

Supporting Information

Shi et al. 10.1073/pnas.1002753107

SI Text

SI Materials and Methods. Modeling ErbB3 and EGF Receptor (EGFR) Kinase Domains in Complex with Mg^{2+} -ATP and Substrate. Our model for the tyrosine kinase domain (TKD) of ErbB3 in the inactive conformation was constructed from the crystal structure of ErbB3 reported here and shown in Fig. 3A. The model for EGFR-TKD in the active conformation was constructed from Protein Data Bank (PDB) entry 1M14 (1). In addition, as a computational reference we constructed a model for ErbB3 TKD in an active-like conformation by homology modeling with the active EGFR-TKD (PDB code 1M14) and ErbB4-TKD (PDB code 3BCE) as templates (1, 2). A sequence alignment between the TKDs of ErbB3, EGFR, and ErbB4 was generated in MODELLER (3, 4) and was used to generate an initial structure of active-like ErbB3-TKD by using the coordinates of active EGFR and ErbB4 TKDs. Missing residues were built by using the loop modeling algorithm in MODELLER, and the structures were further refined by performing additional energy minimization steps and assessed for stereochemical quality by using PROCHECK (5).

Initial models of TKD ternary complexes with ATP, Mg^{2+} , and a 7 amino acid substrate peptide were built on the basis of the 2GS6 crystal structure of EGFR-TKD with a bound bisubstrate analogue of ATP and peptide (6), in which all components are resolved. For ErbB3 (in inactive and active conformations) the kinase structure was superposed on the 2GS6 complex by using an rmsd-based alignment of C^α backbone atoms and a 7-residue peptide (VPEYINQ) derived from the Y1068 site of the EGFR C-terminal tail plus ATP were placed on the basis of rmsd alignment with the bisubstrate analogue in 2GS6. Because the transition state for the phosphoryl-transfer reaction in kinases is believed to involve two divalent metal ions (Mg^{2+}) in the active site, two Mg^{2+} ions were placed in complex with the kinase-bound ATP on the basis of the 1ATP structure of protein kinase A (active conformation of cAMP-dependent protein kinase or PKA) (7).

System Preparation for Quantum Mechanics/Molecular Mechanics (QM/MM) Simulations. Hydrogen atoms were added to the ternary complex structures by using the HBUILD utility in CHARMM (8). The protonation states of histidines were chosen according to recommendations from the WHATIF web interface (<http://swift.cmbi.kun.nl>). Protonation states for all titratable groups distal to the catalytic site were chosen as those expected in aqueous solution on the basis of their individual pK_a values (9). Given that the protonation states of titratable side chains in the active site can have considerable effect on the stability of the ground state as well as on the energetics of the reaction landscape, the different protonation states of the conserved aspartate D813 in EGFR, proposed to function as a catalytic base, were evaluated by using a Poisson–Boltzmann approach (10–12) to evaluate the relative stabilities of unprotonated and protonated forms of the D813 side chain. Multiple protonation states of the conserved catalytic aspartate (D831 in EGFR and D833 in ErbB3) were not considered because the location is relatively removed from the atoms involved in the catalytic reaction pathway. Although recent studies of the Abl kinase have reported that the protonation state of the corresponding aspartate is important in the flipping of the Asp-Phe-Gly motif toward the active conformation (13), such a flip for EGFR is already facilitated, even in the unprotonated state for D831, by kinase dimerization (14). All models were explicitly solvated by using the 3-site (TIP3P) model for water (15), and the buffering distance was set to 15 Å. Sodium (Na^+) and

chloride (Cl^-) ions were added to achieve net electroneutrality of the system and an ionic strength of 75 mM. All Na^+ and Cl^- ions were placed at least 8 Å away from any protein atoms and from each other. The ions were positioned at points of electrostatic extrema by using a Debye–Hückel potential calculated at 300 K by using the program Solvate 1.0 (16). In order to prepare the system for dynamics simulations, the solvated structures were energy minimized by using a conjugate gradient algorithm to remove unfavorable contacts.

QM/MM Simulations. The QM/MM approach we adopted is based on an existing interface between GAMESS-UK (17) and CHARMM (8). The MM region was modeled by using the CHARMM27 force-field parameters (18), and the wave function optimizations in the QM region were performed by using electronic density functional theory. For the quantum region, we chose the two Mg^{2+} ions, water molecules within 5 Å of the Mg^{2+} ions, segments of the ATP and the tyrosine substrate, and two catalytic residues (D813 and D831 in EGFR; N815 and D833 in ErbB3), yielding a total of 60 atoms for EGFR and 64 atoms for ErbB3, including three link atoms each. In order to handle the phosphorous chemistry reliably, the quantum region was treated by using density functional theory with a Becke 3-parameter exchange, Lee, Yang, and Parr exchange correlation function and a 6–31G* basis set (19), and the remaining protein and solvent molecules were treated classically by using the CHARMM27 force field. Because the QM/MM boundary cuts across covalent bonds, we applied the single link atom procedure (20) to satisfy the valences of broken bonds in the QM region. The electrostatic interactions between the quantum region and the classical atoms belonging to groups at the QM/MM boundary were discarded, because this choice has been shown to enhance the accuracy in the resulting energies and geometries (21). We have extensively explored this choice of the QM/MM region, including sensitivity to the size of the QM region and the functional/basis set combination in our prior studies on closely related systems (22–24); others have validated the choice of the link atom method (21). Nonbonded van der Waals (VdW) interactions were treated by applying a switching function at 10 Å and truncating the VdW potential energy smoothly at a cutoff distance of 12 Å. The forces in the quantum region were calculated on the fly, assuming that the system moves on the Born–Oppenheimer surface defined by the QM/MM Hamiltonian. The system was first subjected to 1,200 steps of the adopted basis Newton–Raphson minimization, and subsequently QM/MM MD simulations were performed during which the systems were first heated to 300 K and then subjected to a constant temperature of 300 K dynamics using a Langevin thermostat for 10 ps using a standard 1-fs time step of integration.

Catalytic Reaction Pathways. The exact reaction coordinate is complex and multi- (high) dimensional, and a quantitative estimate of the free energy landscape cost can be obtained only by using efficient long-time sampling approaches (25–29), which remain prohibitively expensive for QM/MM applications because of large computational cost. Although we believe that the true reaction coordinate is complex, in the spirit of a recent study (30), we describe the reaction pathways in terms of several simple reaction coordinates χ_j . For phosphoryl transfer through the associative mechanism, χ_1 is the distance between the reactive tyrosyl oxygen and the ATP γ -phosphate, χ_2 is the distance between the tyrosyl O- and ATP O2/3 β , and χ_3 is the ATP P γ -ATP O2/3 β distance.

For phosphoryl transfer through dissociative mechanisms, in addition to these three distances, χ_4 includes the coordinate for proton abstraction, namely, the distance between the tyrosyl OH- and D830:O δ_2 . In order to drive the system along a reaction coordinate, we performed restrained minimization as well as restrained sampling simulations by using the QM/MM Hamiltonian by augmenting the Hamiltonian in the i th window with a potential bias term $\sum_j w_j (1/2) \times K \times (\chi_j - \chi_j^{i,0})^2$; here, $w_j = 0$ or 1 is the weight for χ_j , K is the effective spring constant (ranging between 10 and 30 kcal \cdot mol $^{-1} \cdot \text{\AA}^{-2}$), and $\chi_j^{i,0}$ are reference values for sampling in different windows. Complete pathways were assembled by varying offset distances $\chi_j^{i,0}$ in increments of $\pm 0.05 \text{\AA}$, where the positive and negative increments represent the directionality of the energy scan. All reaction paths were originated from configurations with tyrosine $\chi_1 = \text{O-ATP Py}$ distance of 3 \AA (resulting from the classical umbrella sampling described below) and computed by energy minimizations in the presence of restraints and recalculating the single-point energies in the absence of the restraints. For our choice of the increments of $\chi_j^{i,0}$, calculations performed in the forward and reverse directions satisfied path reversibility in configurations as well as energy values (see Fig. 4, main text). Following the energy scans, constant temperature restrained dynamics runs of 300 fs–1 ps for each window were also performed to ensure that the reaction path was robust to thermal fluctuations (Fig. S5).

Sampling the Precatalytic Free Energy Landscape. The transition state for the phosphoryl-transfer reaction is thought to involve a trigonal-bipyramidal geometry stabilized by two divalent metal ions in the active site and applicable to several kinases as well as

to DNA and RNA polymerases (31–33). Because the positions and orientations of the ions and ATP relative to the substrate tyrosines in our initial modeled structures were determined either from a bisubstrate ligand complex or by homology with protein kinase A (see above), prior to QM/MM simulations we first explored the classical free energy landscape associated with positioning the nucleophilic oxygen on the tyrosine at a distance of 3 \AA from the target phosphorous and appropriately positioned relative to the catalytic Mg $^{2+}$ ion in our models of EGFR and ErbB3. We achieved this precatalytic sampling of conformational space through classical umbrella sampling simulations, which enable the calculation of the potential of mean force (free energy density) along an a priori chosen set of reaction coordinates from which free energy changes are calculated by numerical integration (34). We mapped out the free energy landscape for preorganizing the coordinate χ_1 (the tyrosyl O-ATP Py distance) at 3 \AA , in order to ensure uniformity across all models with respect to the starting configuration in subsequent QM/MM simulations of phosphoryl-transfer pathways. For computing the free energy landscape, the probability distribution $P(\chi_1)$ is calculated by dividing χ_1 into several (6 or 7) windows. The histograms for each window are collected by harvesting classical molecular dynamics trajectories of length 200 ps per window by using a harmonic restraint $0.5 \times K \times (\chi_1 - \chi_1^{i,0})^2$, $\chi_1^{i,0}$ being the preset of the reaction coordinate characterizing the i th sampling window, and $K = 20 \text{ kcal} \cdot \text{mol}^{-1} \cdot \text{\AA}^{-2}$ being the force constant for the restraint along χ_1 . The sampling data from the different windows are then processed by using the weighted histogram analysis method (34); see Fig. S6.

- Stamos J, Sliwkowski MX, Eigenbrot C (2002) Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. *J Biol Chem* 277:46265–46272.
- Qiu C, et al. (2008) Mechanism of activation and inhibition of the HER4/ErbB4 kinase. *Structure* 16:460–467.
- Fiser A, Do RK, Sali A (2000) Modeling of loops in protein structures. *Protein Sci* 9:1753–1773.
- Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815.
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283–291.
- Zhang X, Gureasko J, Shen K, Cole PA, Kuriyan J (2006) An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125:1137–1149.
- Zheng J, et al. (1993) 2.2 \AA refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallogr D* 49:362–365.
- Brooks BR, et al. (1983) CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 4:187–217.
- Radhakrishnan R, Schlick T (2004) Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymerase beta's closing. *Proc Natl Acad Sci USA* 101:5970–5975.
- Yang AS, Gunner MR, Sampogna R, Sharp K, Hönlig B (1993) On the calculation of pKas in proteins. *Proteins: Struct Funct & Genet* 15:252–265.
- Georgescu RE, Alexov EG, Gunner MR (2002) Combining conformational flexibility and continuum electrostatics for calculating pK(a)s in proteins. *Biophys J* 83:1731–1748.
- Alexov EG, Gunner MR (1997) Incorporating protein conformational flexibility into the calculation of pH-dependent protein properties. *Biophys J* 72:2075–2093.
- Shan Y, et al. (2009) A conserved protonation-dependent switch controls drug binding in the Abl kinase. *Proc Natl Acad Sci USA* 106:139–144.
- Shih AJ, Purvis J, Radhakrishnan R (2008) Molecular systems biology of ErbB1 signaling: Bridging the gap through multiscale modeling and high-performance computing. *Mol Biosyst* 4:1151–1159.
- Jorgensen WL, Chandrasekar J, Madura J, Impey R, Klein M (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935.
- Grubmüller H, Heymann B, Tavan P (1996) Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force. *Science* 271:997–999.
- Schmidt MW, et al. (1993) General atomic and molecular electronic-structure system. *J Comput Chem* 14:1347–1363.
- MacKerell AD, et al. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 102:3586–3616.
- Szabo A, Ostlund NS (1996) *Modern Quantum Chemistry* (Dover Publications, Mineola, NY).
- Field MJ, Bash PA, Karplus M (2002) A combined quantum mechanical and molecular mechanical potential for molecular dynamics simulations. *J Comput Chem* 11:700–733.
- Das D, et al. (2002) Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method. *J Chem Phys* 117:10534–10547.
- Radhakrishnan R (2007) Coupling of fast and slow modes in the reaction pathway of the minimal hammerhead ribozyme cleavage. *Biophys J* 93:2391–2399.
- Radhakrishnan R, Schlick T (2005) Fidelity discrimination in DNA polymerase beta: Differing closing profiles for a mismatched (G:A) versus matched (G:C) base pair. *J Am Chem Soc* 127:13245–13253.
- Radhakrishnan R, Schlick T (2006) Correct and incorrect nucleotide incorporation pathways in DNA polymerase β . *Biochem Biophys Res Comm* 350:521–529.
- Brooks CL, Case DA (2003) Theory and simulation—The control and timescale of structure and reactivity in biological systems: from peptide folding to cellular networks. *Curr Opin Struct Biol* 13:143–145.
- Elber R, Cárdenas A, Ghosh A, Stern H (2003) Bridging the gap between long time trajectories and reaction pathways. *Adv Chem Phys* 126:93–129.
- Chu J-W, Trout BL, Brooks BR (2003) A super-linear minimization scheme for the nudged elastic band method. *J Chem Phys* 119:12708–12717.
- Bolhuis PG, Chandler D, Dellago C, Geissler PL (2002) Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annu Rev Phys Chem* 53:291–318.
- Ren W, Vanden-Eijnden E, Maragakis P, E WN (2005) Transition pathways in complex systems: Application of the finite-temperature string method to the alanine dipeptide. *J Chem Phys* 123:134109.
- Rosta E, Woodcock HL, Brooks BR, Hummer G (2009) Artificial reaction coordinate “tunneling” in free-energy calculations: The catalytic reaction of RNase H. *J Comput Chem* 30:1634–1641.
- Steitz TA (1993) DNA- and RNA-dependent DNA polymerases. *Curr Opin Struct Biol* 3:31–38.
- Steitz TA (1999) DNA polymerases: Structural diversity and common mechanisms. *J Biol Chem* 274:17395–17398.
- Steitz TA, Smerdon SJ, Jäger J, Joyce CM (1994) A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* 266:2022–2025.
- Roux B (1995) The calculation of the potential of mean force using computer-simulations. *Comput Phys Comm* 91:275–282.
- Jura N, Shan Y, Cao X, Shaw DE, Kuriyan J (2009) Structural analysis of the catalytically inactive kinase domain of the human EGF receptor 3. *Proc Natl Acad Sci USA* 106:21608–21613.

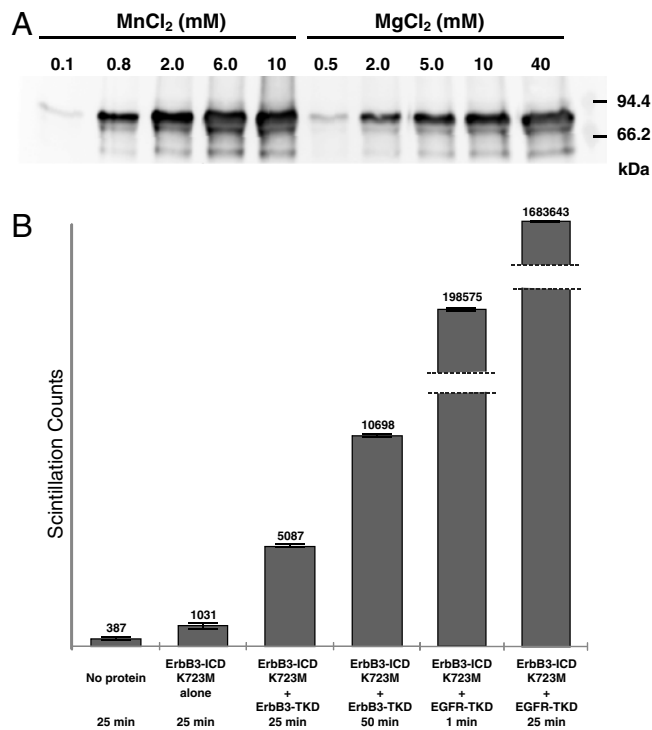


Fig. S1. (A) Dependence on divalent cation concentration of ErbB3-ICD autophosphorylation. Purified ErbB3-ICD⁶⁶⁵⁻¹³²³ (3 μ M) was incubated for 30 min in 100 mM MOPS pH 7, 150 mM NaCl, 5% glycerol, containing 1 mM ATP, 100 μ M phospholipid (in vesicles) with 10% (mol/mol) DOGS-NTA-Ni, and the noted concentrations of MnCl₂ or MgCl₂. Samples were subjected to SDS-PAGE and immunoblotting with antiphosphotyrosine antibodies. (B) Quantitation of ³²P incorporation into ErbB3-ICD by the kinase domains from ErbB3 or EGFR. A final concentration of 3 μ M wild-type (histidine-tagged) ErbB3-TKD⁶⁶⁵⁻¹⁰⁰¹ or EGFR-TKD was coincubated with 6 μ M his₆-ErbB3-ICD⁶⁶⁵⁻¹³²³ (K723M) as substrate in 20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 5% glycerol, containing 0.1 mM DTT, 5 mM MgCl₂, 2 mM MnCl₂, 100 μ M ATP (with 3 μ Ci [γ -³²P]-ATP) and 100 μ M lipid (vesicles containing 10% DOGS-NTA-Ni + 90% DOPC). Reaction mixtures were incubated at room temperature for the times listed, and reactions were stopped by adding EDTA to a final concentration of 50 mM. Samples were spotted onto phosphocellulose paper to immobilize protein. After a series of washes, ³²P incorporation was measured by scintillation counting. Mean \pm range of variation are plotted from at least two independent experiments. ³²P incorporation catalyzed by ErbB3-TKD occurred at approximately 209 cpm per minute (5,087 in 25 min; 10,698 in 50 min). ³²P incorporation catalyzed by EGFR-TKD was clearly saturating (or depleting substrate) by 25 min. On the basis of the 1-min time point, we estimate a maximum incorporation rate of 198,575 cpm per minute, suggesting a rate that is \sim 940-fold higher than that for ErbB3-TKD.

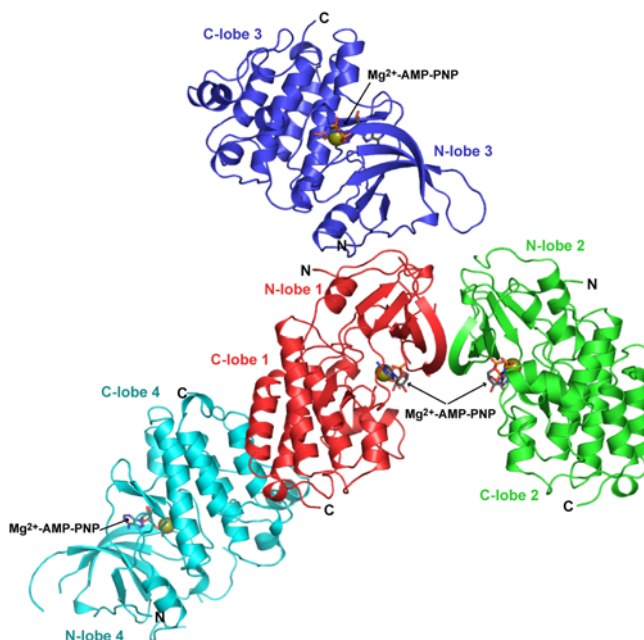


Fig. S2. Crystal packing of ErbB3-TKD⁶⁶⁵⁻¹⁰⁰¹. Cartoon representation of the ErbB3-TKD⁶⁶⁵⁻¹⁰⁰¹ structure, showing major crystal contacts. The green structure is molecule A in the asymmetric unit, and the red structure is molecule B. These associate in the crystals through N-lobe-mediated interactions. In addition, interaction partners in the crystal are shown for molecule B (Red) that associate through alternative N-lobe-mediated (Dark Blue) and C-lobe-mediated (Cyan) contacts. The packing of ErbB3-TKD⁶⁶⁵⁻¹⁰⁰¹ molecules in our crystals appears to be the same as those recently described and discussed in detail by Jura et al. (35).

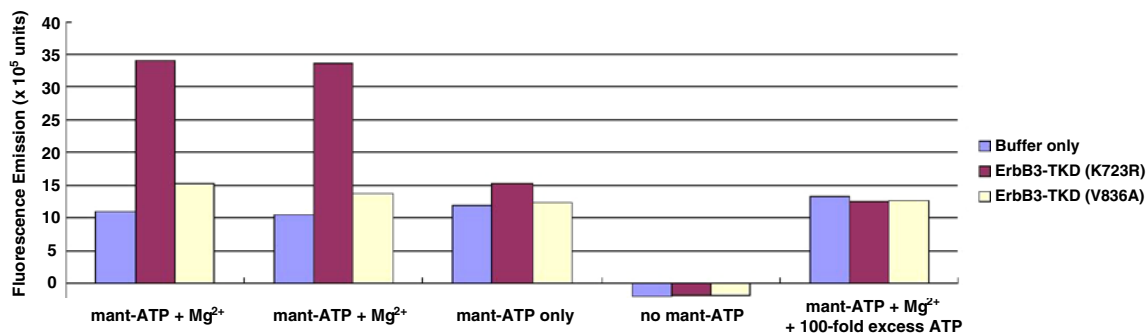


Fig. S3. Mutating V836 in ErbB3-TKD^{648–1001} disrupts mant-ATP binding. Samples containing buffer alone or mutated ErbB3-TKD at 10 μ M were incubated alone (fourth set of bars), with 5 μ M mant-ATP (third set of bars), with 5 μ M mant-ATP plus 10 mM MgCl₂ (first and second sets of bars), or with 5 μ M mant-ATP, 10 mM MgCl₂ plus 500 μ M ATP (fifth set of bars). Assays were performed in 20 mM Tris pH 8.0, containing 200 mM NaCl, 5% glycerol. Fluorescence measurements were taken in triplicate by using a Tecan SAFIRE II plate reader, with excitation wavelength of 340 nm and emission wavelength of 450 nm (10-nm bandwidth for both). The K723R mutant binds mant-ATP with an affinity similar to that seen for wild-type protein and shows a substantial increase in fluorescence emission from mant-ATP that can be competed away by addition of excess ATP. By contrast, the V836A mutant displayed little evidence of binding, consistent with the loss of the ability of V836A-mutated ErbB3-ICD to autophosphorylate as assessed in immunoblotting assays. Note that the first and second sets of bars are experimental repeats.

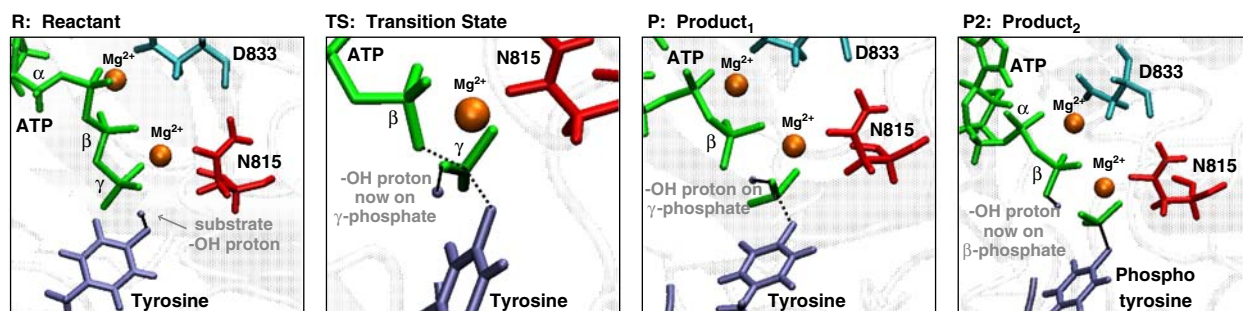


Fig. S4. Representative snapshots of intermediate states in the reaction pathway for phosphoryl transfer catalyzed by ErbB3-TKD in the inactive-like state. States correspond to “R” (reactant); “TS” (transition state with trigonal-bipyramidal geometry around P _{γ}); “P” (product representing the completion of phosphoryl transfer with proton still bound to O1 γ of ATP); and “P2” (product with proton transferred to O2 β of ATP). The reaction pathway is characterized by migration of the proton from the hydroxyl group of the tyrosine substrate to the O γ 1 oxygen of ATP (in TS) and its subsequent transfer to the O β 2 atom of ATP in P2. Phosphoryl transfer coincides with this proton transfer and occurs through an associative mechanism with characteristic values of 1.9 Å for the formation and cleavage of bonds in the transition state (λ_a and λ_c in Fig. 4B, main text). No evidence was found in our simulations for proton transfer mediated through N815—as expected. Thus, only pathway II in Fig. 4B (main text) ensues for ErbB3. Energy changes along the reaction pathway for ErbB3 (in the structure described in this manuscript) are summarized in Fig. 4C, and states (R, TS, P, and P2) are marked. The activation energy (E_a) for the captured associative mechanism of phosphoryl transfer is estimated to be 23 kcal/mol (Fig. 4C).

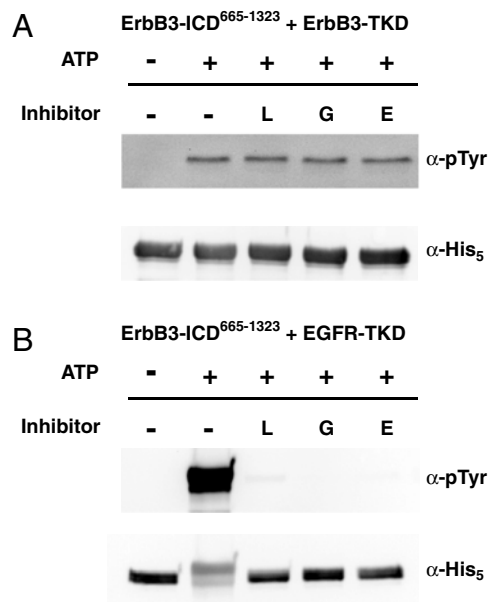


Fig. S7. Western blot analysis to assess the effect of EGFR or EGFR/ErbB2 inhibitors on (A) ErbB3-TKD activity and (B) EGFR-TKD activity. Phosphorylation of ErbB3-ICD⁶⁶⁵⁻¹³²³ was monitored by Western blotting with antiphosphotyrosine antibodies. ErbB3-ICD⁶⁶⁵⁻¹³²³ (0.3 μ M) was mixed with 0.3 μ M ErbB3-TKD (A) or 0.3 μ M EGFR-TKD (B) in 20 mM Tris-HCl, pH8.0, 200 mM NaCl, 5% glycerol, 0.1 mM DTT, 5 mM MgCl₂, 2 mM MnCl₂, 100 μ M ATP, and 10 μ M lipid (10% NTA-NiDOGS + 90% DOPC in vesicles) for 1 hr (A) or 1 min (B) at 25 °C. Samples were then subjected to immunoblotting with pY20 antiphosphotyrosine (Upper) and anti-His₅ antibody for normalization (Lower). Experiments were performed with and without ATP. To assess inhibitor specificity, experiments were performed in the presence of 4 μ M lapatinib (L), 4 μ M gefitinib (G), 4 μ M erlotinib (E), or without inhibitor (DMSO only: "-"). All three inhibitors blocked EGFR activity, but left *trans*-phosphorylation by ErbB3-TKD unaffected. Blots are representative of at least 3 independent experiments.

Table S1. Data collection and refinement statistics (molecular replacement)

Data collection	ErbB3 TKD APS (GM/CA CAT)
Space group	C2
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> (Å), β (°)	193.4, 48.0, 82.2, $\beta = 108.1^\circ$
Resolution (Å)	50.00-2.80 (2.85-2.80)*
R_{sym} or R_{merge}	0.083 (0.499)
$I/\sigma I$	17.1 (2.2)
Completeness (%)	90.0 (55.3)
Redundancy	6.3 (3.4)
Refinement	
Resolution (Å)	50.00-2.80
No. reflections	14,406
$R_{\text{work}}/R_{\text{free}}$	25.2/28.5
Model	
Protein amino acids	aa A680-728, 732-844, 854-959 aa B680-711, 713-729, 732-845, 854-959
No. atoms	4,316
Protein	4,245
Ligand/ion	64
Water	7
rmsd	
Bond lengths (Å)	0.006
Bond angles (°)	1.099

*Values in parentheses are for highest-resolution shell.