Supplemental Table 1. Data collection and refinement statistics.

	V924R	I682Q
Wavelength	0.9998	0.9999
Resolution range	38.56 - 1.55 (1.605 - 1.55)	42.51 - 1.9 (1.968 - 1.9)
Space group	C 2	C 2
Unit cell	155.4 71.8 76.4 90 113.3 90	155.9 72.0 76.8 90 113.3 90
Total reflections	214396 (17502)	122948 (12276)
Unique reflections	110980 (10214)	61664 (6160)
Multiplicity	1.9 (1.7)	2.0 (2.0)
Completeness (%)	0.99 (0.92)	1.00 (1.00)
Mean I/sigma(I)	7.58 (1.32)	8.62 (2.00)
Wilson B-factor	19.27	22.85
R-merge	0.031 (0.501)	0.049 (0.558)
CC1/2	0.997 (0.605)	0.998 (0.54)
Reflections used in refinement (R-free)	110936 (5572)	61611 (3121)
R-work / R-free	0.175 / 0.203	0.182 / 0.211
Number of non-hydrogen atoms	5579	5200
RMS(bonds / angles)	0.017 / 1.68	0.007 / 1.02
Ramachandran favored (%)	97	98
Ramachandran outliers (%)	0	0.17
Rotamer outliers (%)	1.1	0.37
Clashscore	5.13	2.42
Average B-factor	29.35	30.21
macromolecules	28.39	30.07
ligands	20.28	25.07
solvent	39.29	34.00

Statistics for the highest-resolution shell are shown in parentheses.

Supplemental Figure Legends

Supplemental Figure 1 – EGFR sequence alignment

- **A.** Alignment of EGFR sequences from jawed vertebrate species (*Gnathostomes*). The EGFR tail sequence is highly conserved among these jawed vertebrate species, including seven strictly conserved tyrosine residues. All alignments were performed using ClustalW.

- **B.** Alignment of EGFR sequences from vertebrate and invertebrate species, highlighting the breakdown of conservation of the EGFR tail sequence from vertebrate to invertebrate species.

Supplemental Figure 2 – Analysis of EGFR phosphorylated tyrosine antibody specificity

Human EGFR variants expressed in Cos7 cells were used to test antibody specificity towards phosphorylated tyrosines in the activation loop (Tyr 845) and tail segment by Western blot. See methods for antibody details. Data from two replicate experiments are shown.

Supplemental Figure 3 – Full flow cytometry data for the EGFR tail deletions presented in Fig. 3B

- **A.** EGFR tail phosphorylation (Tyr 1173) was measured as a function of EGFR expression level in cell stimulated with EGF by flow cytometry using a phosphorylated-Tyr 1173 antibody. Data is presented for all eight tail deletion constructs presented in Fig. 5.

- **B.** EGFR activation loop phosphorylation (Tyr 845) was measured as a function of EGFR expression level in cell stimulated with EGF by flow cytometry using a phosphorylated-Tyr 845 antibody. Data is presented for all eight tail deletion constructs presented in Fig. 3.

Supplemental Figure 4 – Data analysis of the flow cytometry essay measuring pErk accumulation as a function of EGFR expression and activation

- **A.** Dot plot showing the raw data of the pErk assay used to compare the distribution of pErk level as a function of EGFR expression and activation. Each dot represents one single cell. In the X-axis, the FITC intensity indicates the EGFR expression level. pErk level is measured by the intensity of APC

intensity and plotted in the Y-axis. Data is show for wild-type EGFR without (left) or with (right) EGF stimulation.

B. Untransfected cells present low FITC intensity (colored grey) and are excluded, defining the lower limit of the cell selection gate.

C. Cells with high level of EGFR expression that present a clear increase in pErk level in the absence of EGF stimulation (blue) are also excluded. This defines the upper limit of the cell selection gate.

D. Only the cells expressing moderate level of EGFR (colored red), were pooled to make the histogram showing the pErk level histograms.

Supplemental Figure 5 – Crystal structures of EGFR V924R and I682Q mutants

A. Asymmetric unit of the EGFR V924R crystal structure. Strong electron density from a proteolytic fragment was found sandwiched between the N-lobes of the two EGFR V924R kinase domains in the crystal structure.

B. Electron density and model of the proteolytic peptide fragment.

C. Superposition of the original (white; PDB code 3GTO) and new (light blue) structures of the EGFR V924R mutant.

D. Detailed view of the sequence register correction in the EGFR V924R tail structure.

E. Orthogonal views of a superposition between monomer A (blue) of the EGFR I682Q structure here presented with the structure of EGFR bound to Mig6 (white; Mig6 colored green). A crystal contact between monomers in the EGFR I682Q structure is incompatible with tail binding in *cis* to the C-lobe of the kinase domain. In this superposition Mig6 clashes with the C-lobe of the other kinase domain in the inactive dimer.

Supplemental Figure 6 – Molecular dynamics simulations

A. α-Helix formation over time**.** The Amber module *cpptraj* was used to analyze secondary structure formation over time during replicate simulations started from a model ~10 ps after the initial formation of the FF helix. For replicate simulations, solvent atoms were left in place, but all atoms were re-assigned random velocities, and the system was equilibrated with harmonic constraints at constant temperature and pressure before unconstrained molecular dynamics. α-helix formation is plotted as black bars for each 1 ps frame in which the listed residue participates in an α-helix according to the DSSP method of Kabsch and Sander (Kabsch and Sander 1983).

B. β-sheet Hydrogen Bond Distances for the Mig6-like tail segment. The Amber module *cpptraj* was used to measure hydrogen bond distance over time during replicate

simulations started from a model ~7 ps after the initial formation of the FF helix, as described in Supplemental Figure 4A. The amide nitrogen to carbonyl oxygen distance is plotted for each 1 ps frame. A line is shown for reference at 3 Å, a typical distance between heavy atoms in a hydrogen bond. The bond between the Gly 906 carbonyl oxygen and the Leu 1014 amide nitrogen forms intermittently throughout these and all other simulations performed (not shown).

Supplemental Figure 7 – Roles of the LLSSL (residues 1010-1014) segment in autoinhibition

In these experiments, we pooled data from cells expressing low to medium levels of EGFR, in order to not include cells that are overexpressing EGFR and exhibiting non-EGF stimulated Erk phosphorylation. Histograms that show the distribution of pErk levels in the selected population of cells were generated and used for comparison between wild-type receptor and the mutants. For wild-type receptor, the pErk distribution shifts from low levels to high levels upon EGF stimulation. In contrast for the Δtail-3 construct, in which residues 1005-1051 are deleted, there is a marked shift of pErk level towards higher end of the distribution even without EGF stimulation (Figure 3C). Also, with EGF stimulation, the Δtail-3 construct shows slightly higher level of pErk in comparison to wild-type (Figure 3C).

A. Mutating only the LLSSL segment also shows an activating effect on the pErk level. For the EESSE construct, replacement of the hydrophobic leucine to negatively charged glutamine causes roughly half of the cells to show a higher level of pErk, in the absence of EGF. However, there is no further shift of pErk distribution for this construct after EGF stimulation.

B. Replaceing leucine by glutamine residues in this LLSSL segment failed to activate the receptor, as shown by the similar distributions of pErk in QQSSQ and wild-type constructs. These results are consistent with the docking being mediated principally by backbone interactions, which are disrupted by the introduction of negative charge, but not by the neutral glutamine replacements.

Supplemental Figure 8 – HER3 tail phosphorylated tyrosine antibody specificity

Flow cytometry was used to evaluate anti-HER3-pY1289 antibody (Cell Signalling Cat# 4791) specificity toward EGFR and an EGFR-HER3 tail chimera. Strong signal was obtained with cells expressing the EGFR-Her tail chimera upon EGF stimulation and the virtually no signal was measured with cells expressing EGFR, confirming the specificity of the antibody used for the data shown in Fig. 9.

Supplemental Figure 9 – Vesicle-reconstituted TM-ICM EGFR exhibits the same biphasic autophosphorylation kinetics after phosphatase treatment

Time courses of phosphorylation reactions as quantified by Western blotting. Data is presented before and after phosphatase treatment of the samples.

Supplemental Figure 10 - Incorporation of a TM-ICM construct of EGFR into

unilamellar vesicles induces membrane stacking

A. Electron micrographs of vesicles with or without incorporation of purified EGFR protein. Extruded DOPC vesicles were reconstituted with the proteins indicated above each micrograph, incubated with or without ATP, and prepared for either negative stain (left three columns) or electron cryomicroscopy (right-most column) as indicated. Protein constructs are as follows: His_6–kinase domain + tail, EGFR, residues 672–1186 with an N-terminal His_6 tag; TM-ICM, EGFR residues 618–1186 with a C-terminal GB1 tag.

B. Illustration showing the measured dimensions of the stacked structures seen in negative stain preparations shown in **A**.

C. Hypothetical model for the transition of protein-loaded vesicles to flattened membrane stacks. N-lobe to C-lobe asymmetric dimer interactions form between kinases in separate bilayers, and membranes flatten and rearrange to satisfy many such interactions.

Supplemental Figure 11 – Electron cryomicroscopy images

A. Empty vesicles. Enlarged image of panel in Fig. 11A.

B. EGFR TM-ICM reconstituted in vesicles leads to membrane stacking. Enlarged image of panel in Fig. 11A.

Supplemental Figure 12 – Kinetic model of EGFR autophosphorylation

Data from stochastic simulations described in the Methods section.

Supplemental References

Kabsch, W. and C. Sander (1983). "Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features." Biopolymers **22**(12): 2577-2637.

Vedhachalam, C., P. T. Duong, M. Nickel, D. Nguyen, P. Dhanasekaran, H. Saito, G. H. Rothblat, S. Lund-Katz and M. C. Phillips (2007). "Mechanism of ATP-binding cassette transporter A1-mediated cellular lipid efflux to apolipoprotein A-I and formation of high density lipoprotein particles." J Biol Chem **282**(34): 25123-25130.

Waggoner, T. A., J. A. Last, P. G. Kotula and D. Y. Sasaki (2001). "Self-assembled columns of stacked lipid bilayers mediated by metal ion recognition." J Am Chem Soc **123**(3): 496-497.

Zhang, L., J. Song, G. Cavigiolio, B. Y. Ishida, S. Zhang, J. P. Kane, K. H. Weisgraber, M. N. Oda, K. A. Rye, H. J. Pownall and G. Ren (2011). "Morphology and structure of lipoproteins revealed by an optimized negative-staining protocol of electron microscopy." J Lipid Res **52**(1): 175-184.

Zhang, X., J. Gureasko, K. Shen, P. A. Cole and J. Kuriyan (2006). "An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor." Cell **125**(6): 1137-1149.

Supplemental Figure 1

A

B
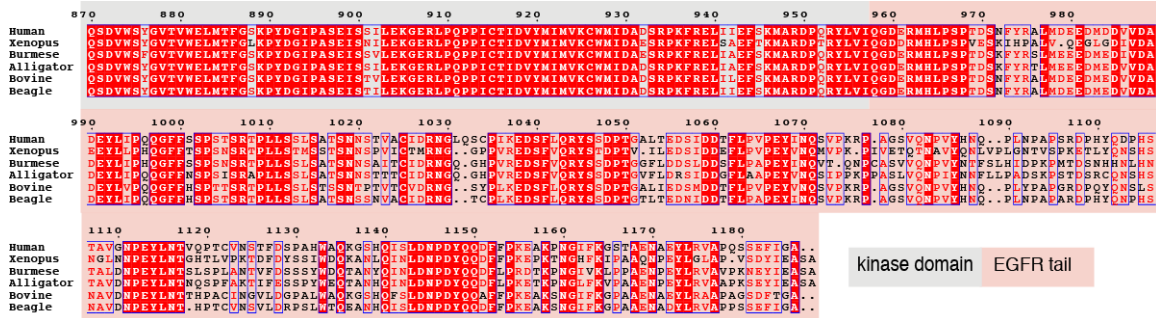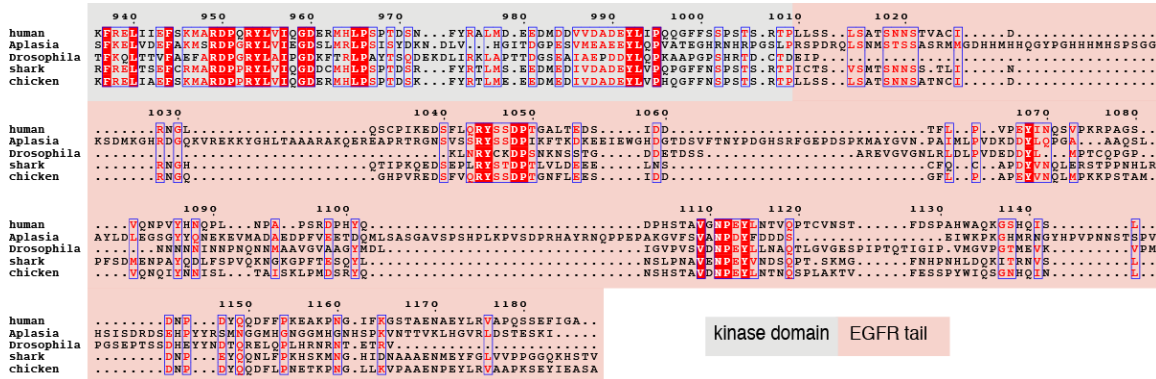
Supplemental Figure 2
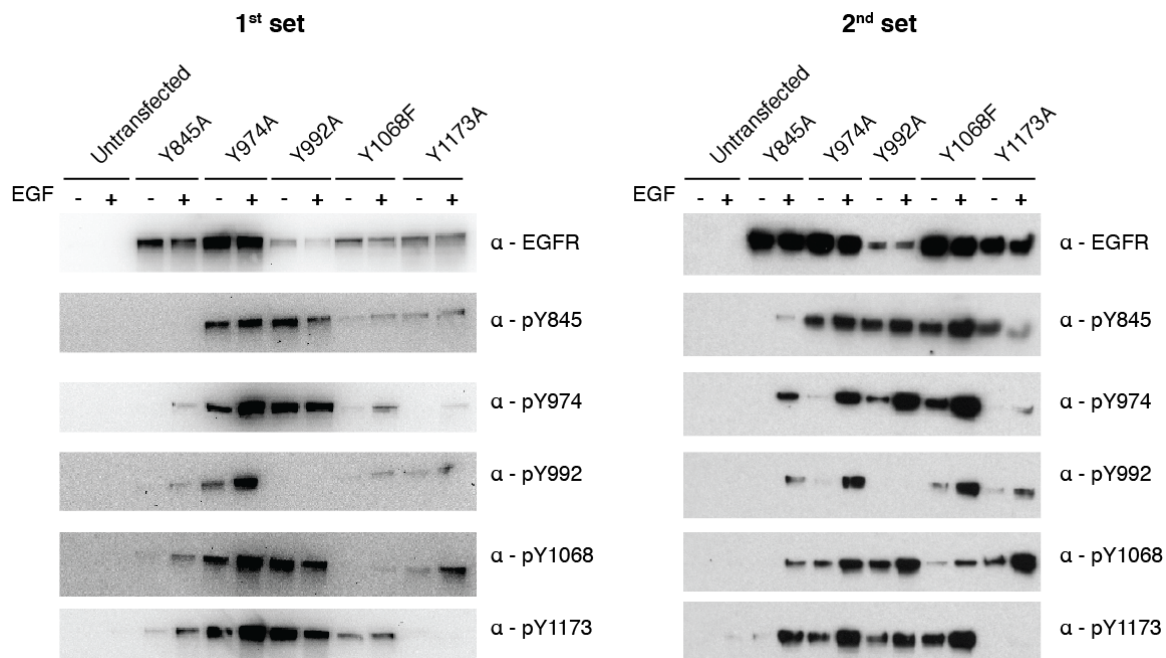
Supplemental Figure 3A
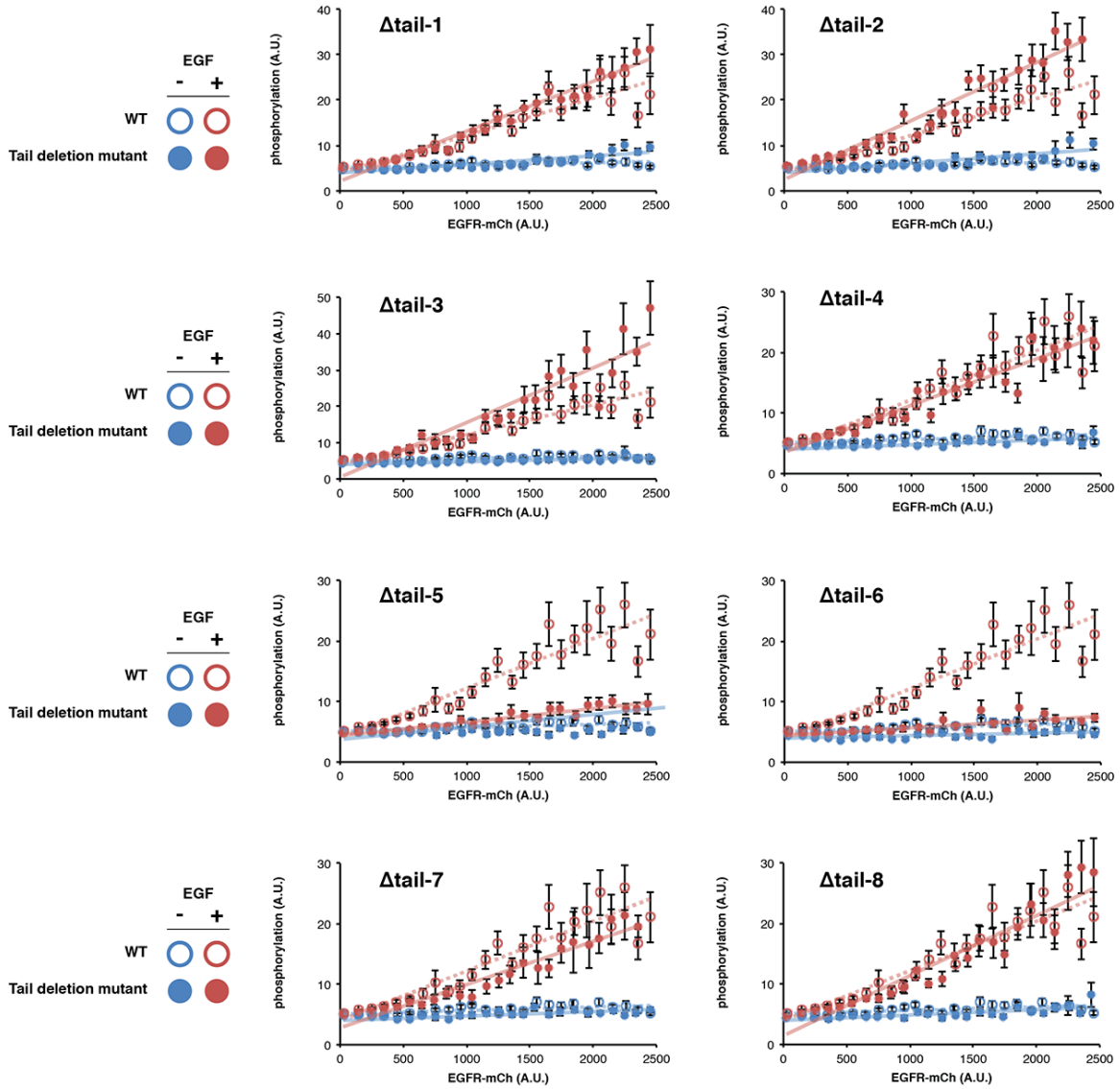
Supplemental Figure 3B

Supplemental Figure 4

- EGF

+ EGF

pErk level

EGFR expression

pErk level

EGFR expression

pErk level

EGFR expression

pErk level

EGFR expression

pErk level

EGFR expression

pErk level

EGFR expression

Cell #

pErk

Cell #

pErk

Supplemental Figure 5

A

EGFR V924R (monomer A and
- ● Proteolytic fragment
- ● AP2-helix
- ● Activation loop

B

C

○ Original EGFR V924R structure
● New EGFR V924R structure
- ● αC helix
- ● Activation loop
- ● AP2-helix

D

Met 983

E

Mig6 clash

EGFR I682Q structure (monomer A / B)
- ● αC helix
- ● Activation loop
- ● AP2-helix
○ EGFR / Mig6 complex structure
- ● Mig6

Supplemental Figure 6A

Supplemental Figure 6B

Supplemental Figure 7

A

EESSE - EGF
EESSE +EGF
WT – EGF
WT + EGF

B

QQSSQ - EGF
QQSSQ +EGF
WT – EGF
WT + EGF
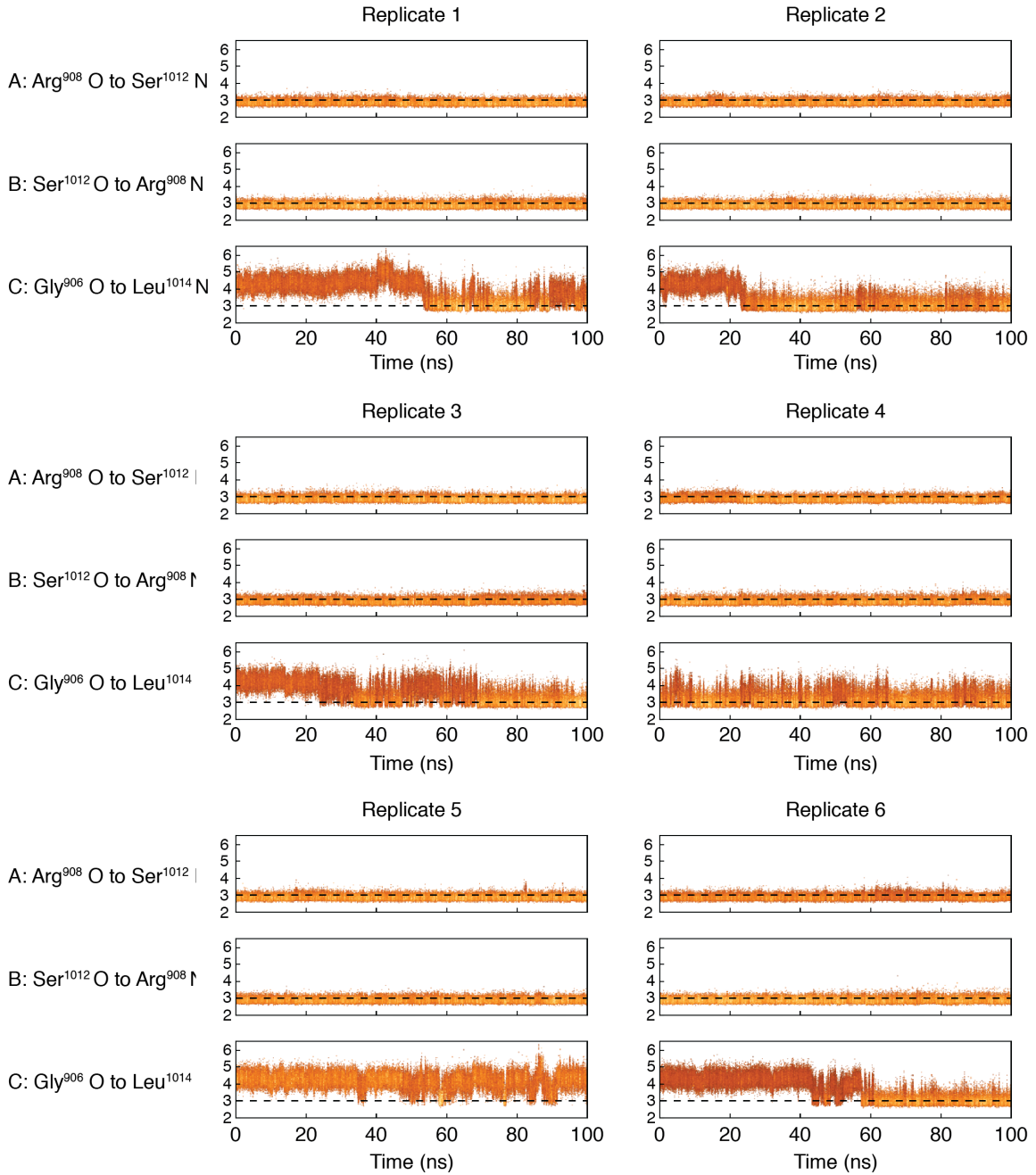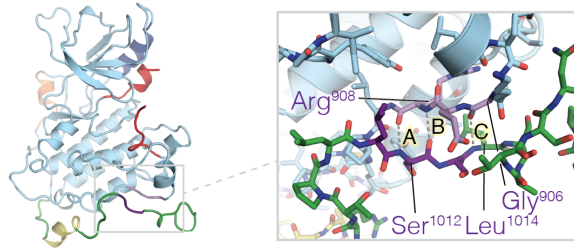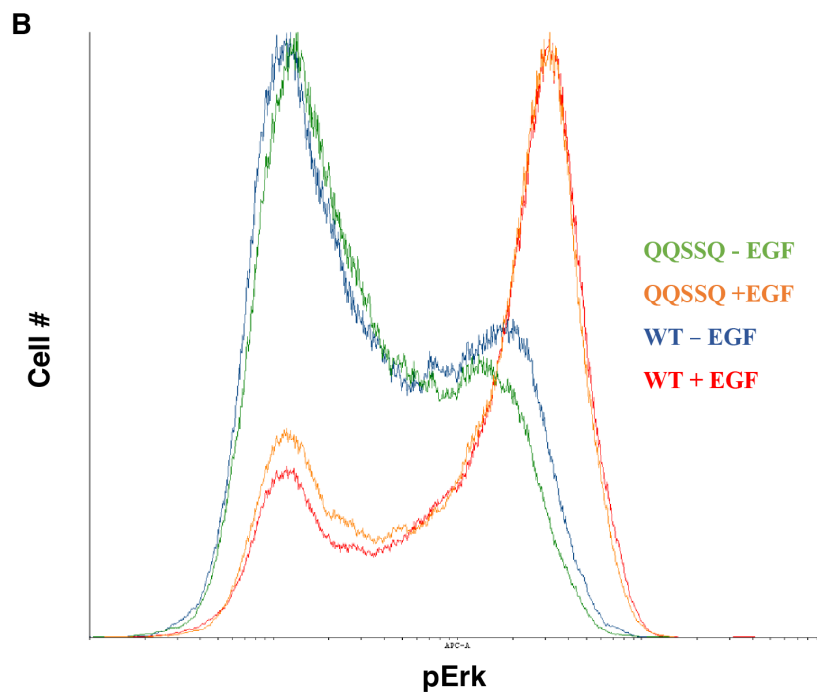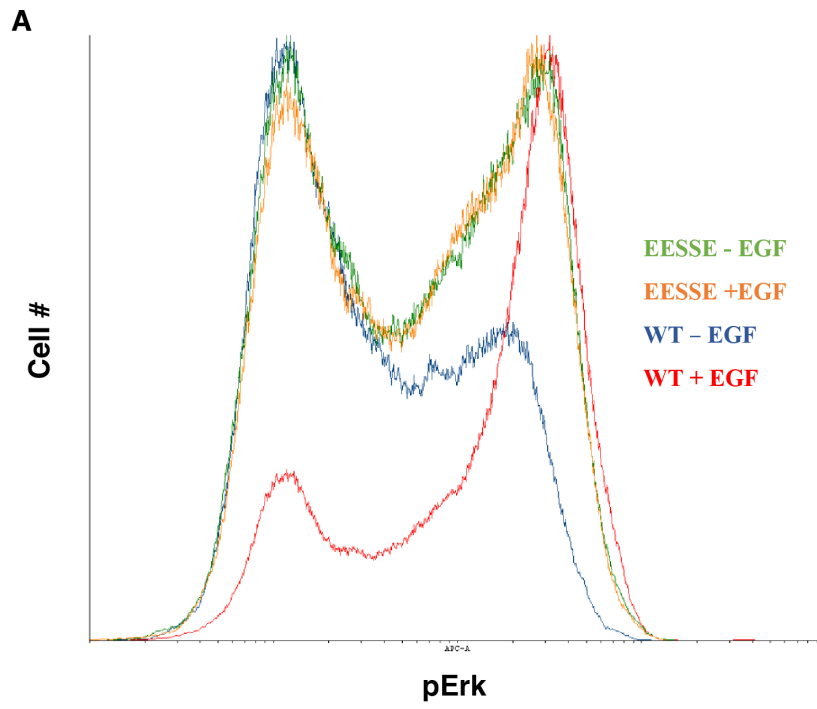
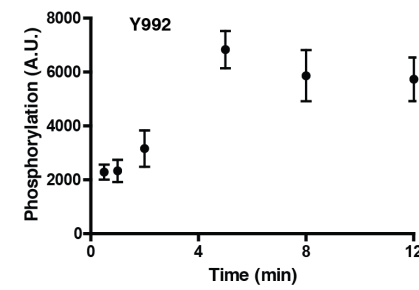**Supplemental Figure 8**
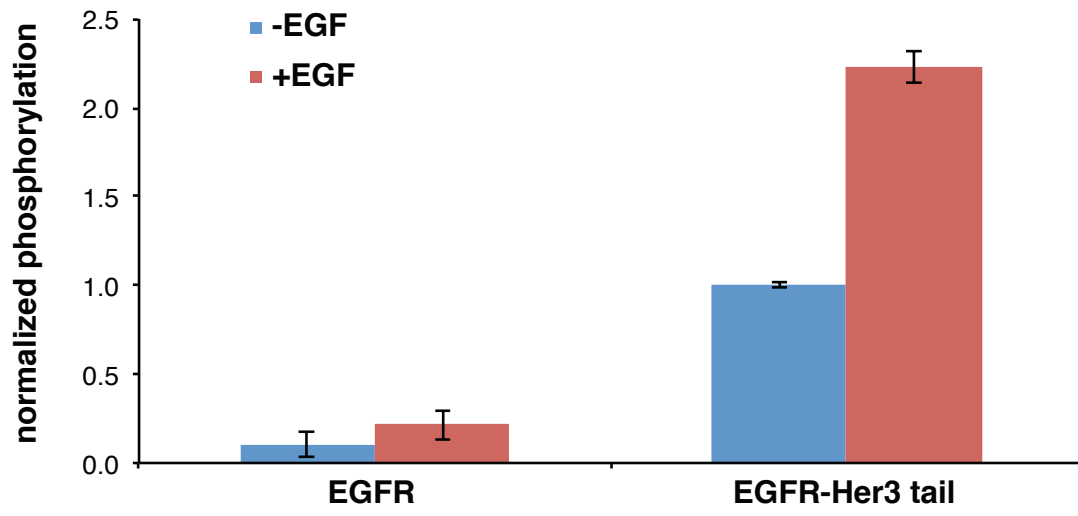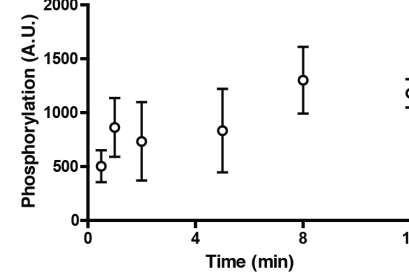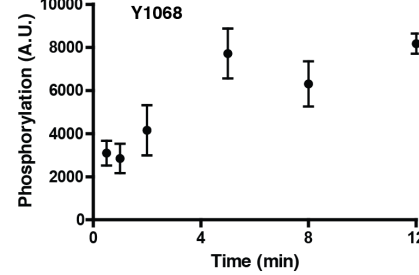
Supplemental Figure 9
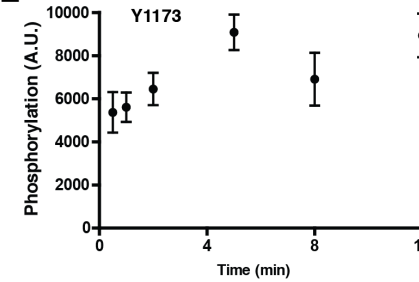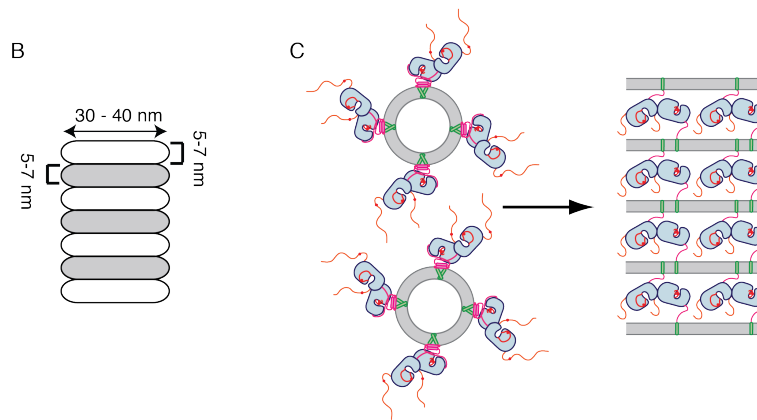
A

● before phosphatase   ○ after phosphatase

B

C

D

E

Supplemental Figure 10

A

vesicles only

His$_6$–kinase domain + tail

TM-ICM V924R

vesicles only cryoEM

100 nm

TM-ICM + ATP

TM-ICM + ADP

TM-ICM V924R + TM-ICM I682Q

TM-ICM cryoEM

B

30 - 40 nm

5-7 nm

5-7 nm
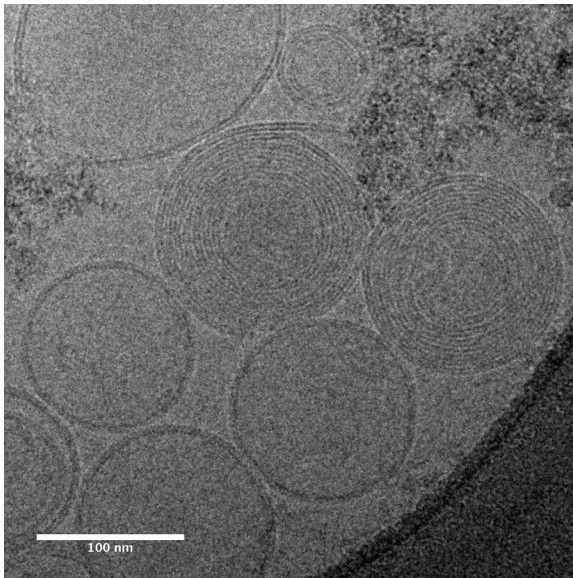
C
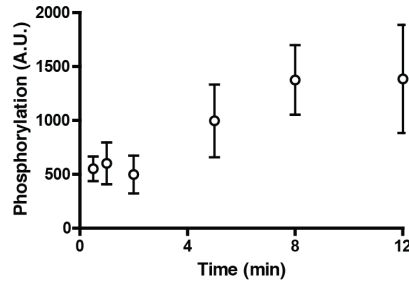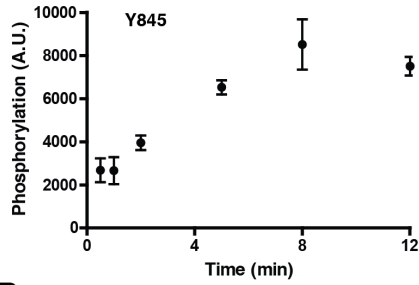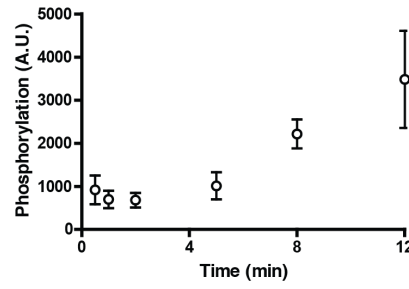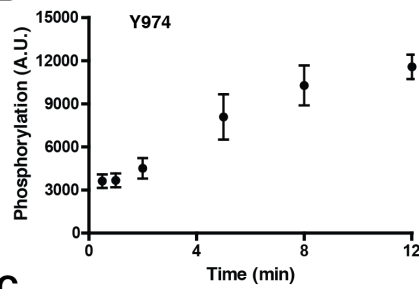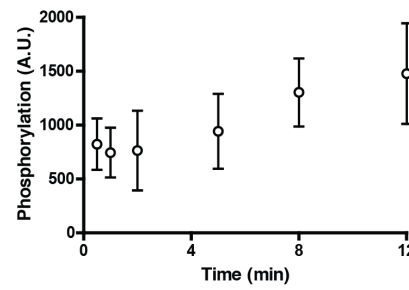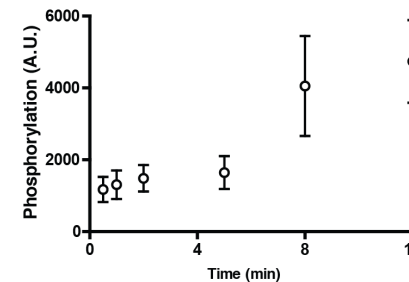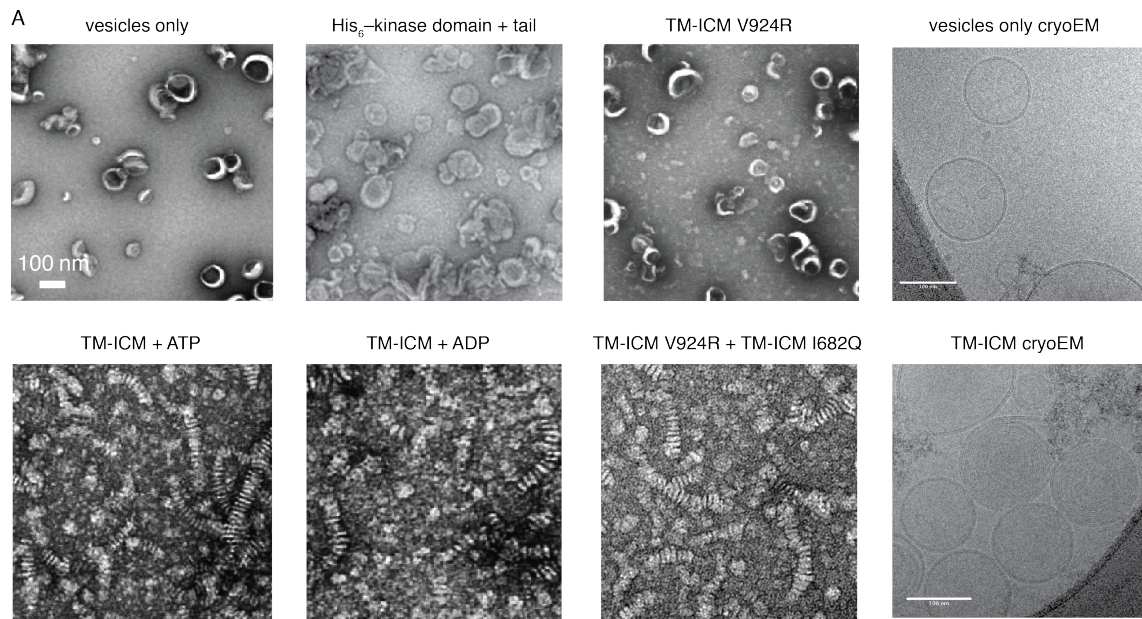
Supplemental Figure 11

A

B

Supplemental Figure 12

$$2\text{EGFR} \xrightarrow{\textit{fast}} (\text{EGFR})_2 \qquad \text{(constitutive dimerization on membranes)}$$

$$(\text{EGFR})_2 \xrightarrow{\textit{fast}} \text{pEGFR·EGFR} \qquad \text{(autophosphorylation in a dimer)}$$

$$\text{pEGFR·EGFR} \xrightarrow{\textit{slow}} \text{pEGFR} + \text{EGFR} \qquad \text{(dimer disassembly)}$$

$$\text{pEGFR} + \text{EGFR} \xrightarrow{\textit{fast}} 2\text{pEGFR} \qquad \text{(monomer \textit{trans} - autophosphorylation)}$$