Co-Conserved Features Associated with *cis* Regulation of ErbB Tyrosine Kinases

Amar Mirza[1], Morad Mustafa[1], Eric Talevich[2], Natarajan Kannan[1,2]*

1 Department of Biochemistry and Molecular Biology, University of Georgia, Athens, Georgia, United States of America, **2** Institute of Bioinformatics, University of Georgia, Athens, Georgia, United States of America

Abstract

Background: The epidermal growth factor receptor kinases, or ErbB kinases, belong to a large sub-group of receptor tyrosine kinases (RTKs), which share a conserved catalytic core. The catalytic core of ErbB kinases have functionally diverged from other RTKs in that they are activated by a unique allosteric mechanism that involves specific interactions between the kinase core and the flanking Juxtamembrane (JM) and COOH-terminal tail (C-terminal tail). Although extensive studies on ErbB and related tyrosine kinases have provided important insights into the structural basis for ErbB kinase functional divergence, the sequence features that contribute to the unique regulation of ErbB kinases have not been systematically explored.

Methodology/Principal Findings: In this study, we use a Bayesian approach to identify the selective sequence constraints that most distinguish ErbB kinases from other receptor tyrosine kinases. We find that strong ErbB kinase-specific constraints are imposed on residues that tether the JM and C-terminal tail to key functional regions of the kinase core. A conserved RIxKExE motif in the JM-kinase linker region and a glutamine in the inter-lobe linker are identified as two of the most distinguishing features of the ErbB family. While the RIxKExE motif tethers the C-terminal tail to the N-lobe of the kinase domain, the glutamine tethers the C-terminal tail to hinge regions critical for inter-lobe movement. Comparison of the active and inactive crystal structures of ErbB kinases indicates that the identified residues are conformationally malleable and can potentially contribute to the *cis* regulation of the kinase core by the JM and C-terminal tail. ErbB3, and EGFR orthologs in sponges and parasitic worms, diverge from some of the canonical ErbB features, providing insights into sub-family and lineage-specific functional specialization.

Conclusion/Significance: Our analysis pinpoints key residues for mutational analysis, and provides new clues to cancer mutations that alter the canonical modes of ErbB kinase regulation.

Introduction

The epidermal growth factor receptor (EGFR) and related kinases, ErbB2, ErbB3, and ErbB4 (collectively called the ErbB family) [1], are key components of our cellular machinery that control signaling pathways associated with cell migration, proliferation, and differentiation. Understanding how ErbB kinases respond to specific activation and regulatory signals in these pathways is essential for the development of new therapies for human cancers that are associated with abnormal regulation of ErbB kinase activity [2]. The domain architecture of ErbB kinases, like most receptor tyrosine kinases (RTKs), is characterized by an extracellular ligand-binding domain, a transmembrane domain, a juxtamembrane (JM) segment, a kinase domain, and a COOH-terminal tail (C-terminal tail). The kinase domain adopts a bi-lobal structure consisting of an N-terminal ATP-binding lobe (N-lobe) and a C-terminal substrate-binding lobe (C-lobe) [3,4].

Extensive studies on the extracellular ligand-binding domain [5,6,7,8,9] and more recently on the intracellular kinase domain of

EGFR [10,11,12] have provided key insights into how EGFR converts an extracellular signal into an intracellular response. Ligand binding to the extracellular receptor induces dimerization and activation of the intracellular kinase domain, which, upon activation, autophosphorylates conserved tyrosine residues in the C-terminal tail [13,14]. Autophosphorylation of the tyrosine residues activates downstream signaling pathways by recruiting signaling and docking proteins to the C-terminal tail [2,15,10]. In the absence of an extracellular ligand, EGFR is maintained in an inactive dimeric form [11], which prevents formation of the active dimer. ErbB3 is believed to diverge from this canonical mechanism of action because of its inability to catalyze phosphoryl transfer [16]. Recent studies, however, have challenged this view by showing low, but detectable levels of ErbB3 autophosphorylation [17].

A key step in ErbB kinase signaling is the activation of the intracellular kinase domain, which is achieved by an intermolecular interaction between two kinase molecules in an asymmetric dimer [10]. In the asymmetric dimer, the C-lobe of one kinase molecule (the "activator") allosterically activates the other ("receiver") by

inducing conformational changes in key regions of the receiver [10]. In particular, the regulatory C-helix in the N-lobe of the receiver kinase switches from an inactive "out" to an active "in" conformation upon dimerization, and the flexible activation loop in the C-lobe of the kinase switches from a substrate-inaccessible conformation to a substrate-accessible conformation [10,11,12,18]. Also, the N-lobe of the kinase domain moves from an ATP-accessible "open" conformation to an ATP-inaccessible "closed" conformation [15,19,20]. These conformational changes, which occur upon activation of many protein kinases [21,22,23,24], are tightly regulated to avoid physiological catastrophes [25,26].

The catalytic activity of EGFR is also regulated by conformational changes in the JM and C-terminal tail—two sequence segments flanking the kinase domain. The JM segment functions as an activation domain [12] by facilitating the formation of the asymmetric dimer [11]. Specifically, the JM segment of the receiver docks to the C-lobe of the activator to stabilize the asymmetric dimer [11,12]. This docking interaction is prevented in the inactive dimer of EGFR [11], where the JM docking surface on the C-lobe is shielded by the C-terminal tail. Presumably, the conformational changes associated with the JM and C-terminal tail during kinase activation are closely coupled with the conformational changes in the kinase core (described above) for the tight regulation of kinase activity [27,28]. The atomic details of how this coupling is achieved are not fully understood.

Receptor tyrosine kinases (RTKs) outside of the ErbB family also contain flexible JM and C-terminal tail segments that play important regulatory roles. In c-KIT and Ephrin receptor tyrosine kinases, for example, the JM segment plays an autoinhibitory role, in contrast to its activating role in EGFR, by interacting with the active site [29] and the substrate-binding regions of the kinase domain [30]. Likewise, the C-terminal tail in Tie2 inhibits catalytic activity by an autoinhibitory mechanism [31], which is distinct from EGFR [27]. Thus, individual RTKs have evolved unique mechanisms to regulate catalytic activity by the JM and C-terminal tails. Information regarding these family-specific regulatory mechanisms are encoded in the protein sequences—the cell's own medium for specifying molecular mechanisms. However, despite the availability of RTK sequences from diverse organisms (~3000 sequences), the sequence features that contribute to the unique modes of regulation in individual RTKs have not been systematically delineated.

We have shown using several case studies that Bayesian analysis of the evolutionary constraints distinguishing functionally divergent kinases is a viable approach for investigating the functional specificity of kinases in signaling pathways [32,33]. Using this approach, we recently identified that a conserved C-terminal tail which wraps around the kinase core of AGC kinases is a distinguishing feature of the AGC group [34]. This study also revealed novel AGC kinase-specific motifs in the C-terminal tail that were experimentally shown to be important for AGC kinase functions [35,36]. In this study, we compare the functional constraints acting on ErbB and related RTKs to identify the key residues/motifs that contribute to ErbB kinase functional divergence. We show, for the first time, that nearly all the residues that distinguish ErbB kinases from other RTKs are involved in tethering the JM and C-terminal tail to key functional regions of the kinase core. Analysis of these tethering interactions in light of the wealth of structural and functional data available on the ErbB kinases suggests a model in which the identified residues contribute to ErbB kinase functional specialization by facilitating a unique *cis* interaction between the kinase core and the flanking JM and C-terminal tails. Our analysis provides new testable hypotheses regarding the *cis* regulation of the kinase core by the JM and C-

terminal tails, and provides new insights into cancer mutations that alter this mode of regulation.

Results and Discussion

A co-conserved sequence pattern characteristic of the ErbB kinase domain

To identify which sequence features most distinguish ErbB kinases from other RTKs, we measured and analyzed the selective constraints imposed on ErbB kinase sequences from diverse organisms (see Methods). These constraints generally correspond to residues that are highly conserved within the ErbB family, but strikingly different in RTKs outside of the ErbB family (Figure 1). Within the catalytic core, these residues correspond to W731, P733, G735, E736, V738, K739, P741 in the β2-β3 loop; Y764 and S768 in the C-helix; S784 in the β4-β5 loop; Q791, P794, C797 in the inter-lobe linker; V802 in the D-helix; G810, N816, W817 in the E-helix; P848 in the β8 strand (not shown); L861 in the activation loop (not shown); I/V904 in the F-helix (not shown); I938 in the αG-αH loop; I/L941 and D942 in the H-helix. Among these residues, Q791 in the inter-lobe linker contributes the most (indicated by the height of the histogram in Figure 1) to ErbB functional divergence, since none of the RTKs outside of the ErbB family conserve a glutamine at the 791 position (Background alignment in Figure 1). The residues described above also distinguish ErbB kinases from non-receptor tyrosine kinase (NRTK) sequences, as NRTKs also conserve strikingly different residues at these positions. The only exceptions are W731 in the β2-β3 loop and L861 in the activation loop. These two residues are conserved in ErbB's as well as in some NRTKs.

The JM and C-terminal tail contribute to ErbB kinase functional divergence

In addition to the kinase domain, strong ErbB-specific constraints are also imposed on residues flanking the kinase core, namely, the Juxtamembrane segment, the JM-kinase linker and the COOH-terminal tail.

ErbB-specific constraints in the JM and JM-kinase linker. The JM segment is conserved across diverse organisms within the ErbB family (Figure 1). However, across RTKs, the JM segment displays little or no detectable sequence similarity. This indicates that the JM segment is unique to the ErbB family and likely contributes to its functional divergence. Some of the most distinguishing residues in the JM region include: E690, P694, S695 and N700. Unlike the JM, the JM-kinase linker of ErbB kinases share significant structural similarity with the JM-kinase linker of other RTKs [37], despite very low sequence similarity. This is indicated by the shared hydrophobic residues (L/I 707, L712) between ErbB kinases (Foreground alignment in Figure 1) and RTKs (Background alignment in Figure 1). The JM-linker region also contains several residues that distinguish ErbB kinases (Foreground) from other RTKs (Background). These include: R705, I706, K708, E709, E711 and K714 (Figure 1).

The C-terminal tail is a distinguishing feature of ErbB kinases. The C-terminal tail is also a distinctive feature of the ErbB family. In particular, the sequence segment immediately following the kinase domain (residues 971-1020 in Fig 1) is highly conserved in ErbB kinases, but strikingly different in RTKs outside of the ErbB family. The C-terminal tail segment is also co-conserved with key regions of the kinase domain (see below). Some of the distinctive residues/motifs in the C-terminal tail segment include: [MF][AC][RK]DPxR[YF]LVI motif in the beginning of the C-tail, D/E994 and F/L997 in the middle, and [DE]x[DE]xYL motif at the C-terminal end (Figure 1).
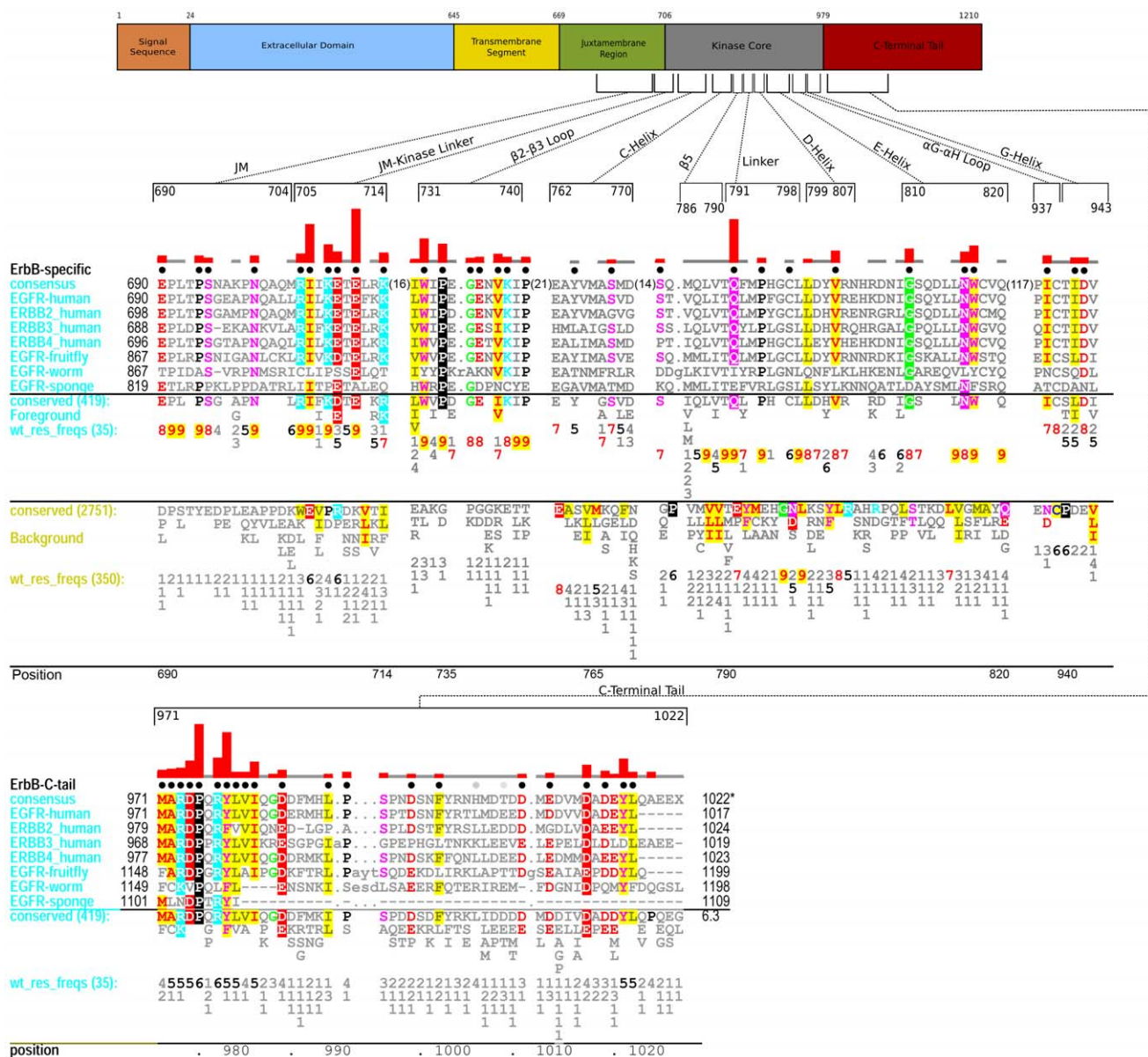
Figure 1. Contrast Hierarchical Alignment showing sequence patterns that most distinguish ErbB kinases (foreground alignment) from other receptor tyrosine kinase (RTK) sequences (background). The residues identified by the Bayesian pattern partitioning procedure (see Methods) as distinctive of the ErbB family are indicated by black dots above the alignment. The histograms on top of the alignment indicate the degree to which residue composition in the foreground alignment (ErbB sequences) contrast with residues observed at the corresponding position in the background alignment (other RTKs). The foreground set includes the sequences shown in the alignment and others whose conservation levels are denoted by the consensus pattern and corresponding weighted residue frequencies (wt_res_freqs) below the alignment. Residue frequencies are indicated in integer tenths where, for example, a '6' indicates that the corresponding residue directly above it occurs 60-70% of the weighted sequences. The number of weighted sequences and the total number of alignment sequences are indicated in parentheses next to 'wt_res_freqs' and 'conserved', respectively. The background alignment and the corresponding residue frequencies are shown directly below the foreground alignment. The structural location of the ErbB kinase-conserved residues and the overall domain organization of the ErbB family are shown above the alignment. The numbering used in the alignment and in the text is according to the pre-mature EGFR numbering, which includes the 24 amino acid signaling sequence. A background alignment for the C-terminal tail region is not shown because the C-terminal tail of ErbB kinases shares no detectable sequence similarity with the C-terminal tail of other RTKs. Thus, a standard background alignment consisting of protein sequences from NCBI-nr database was used to quantify the constraints acting on the C-terminal tail residues. The NCBI sequence identifiers used in the query display alignment are: EGFR-human: 134104655; ERBB2_human: 119533; ERBB3_human: 119534; ERBB4_human: 3913590; EGFR-fruitfly: 4588511; EGFR-worm: 212645651; EGFR-sponge: 18146642. doi:10.1371/journal.pone.0014310.g001

Lineage and sub-family-specific variations within the ErbB family

The ErbB prototypic features, described above, are generally well conserved across diverse eukaryotic phyla (Figure 1). However, some lower eukaryotes and parasitic worms diverge from the canonical ErbB features, and display correlated sequence changes in the JM, kinase, and C-terminal tail regions (Figure 1). For example, a distinctive glutamine (Q791) in the inter-lobe linker region is conserved as a glutamate (E) in sponges (Figure 1). This change is correlated with the absence of the C-terminal tail [DE]x[DE]xYL motif (Figure 1), which typically interacts with Q791 in mammalian ErbB's (Figure 2). Thus, the correlated

sequence change observed in the kinase and C-terminal tail suggests possible co-evolution of these two regions during mammalian ErbB kinase evolution.

ErbB3, an atypical member of the ErbB family, also displays significant variations in some of the canonical ErbB motifs/residues. For example, a phosphorylatable tyrosine (Y1016) within the C-terminal tail [DE]x[DE]xYL motif is conserved as an aspartate (D) in ErbB3. Likewise, ErbB3 and ErbB4 replace a canonical tyrosine (Y764) in the C-helix by a leucine (L) (Figure 1). The structural and functional implications of these family-specific variations are discussed in the sections below.

ErbB conserved residues are frequently mutated in human cancers

Since ErbB kinases are one of the most frequently mutated gene families in human cancers, we investigated whether any of the identified ErbB conserved residues are among those known to be associated with human cancers. Mapping of somatic mutations identified in the ErbB family (see Methods) to available crystal structures indicates that several of the ErbB kinase conserved residues are indeed mutated in human cancers. S768 in the C-helix, and L861 in the activation loop, are two of the most frequently mutated residues in EGFR (Table 1). In addition, ErbB conserved residues in the JM-kinase linker and β2-β3 loop are also frequently mutated in lung, esophagus and upper digestive track cancers (Table 1). The structural/functional impacts of these mutations, however, are not fully understood.

Structural analysis of ErbB kinase-conserved residues and proposed roles

To understand how the identified ErbB kinase conserved residues contribute to ErbB kinase functional specialization, and how mutations of these residues contribute to disease, we performed crystal structure analysis of the identified residues (see methods). As shown in Figure 2, nearly all the ErbB conserved residues, although widely dispersed in sequence, spatially interact
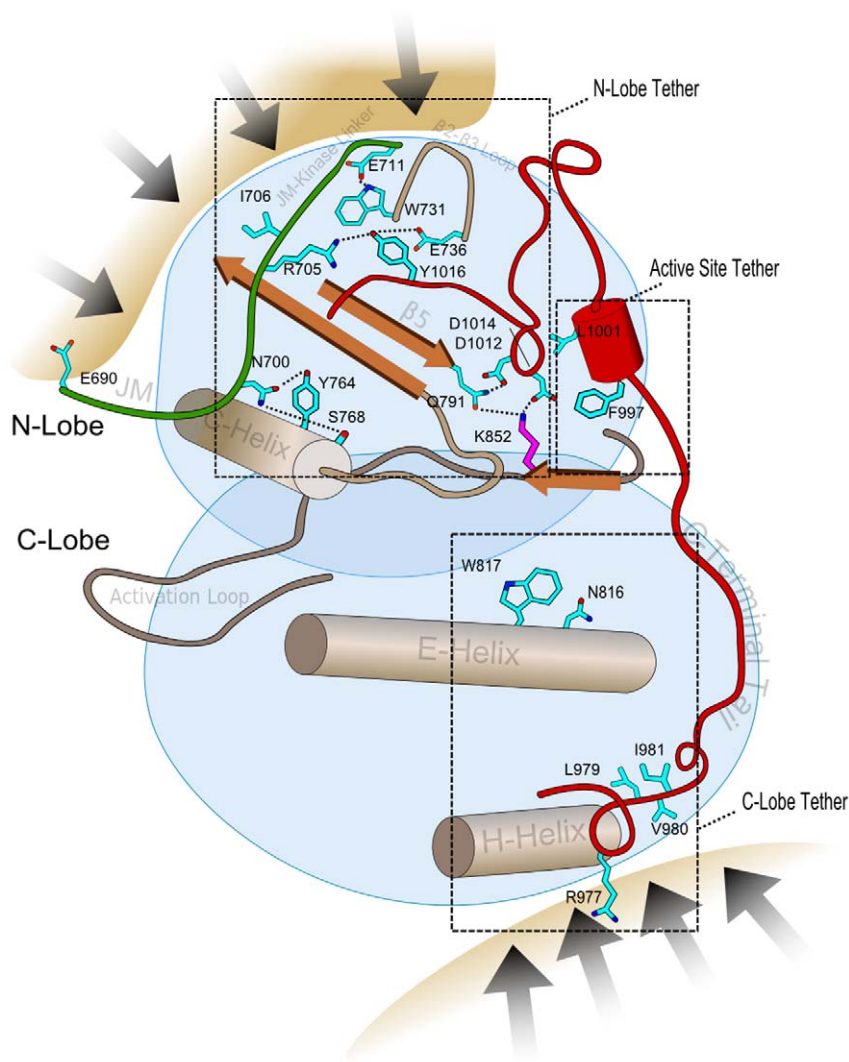

Figure 2. A schematic of the ErbB kinase domain showing the distinctive ErbB residues and associated interactions (based on PDB: 2J5F). These residues are broadly classified into three categories based on their structural location and their role in tethering the JM and C-terminal tail: (i) N-lobe tether (NLT), (ii) Active site tether (AST) and (iii) C-lobe tether (CLT). The residues are shown in stick representation. ErbB kinase-conserved residues are colored in cyan and kinase conserved residues are colored in magenta. The sites of homo/hetero-dimerization are shown by dark arrows.
doi:10.1371/journal.pone.0014310.g002

Table 1. Somatic mutations in EGFR targeting the canonical ErbB kinase-specific residues.

Mutation	Cancer Primary Site	Structural location
P694L(1) P694S(1)	Lung	JM
S695G(1)	Thyroid	JM
I706T(1)	Lung	JM-kinase linker
K708M(1)	Lung	JM-kinase linker
E709A(8) E709G(4) E709H(2) E709K(9) E709V(2)	Lung, prostate	JM-kinase linker
E711K(1)	Lung	JM-kinase linker
W731R(1)	Lung	β2-β3 loop
P733L(1) P733S(1) P733T(1)	Lung	β2-β3 loop
G735S(3)	Prostate, Lung	β2-β3 loop
V738G(2)	Prostate	β2-β3 loop
P741H(1) P741L(2)	Thyroid Central-nervous-system	β3 strand aa
S768C(1) S768I(21) S768I(23) S768N(1)	Lung, Oesophagus, Central-nervous-system	C-helix
S784F(2) S784P(1) S784Y(1)	Lung, Upper_aerodigestive_tract	β4-β5 loop
V802F(2) V802I(1)	Lung, Upper_aerodigestive_tract	αD-helix
G810D(2) G810S(1)	Upper_aerodigestive_tract, Lung	αE-helix
P848L(3) P848L(4)	Upper_aerodigestive_tract, Lung	β7-β8 loop
L861F(1) L861P(1) L861Q(26) L861Q(30) L861R(5) L861R(3) L861V(1)	Lung, Central-nervous-system	Activation loop
F968L(1)	Central-nervous-system	I-helix
D1012H(1)	Lung	C-terminal Tail

doi:10.1371/journal.pone.0014310.t001

with the flexible JM and C-terminal tail segments to "tether" them to three regions of the kinase core, namely, the N-lobe, the C-lobe and the active site (Figure 2). Because these interactions are malleable in crystal structures (Table S1), we use the term "tether" [34] to describe these interactions. Broadly, ErbB conserved residues can be classified into three categories based on their structural location and interaction: (i) N-Lobe Tether (NLT): residues that tether the JM and C-terminal tail to the kinase N-lobe; (ii) Active Site Tether (AST): residues that tether the C-terminal tail to the ATP binding site; and (iii) C-Lobe Tether (CLT): residues that tether the JM and C-terminal tail to the kinase C-lobe (Figure 2; Figure 3).

The NLT, AST and CLT residues are highly conserved in ErbB kinases, but strikingly different in RTKs outside of the ErbB family. This selective conservation is likely to be important for

ErbB kinase functions, rather than for maintaining ErbB kinase structure or fold, because RTKs that lack these residues essentially adopt the same fold as ErbB kinases [30,31,38]. Indeed, recent studies on the activation mechanism of EGFR and related ErbB kinases support the functional importance of some of the identified residues. For example, the C-terminal tail [MF][AC][RK]DPxR [YF]LVI motif, which is part of the CLT, has been shown to play an important role in ErbB3-ErbB2 hetero-dimerization [4,39], and EGFR activation [11] (Figure 3A). Similarly, D984, a distinctive aspartate (D984) in the C-terminal tail, was recently shown to control C-terminal tail movement and kinase activation [40]. Likewise, N816 and W817 in the CLT have been noted to provide a malleable docking surface for the JM and C-terminal tail in the active [12], and inactive states [11] of EGFR, respectively (Figure 3B-C) [22]. We note that the JM and C-terminal tail docking surface on the C-lobe is coupled to the substrate binding αD-helix [41] via hydrophobic interactions between W817 and V802 in the CLT (Figure 3C).

Whereas the role of CLT residues is well understood, little is known about the role of the NLT and AST residues in ErbB kinase functions. In particular, the selective conservation of residues in the β2-β3 loop and JM-kinase linker are largely mysterious. To obtain insights into these mysterious residues, we performed crystal structure analysis of NLT and AST residues, and interpreted our observations in light of the wealth of functional data available on ErbB kinases. Our analysis suggests important functional roles for the NLT and AST residues, and provides new clues to cancer mutations that alter these residues.

NLT: A structural framework for coupling C-helix and inter-lobe movement in ErbB kinases

As mentioned earlier, activation of EGFR kinase by dimerization involves conformational changes in key regions of the N-lobe, including repositioning of a regulatory C-helix from an inactive to active conformation, and movement of the N-lobe relative to the C-lobe [11,42]. We find that these flexible regions of the N-lobe are tethered to the JM and C-terminal tail via ErbB kinase-conserved interactions described below.

Interactions tethering the JM and regulatory C-helix. Tethering of the JM to the C-helix is mediated through ErbB conserved residues in the C-helix and JM segment. In particular, a conserved asparagine (N700) in the JM segment hydrogen bonds to the side-chain of Y764 and S768 in the C-helix (Figure 4A). While these interactions are stable in the active state of EGFR (see Table S1), in the inactive state these interactions are disrupted, in part, because of repositioning of the C-helix in an inactive "out" conformation [11,42] (Figure 4B). In particular, Y764 in the C-helix moves away from N700 in the inactive state to interact with hydrophobic residues in the β4 strand. This malleable tethering of the JM to the C-helix is likely to be functionally significant, as this may provide a framework for the JM and the activating monomer to dynamically control C-helix movement [11,12]. Consistent with this view, mutation of Y764 to a phenylalanine [43], or N700 to an alanine [12], have been found to significantly impair EGFR kinase activity. Notably, ErbB3 and ErbB4 conserve a leucine at the Y764 position (Figure 4C). This variation may reflect the unique ability of ErbB3 and ErbB4 to form inactive N-lobe-N-lobe dimers, as the leucine, which replaces Y764, is part of the N-lobe-N-lobe dimer interface in ErbB3 and ErbB4 [15,16,17].

Oncogenic mutations, S768I and L861Q, may alter the canonical interactions at the JM-C-helix interface. *S768I*: S768I is a frequently occurring mutation (Table 1) in the C-helix of EGFR that increases basal kinase activity [44]. S768 is located
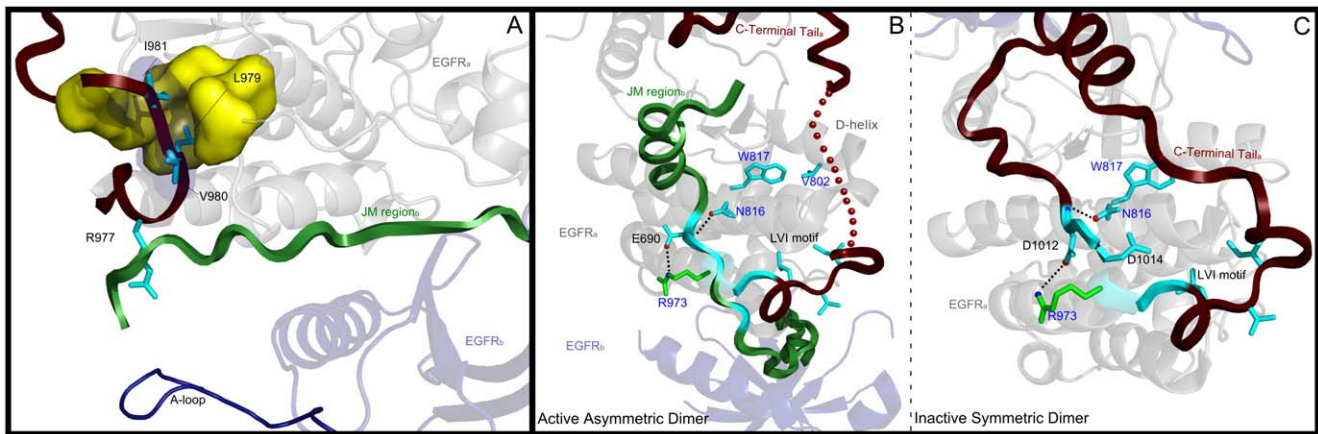
Figure 3. Role of CLT residues in tethering the JM and C-terminal tails. A) Role of the PxR[YF]LVI motif in tethering the C-terminal tail to the C-lobe (PDB:2J5F). B) Tethering of the JM segment to the C-lobe in the active state (PDB: 3GOP). C) Tethering of the C-terminal tail to the C-lobe in the inactive state (PDB: 3GT8). EGFRa and EGFRb correspond to monomer A and monomer B in the asymmetric dimer, respectively. The PxR[YF]LVI motif in the C-terminal tail is labeled as LVI motif. The structure images were generated using Pymol (www.pymol.org). ErbB kinase-specific residues are colored in cyan and hydrogen bonds are depicted as black dotted lines. The disordered segment of the C-terminal tail in the active state (Figure 3B) is shown in dotted representation. Note that R973 is colored in green in Figures 3B and C because this residue is shared by some tyrosine kinases outside of the ErbB family.
doi:10.1371/journal.pone.0014310.g003

at the asymmetric dimer interface and is known to get phosphorylated by Calcium calmodulin-dependent kinase II (CAMK2), which modulates EGFR autokinase activity by phosphorylation of S768 and C-terminal tail serine residues [45]. Thus, mutation of S768 to isoleucine can contribute to abnormal EGFR kinase activity by impacting one or more of the following functions: (i) altering the tethering interactions between the JM and C-helix, (ii) changing the dimer interface [46], and (iii) preventing CAMK2 phosphorylation.

L861Q: L861Q is a frequently occurring activating mutation in the activation loop of EGFR [47]. L861 is specific to ErbB kinases (Figure 5A) and is typically conserved as an aspartate (D) in RTKs outside of the ErbB family (Figure 5A). In the inactive state of

EGFR, L861 packs up against hydrophobic residues in the C-helix [20], and this observation previously led to the suggestion that the L861Q mutation may activate EGFR by destabilizing the hydrophobic interactions in the inactive state [10,46,48]. However, the structural interactions that stabilize the active form of the L861Q mutant have not been proposed before. Modeling of a glutamine in the active form of EGFR indicates that a glutamine at the L861 position can potentially form a hydrogen bond with Y764 (in the C-helix) in the active form, but not in the inactive form of EGFR (Figure 5B). Furthermore, molecular dynamics studies on the L861Q mutant (Figure S1) indicates that the hydrogen bond between Q861 and Y764 is stable during the course of the simulation (Figure 5C), and can likely prevent Y764
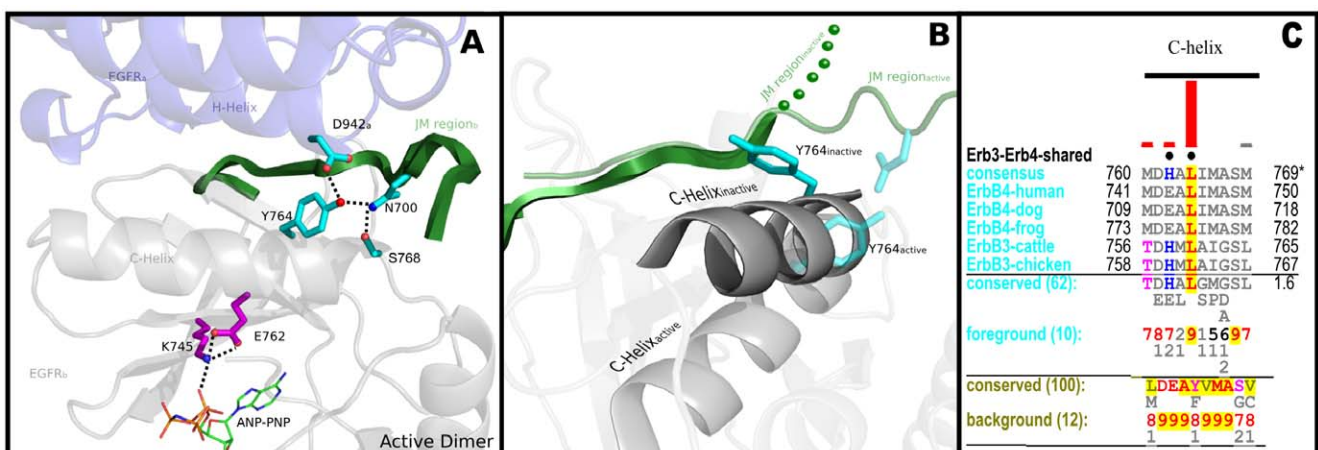

Figure 4. Interactions tethering the JM to the C-helix in the N-lobe. A) Tethering interactions in the active state of EGFR (PDB: 2J5F). The modeled nucleotide (ANP-PNP) is shown in sticks representation. Key hydrogen bonds mediated by ErbB kinase conserved residues are shown. B) Conformational changes associated with ErbB kinase conserved residues in the active (PDB: 2J5F) and inactive state of EGFR (PDB: 2RFE). C) A multiple sequence alignment showing the replacement of Y764 by a leucine in ErbB3 and ErbB4. The foreground alignment in Figure 4C corresponds to ErbB3 and ErbB4 sequences (62 sequences) and the background alignment corresponds to ErbB1 (EGFR) and ErbB2. The NCBI sequence identifiers numbers for the sequences used in the query alignment of Figure 4C are as follows: ErbB4-human: 167745042; ErbB4-dog: 74005688; ErbB4-frog: 147906005; ErbB3-cattle: 156718140; ErbB3-chicken: 113206130.
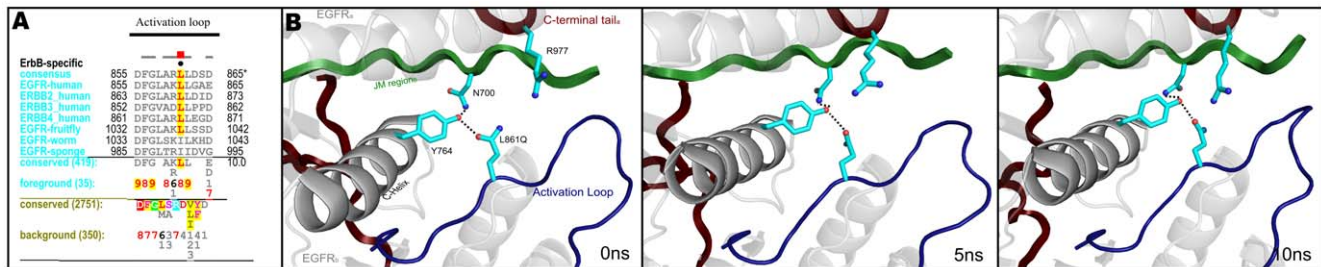doi:10.1371/journal.pone.0014310.g004

Figure 5. Selective conservation and modeling of L861Q mutation (based upon PDB: 2JIU and 2ITN). A) Selective conservation of L861 in the ErbB family. B) Local structural interactions mediated by the L861Q lung cancer mutation in EGFR and snapshots of this interaction at various time points (0, 5 and 10 ns) during the MD run. RMSD plots of the backbone atoms during the course of the simulation for the active (dimer) and inactive (monomer) are shown in Figure S1.
doi:10.1371/journal.pone.0014310.g005

from switching to an inactive conformation (Figure 5B-D). Thus, in addition to destabilizing the inactive state, the L861Q mutation may activate EGFR by stabilizing the C-helix tyrosine (Y764) in an active conformation. We also predict that the L861Q mutation in ErbB3 and ErbB4 may not have the same functional impact as in EGFR because ErbB3 and ErbB4 conserve a leucine at the Y764 position (Figure 4C).

Interactions tethering the C-terminal tail and inter-lobe hinge. The opening and closing of the N-terminal ATP binding

lobe relative to the C-terminal substrate-binding lobe is an essential part of catalysis [23]. Inter-lobe movement in eukaryotic protein kinases is facilitated by the inter-lobe linker [49], which connects the N and C lobes, and lobe-spanning salt bridges, which serve as hinge points for domain movements [50].

In ErbB kinases, the hinge regions of the kinase domain are tethered to the C-terminal tail via ErbB kinase-conserved residues (Figure 6A). In particular, an ErbB kinase-conserved glutamine (Q791) tethers the C-terminal tail to the inter-lobe linker by
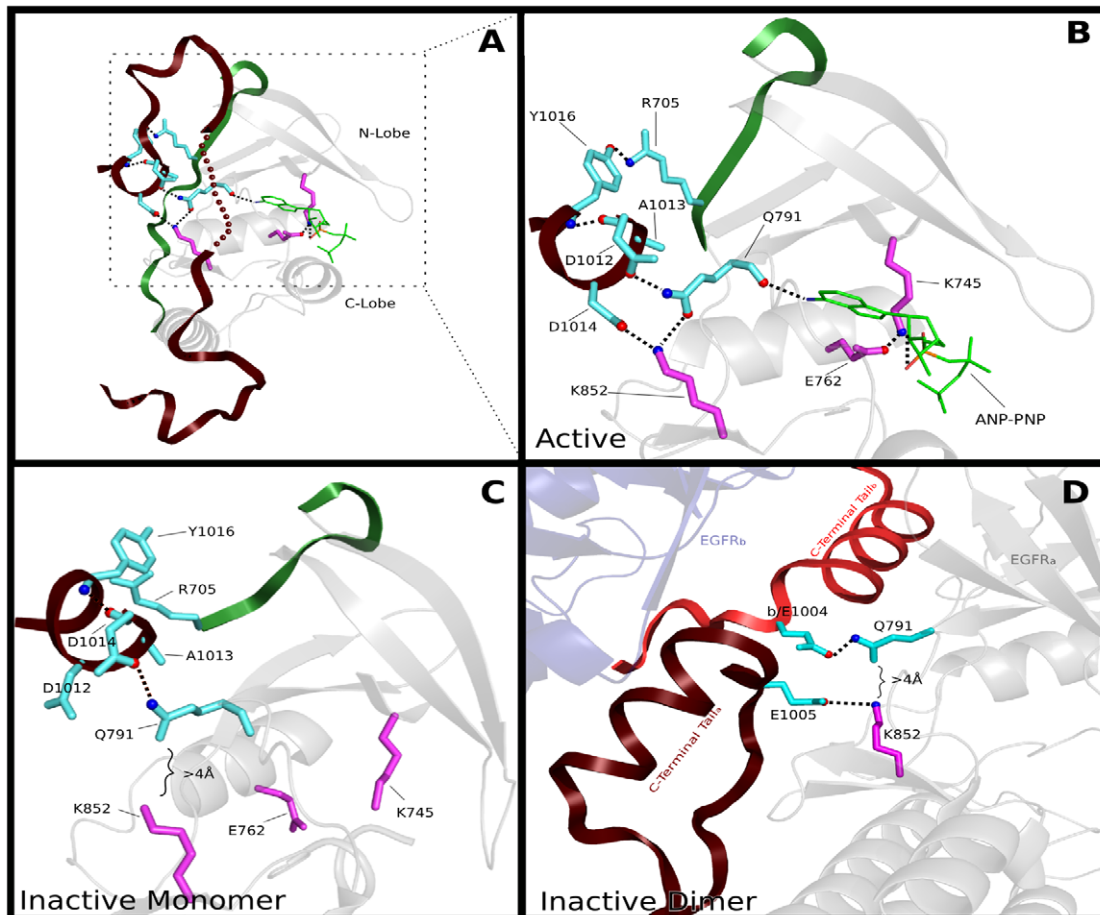

Figure 6. Interactions tethering the C-terminal tail to hinge regions of the kinase core. A) Structural location of the hinge tether. B-D) A close-up view of the tethering interactions in the (B) active dimer (PDB: 2ITN), (C) inactive monomer (PDB: 2RFD), and (D) inactive dimer (PDB: 3GT8). ANP-PNP is shown as green sticks and hydrogen bonds are depicted as black dotted lines. Some hydrogen bonds have been omitted for clarity.
doi:10.1371/journal.pone.0014310.g006

hydrogen bonding to two conserved asparates (D1012 and D1014) in the C-terminal tail (Figure 6B). In EGFR, one of the aspartates (D1014) in the C-terminal tail also hydrogen bonds to a kinase conserved lysine (K852 in the C-lobe), which has been noted to serve as a pivot point for inter-lobe movement [50]. Thus, ErbB kinases have diverged from other RTKs to uniquely tether the C-terminal tail to hinge regions of the kinase domain critical for inter-lobe movement. Why would such tethering be important for ErbB functions? One possibility is that this may provide an additional layer of regulation by allowing the C-terminal tail to internally control inter-lobe movement, and consequently kinase activity. Notably, in the inactive structure of EGFR, where the two lobes are in a closed conformation, the lobe-spanning salt bridge between the glutamine (Q791) and the lysine (K852) is lost, in part because of the movement of C-terminal tail away from the lysine (K852) (Figure 6C) [42]. Also, the C-terminal tails in the inactive dimer [11] prevent the formation of inter-lobe salt bridge by engaging Q791 and K852 in different interactions (Figure 6D).

The control of inter-lobe movement by the C-terminal tail is likely to differ in parasitic worms and sponges because the C-terminal tail residues that interact with the inter-lobe hinge are different in these organisms. Sponges lack the C-terminal tail aspartates (D1012 and D1014), and parasitic worms contain a glutamine at the D1014 position. Notably, both sponges and parasitic worms replace the glutamine (Q791) in the inter-lobe linker by an isoleucine (I) and glutamate (E), respectively (Figure 1). Although the functional implication of this lineage specific variation is unclear, it is likely that EGFR orthologs in sponges and parasitic worms do not require regulation of catalytic activity by the C-terminal tail. We note that ErbB3 differs from other ErbB members in the inter-lobe hinge. In particular, the kinase

conserved lysine (K852), which forms a lobe spanning salt bridge with Q791, is conserved as a glutamine (Q) in ErbB3. This ErbB3-specific variation may contribute to the low levels of kinase activity [17] by preventing opening and closing motion during catalysis.

Interactions coupling the C-terminal tail and the JM segment. As mentioned earlier, some of the strongest ErbB kinase-specific constraints are imposed on residues in the JM-kinase linker and β2-β3 loop. ErbB-conserved residues in these two regions structurally couple a phosphorylatable tyrosine (Y1016) in the C-terminal tail to the JM and N-lobe regions involved in dimerization (Figure 7). Some of the distinctive residues involved in this coupling include, R705 in the JM-kinase linker, W731 in the β2-β3 loop, and E736 in the β2-β3 loop. Specifically, R705 and E736 form hydrogen bonds with the hydroxyl group of Y1016, and W731 provides a favorable docking surface for the aromatic ring of Y1016. These interactions are further coupled to the C-helix and dimerization sites in the N-lobe by I706 and E711 in the JM-kinase linker (Figure 7). Specifically, I706 packs up against hydrophobic residues in the C-helix, and E711 hydrogen bonds to the side-chain of W731, as well as to the backbone of K708, which is involved in the asymmetric dimer interface [10,12] (Figure 7).

The canonical ErbB kinase-conserved interactions between the C-terminal tail, JM-kinase linker, and β2-β3 loop are altered in the inactive dimer [11] (Figure 7). In particular, a conserved C-terminal tail phenylalanine (F997) occupies a position analogous to Y1016 in the inactive state (Figure 7). ErbB kinase-conserved residues in JM-kinase linker and β2-β3 loop also undergo concerted changes in the inactive dimer. In particular, R705, which hydrogen bonds to Y1016 in the active state, moves away to interact with E711 in the JM-kinase linker. Likewise, E736, which
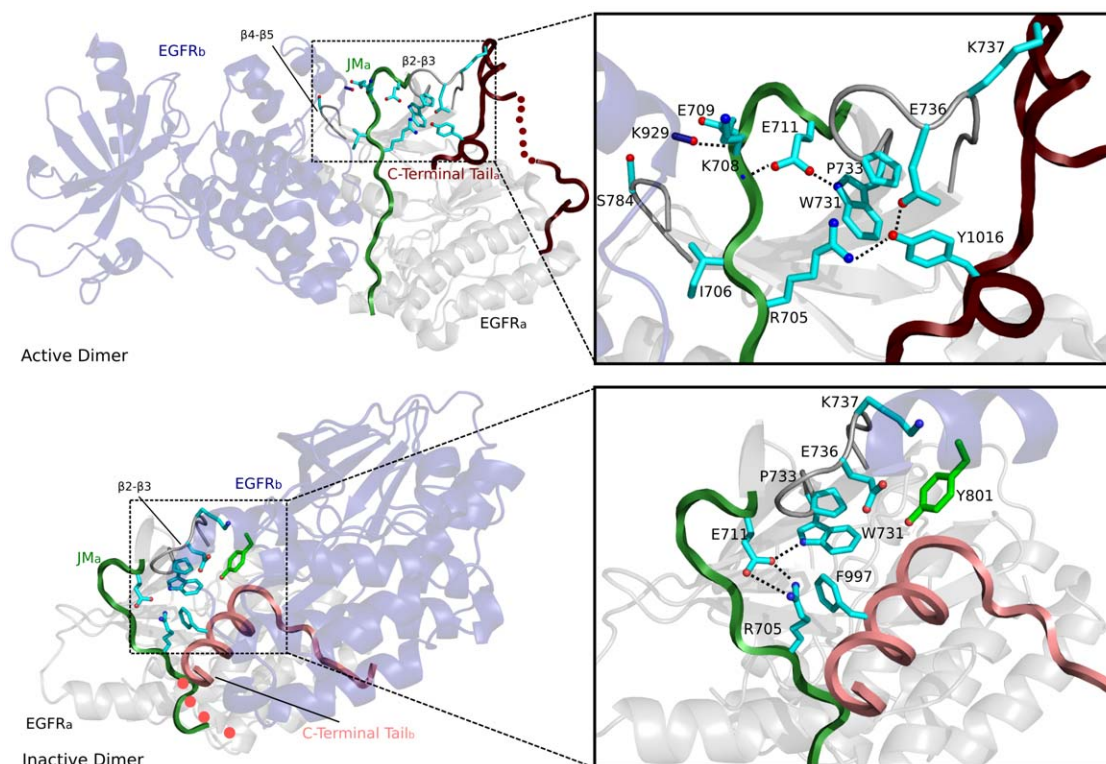

Figure 7. Interactions tethering the C-terminal tail, JM-kinase linker and the β2-β3 loop (labeled as beta2-beta3). The top panel shows structural interactions associated with ErbB kinase-conserved residues in the active asymmetric dimer (PDB: 2ITN) and the bottom panel shows the inactive dimer (PDB: 3GT8). The right panel shows a close-up view of the interactions.
doi:10.1371/journal.pone.0014310.g007

typically interacts with Y1016 in the active state, moves towards a phosphorylatable tyrosine (Y801 in the D-helix) in the inactive state [11]. Together, these concerted changes appear to dynamically couple the C-terminal tail with the JM and C-helix regions involved in dimerization. Notably, a similar coupling between the SH2-kinase linker, β2-β3 loop and SH3 domain have been noted for Src tyrosine kinase [51,52], where the SH3 domain performs a function analogous to the C-terminal tail of EGFR [11].

Our analysis suggests that the structural coupling between the C-terminal tail and JM-kinase linker (described above) is likely to differ in ErbB3, and EGFR orthologs in sponges and parasitic worms because ErbB3 conserves an aspartate at the Y1016 position, and EGFR orthologs in sponges and parasitic worms lack some of the canonical ErbB residues in the JM-kinase linker and β2-β3 loop (Figure 1). Our analysis also suggests that cancer mutations in the β2-β3 loop and the JM-kinase linker (Table 1) may contribute to abnormal regulation by altering the conformational coupling between these two regions.

AST: A hypothetical mechanism for regulating ATP binding by the C-terminal tail

The AST is largely formed by a helical segment (residues 997–1001) in the C-terminal tail (Figure 8A), also referred to as the AP-2 helix [11]. The AST is typically disordered in most ErbB structures; however, in two structures of EGFR (PDB:1XKK and 2JIU) [19,53], the AST segment adopts two distinct conforma-

tions. In one conformation, it protrudes into the ATP binding pocket, thereby tethering the C-terminal tail to the ATP binding site (via two hydrophobic residues, F/L997 and L1001), while in the other conformation the AST swings away from the ATP binding pocket to become solvent-exposed (Figure 8A–B) [19,53]. This mode of dynamically tethering the C-terminal tail to the ATP binding pocket is remarkably similar to PKA (Figure 8C), where a conserved phenylalanine (F327^{PKA}) in the C-terminal tail moves in and out of the ATP binding pocket to serve as a gate for nucleotide binding [23,54]. An analogous role for F997 in EGFR would suggest a similar gating mechanism, wherein nucleotide binding is controlled in a conformation dependent manner by the AST. Such a function may also explain the paradoxical experimental observations, where mutations in the AST both increases [11] and decreases catalytic activity [10,55].

Concluding Remarks

Bayesian analysis of the evolutionary constraints acting on receptor tyrosine kinase sequences has revealed a co-conserved pattern characteristic of the ErbB family. Analysis of this co-conserved pattern, in light of the wealth of structural and functional data available on ErbB kinases, suggests a model in which the identified residues contribute to ErbB kinase functional divergence by providing a structural framework for the JM and C-terminal tail to uniquely regulate ErbB kinase activity (Figure 9). A compelling aspect of this model is that it readily explains the
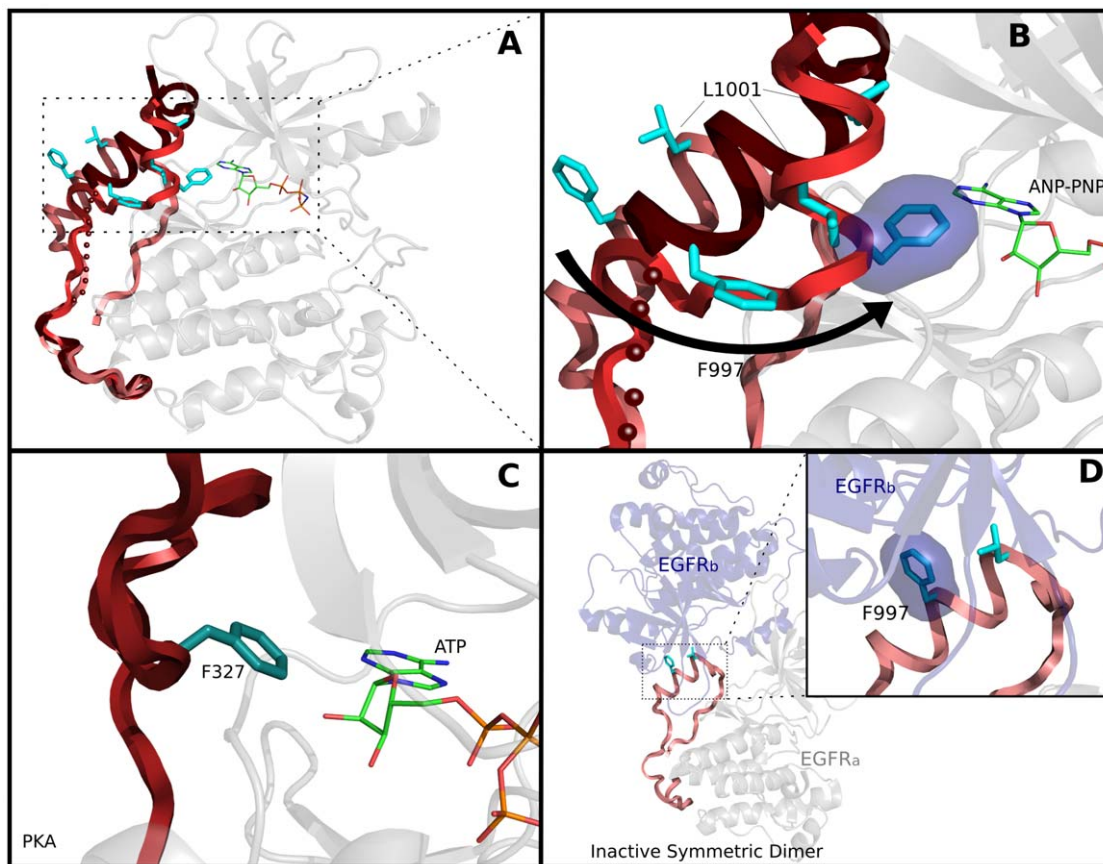

Figure 8. Active site tether (AST) and its role in ATP binding. A) Superposition of the C-terminal tail in the active (PDB: 2JIU), inactive monomer (PDB: 1XKK) and inactive dimer (PDB: 3GT8). B) A close up view of F997 and its proposed role in ATP binding. The modeled ANP-PNP is shown in sticks representation. C) Structural location of F327 (PKA numbering) in the C-terminal tail of PKA (PDB: 1ATP [49]). D) Conformational changes associated with F997 in the inactive symmetric dimer (PDB: 3GT8).
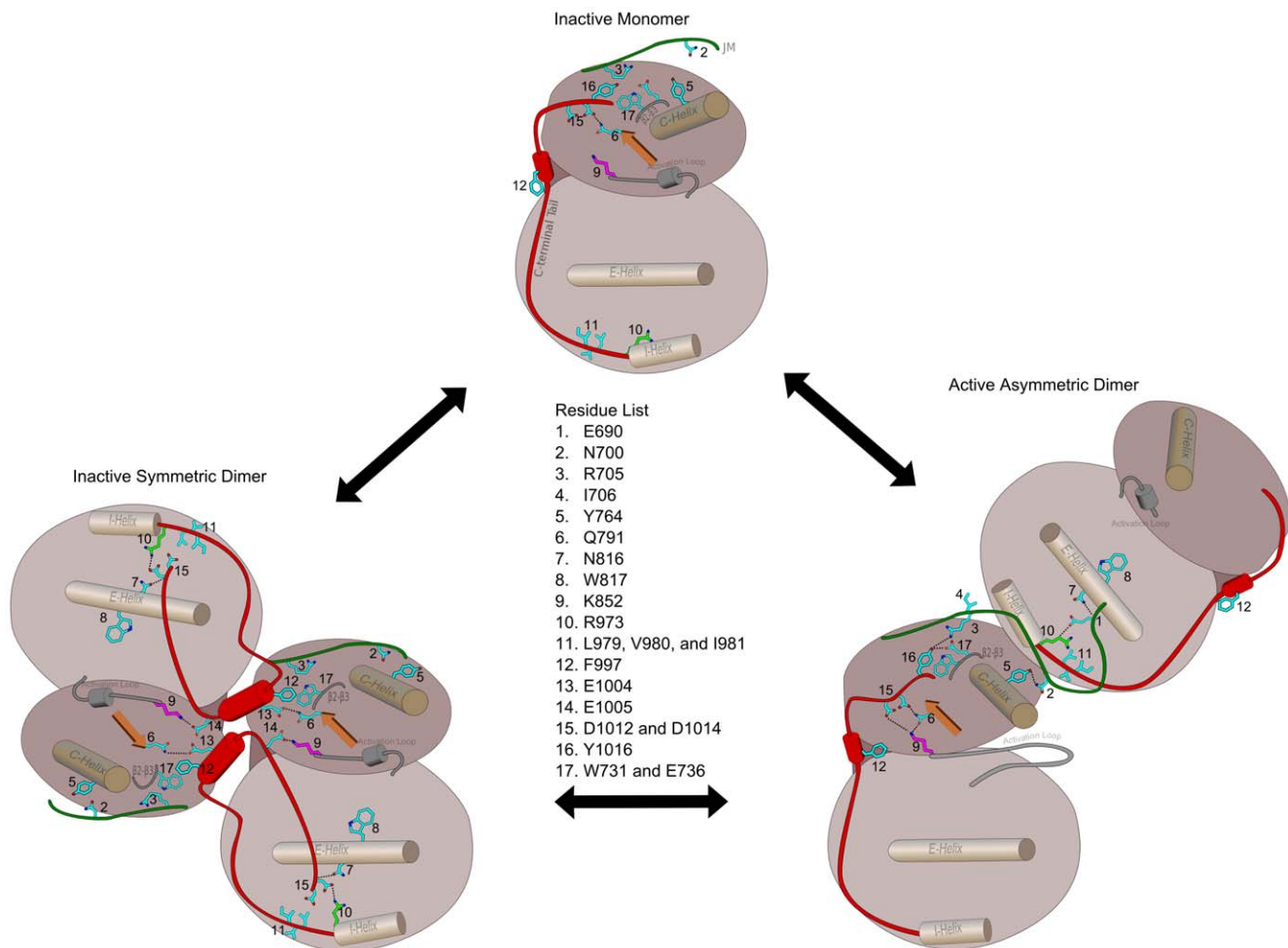doi:10.1371/journal.pone.0014310.g008

Figure 9. Mechanistic model of EGFR kinase activation involving ErbB kinase-conserved residues. Three functional states of EGFR are used (active dimer based upon PDB: 3GOP), inactive dimer (based upon PDB: 3GT8) and inactive monomer (based upon PDB: 1XKK) to illustrate the coordinated functions of the identified residues in kinase activation. The JM and C-terminal tails are colored in green and red, respectively. The residue numbers for the labeled residues are shown. The opening and closing motion of the two lobes are shown to illustrate how the C-tail may regulate catalytic activity in a conformation dependent manner.
doi:10.1371/journal.pone.0014310.g009

inhibitory and activating functions of the C-terminal tail segment, and provides new testable hypotheses for experimental studies. For example, the hypothesis that the ErbB kinase-conserved residues contribute to the *cis* regulation of the kinase core by the JM and C-terminal tail can be tested experimentally. Likewise, the hypothesis that the activation mechanism of sponges and roundworms differs from their mammalian counterparts can also be tested experimentally. Finally, by identifying a potential role for the C-terminal tail in ATP binding (AST), our study provides new avenues for designing selective ErbB kinase inhibitors.

Materials and Methods

Identification of ErbB-specific selective constraints

ErbB and related receptor tyrosine kinase (RTK) sequences from diverse organisms were identified within NCBI nr, env_nr, and translated EST databases using PSI-BLAST and motif-based search procedures. These sequences were multiply aligned using the MAPGAPS program [56] by building curated alignments/profiles for each of the 18 families within the RTK sub-group. These alignments included the Juxtamembrane region, the kinase

domain and ~50 residue segments C-terminal to the kinase domain. The aligned sequences (~3,170 sequences) were partitioned into two functionally divergent subgroups using a Bayesian partitioning with pattern selection (BPPS) procedure [57]. This identified a distinctive pattern that most optimally distinguishes ErbB kinase sequences from other receptor tyrosine kinase sequences (Figure 1). The extent to which these residues contribute to the divergence of ErbB kinases was quantified using a ball-in-urn statistical model [58], and indicated by the height of the histogram above the alignments in Figures 1, 4C, and 8A.

Visualization of ErbB-specific selective constraints

The residues that contribute the most to ErbB kinase evolutionary divergence are shown using a "Contrast Hierarchical Alignment (CHA)" [58] (Figure 1). A CHA is based on three categories of related sequences: (i) a query set, (ii) a foreground set and (iii) a background set. In the Figure 1 alignment, representative ErbB sequences from diverse organisms constitute the query set, all ErbB kinase sequences (419 sequences) correspond to the foreground set, and receptor tyrosine kinases outside of the ErbB family (2,751 sequences) correspond to the background set. The

residues that contribute to ErbB kinase evolutionary divergence, as identified by the BPPS procedure, are shown by block dots above the alignment. Notably, the residues identified by the BPPS procedure are highly conserved in the ErbB family (foreground alignment) and strikingly different in receptor tyrosine kinases outside of the ErbB family (background alignment) (Figure 1).

ErbB3 is considered an atypical member of the ErbB family. To determine to what extent ErbB3 contributes to the divergence of the ErbB family, we ran the BPPS procedure by removing ErbB3 sequences from our alignments. Removing ErbB3 sequences did not significantly alter the pattern-partitions created by the BPPS procedure.

Identification of ErbB3-ErbB4 shared patterns

Representative sequences from ErbB3 and ErbB4 (query set) were multiply aligned against mammalian ErbB kinase sequences (~164 sequences). The BPPS procedure was applied on this alignment to identify sequence patterns that most distinguish ErbB3 and ErbB4 from other ErbB members (ErbB1 and ErbB4). Among other residues, L794 in the C-helix (Figure 4C) was identified as one of the most contributing residues to ErbB3-ErbB4 functional divergence.

Structural analysis of ErbB kinase conserved residues

Crystal structures of ErbB kinases solved in various function states (see below) were obtained from the PDB database (http://www.rcsb.org). Protein hydrogen bonds were added to structural coordinates using the Reduce program [59]. Hydrogen bonds, van der Waals interaction and CH-π interaction were calculated using the CHAIN suite of programs [58]. The identified interactions were further quantified by calculating the frequency of occurrence of each interaction across multiple crystal structures (Table S1). The structural interactions were visualized using PyMOL (http://www.pymol.org). The symmetry related molecules in Figs 3b, 4a, 5b and 7a were generated using the "symmetry mates" utility in PyMOL. The following PDB files were used in our analysis:

Active State: 1M14 [4]; 1M17 [4]; 2EB2(not published); 2EB3(not published); 2GS2 [10]; 2GS6 [10]; 2ITN [60]; 2ITO [60]; 2ITP [60]; 2ITQ [60]; 2ITT [60]; 2ITU [60]; 2ITV [60]; 2ITW [60]; 2ITX [60]; 2ITY [60]; 2ITZ [60]; 2J5E [61]; 2J5F [61]; 2J6M [60]; 2JIT [53]; 2JIU [53]; 3GOP [12]

Inactive State: 1XKK [18]; 2GS7 [10]; 2JIV [53]; 2RF9 [42]; 2RFD [42]; 2RFE [42]; 2RGP [62]; 3BEL [63]; 3GT8 [11]

Retrieval of cancer mutations, modeling and molecular dynamics simulations of the mutant forms

Cancer-associated mutations in EGFR were identified by mining the COSMIC database [64], a repository for somatic mutations in human cancers. A structural model of the lung cancer-associated L861Q mutant was built using two crystal structures of active EGFR kinase domain (PDB ID: 2ITN and 2JIU) as templates. Because all available crystal structures of EGFR have disordered regions, we chose structures with non-overlapping disordered regions (PDB:

2ITN and 2JIU) for our modeling studies. ANP-PNP was modeled in some of the nucleotide unbound structures to show the proximity of ErbB conserved residues to the ATP binding pocket. ANP-PNP was modeled by superimposing the nucleotide unbound structure to the ANP-PNP bound structure of EGFR (PDB: 2ITN). Superposition was done using the "cealign" plugin in Pymol.

For molecular dynamics studies, water molecules, bound inhibitors, and other heteroatom's were removed. The missing residues were modeled using MODELLER [65]. Missing hydrogen and heavy atoms were added using the LEaP program in the Amber software suite [66]. Each protein was solvated with TIP3P water model [67] and counterions were added for neutralization. Molecular dynamics (MD) simulations were done using NAMD software [68], version 2.7b1, and all-atom ff03 force fields from the Amber package. Prior to the regular MD production run, a smoothing function was applied to both the electrostatic and the van der Waals forces at a distance of 10 Å, and a pair list distance of 14 Å with a switching cutoff distance of 12 Å. All bonds with hydrogen were kept rigid by applying the ShakeH algorithm, and the protein backbone atoms were restrained with a harmonic restraint ($k_f = 10$ kcal/(mol · Å^2)). Conjugate-gradient energy minimization was performed on the solvated protein for 10,000 steps, followed by heating from 0 to 298.15 K. The restraints on the protein backbone atoms over multiple stages of equilibration under NPT ensemble (P = 1 atm, T = 298.15 K) were released to obtain a relaxed protein. The unrestrained MD productions were run for 10 ns using a time step of 2 fs and the NPT ensemble. Root-mean-square deviation (RMSD) calculations were performed and monitored to ensure that the simulation was stable during the 10ns time scale. Hydrogen bonding analysis was done using the ptaj program in Amber suite of programs.

Supporting Information

Table S1
Found at: doi:10.1371/journal.pone.0014310.s001 (0.09 MB DOC)

Figure S1
Found at: doi:10.1371/journal.pone.0014310.s002 (0.07 MB DOC)

Author Contributions

Conceived and designed the experiments: NK. Performed the experiments: AM NK. Analyzed the data: AM NK. Contributed reagents/materials/analysis tools: MM ET NK. Wrote the paper: AM NK.

References

1. Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. Science 298: 1912–1934.
2. Yarden Y, Sliwkowski MX (2001) Untangling the ErbB signalling network. Nat Rev Mol Cell Biol 2: 127–137.
3. Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, et al. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science 253: 407–414.
4. 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.
5. Cho H-S, Leahy DJ (2002) Structure of the Extracellular Region of HER3 Reveals an Interdomain Tether. Science 297: 1330–1333.
6. Ferguson KM, Berger MB, Mendrola JM, Cho H-S, Leahy DJ, et al. (2003) EGF Activates Its Receptor by Removing Interactions that Autoinhibit Ectodomain Dimerization. Molecular Cell 11: 507–517.
7. Garrett TPJ, McKern NM, Lou M, Elleman TC, Adams TE, et al. (2002) Crystal Structure of a Truncated Epidermal Growth Factor Receptor Extracellular Domain Bound to Transforming Growth Factor [alpha]. Cell 110: 763–773.
8. Ogiso H, Ishitani R, Nureki O, Fukai S, Yamanaka M, et al. (2002) Crystal Structure of the Complex of Human Epidermal Growth Factor and Receptor Extracellular Domains. Cell 110: 775–787.
9. Schlessinger J (2002) Ligand-Induced, Receptor-Mediated Dimerization and Activation of EGF Receptor. Cell 110: 669–672.

10. 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.

11. Jura N, Endres NF, Engel K, Deindl S, Das R, et al. (2009) Mechanism for Activation of the EGF Receptor Catalytic Domain by the Juxtamembrane Segment. Cell 137: 1293–1307.

12. Red Brewer M, Choi SH, Alvarado D, Moravcevic K, Pozzi A, et al. (2009) The Juxtamembrane Region of the EGF Receptor Functions as an Activation Domain. Molecular Cell 34: 641–651.

13. Schreiber A, Libermann T, Lax I, Yarden Y, Schlessinger J (1983) Biological role of epidermal growth factor-receptor clustering. Investigation with monoclonal anti-receptor antibodies. J Biol Chem 258: 846–853.

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

15. Qiu C, Tarrant MK, Choi SH, Sathyamurthy A, Bose R, et al. (2008) Mechanism of Activation and Inhibition of the HER4/ErbB4 Kinase. Structure 16: 460–467.

16. 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 U S A 106: 21608–21613.

17. Shi F, Telesco SE, Liu Y, Radhakrishnan R, Lemmon MA (2010) ErbB3/HER3 intracellular domain is competent to bind ATP and catalyze autophosphorylation. Proc Natl Acad Sci U S A 107: 7692–7697.

18. Wood ER, Truesdale AT, McDonald OB, Yuan D, Hassell A, et al. (2004) A Unique Structure for Epidermal Growth Factor Receptor Bound to GW572016 (Lapatinib): Relationships among Protein Conformation, Inhibitor Off-Rate, and Receptor Activity in Tumor Cells. Cancer Res 64: 6652–6659.

19. Wood ER, Shewchuk LM, Ellis B, Brignola P, Brashear RL, et al. (2008) 6-Ethynylthieno[3,2-d]- and 6-ethynylthieno[2,3-d]pyrimidin-4-anilines as tunable covalent modifiers of ErbB kinases. Proceedings of the National Academy of Sciences 105: 2773–2778.

20. Yun CH, Boggon TJ, Li Y, Woo MS, Greulich H, et al. (2007) Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. Cancer Cell 11: 217–227.

21. Jeffrey PD, Russo AA, Polyak K, Gibbs E, Hurwitz J, et al. (1995) Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature 376: 313–320.

22. Sicheri F, Moarefi I, Kuriyan J (1997) Crystal structure of the Src family tyrosine kinase Hck. Nature 385: 602–609.

23. Narayana N, Cox S, Xuong N-h, Eyck LFT, Taylor SS (1997) A binary complex of the catalytic subunit of cAMP-dependent protein kinase and adenosine further defines conformational flexibility. Structure 5: 921–935.

24. Huse M, Kuriyan J (2002) The conformational plasticity of protein kinases. Cell 109: 275–282.

25. Ferguson KM (2008) Structure-based view of epidermal growth factor receptor regulation. Annu Rev Biophys 37: 353–373.

26. Cheng HC, Johnson TM, Mills RD, Chong YP, Chan KC, et al. (2010) Allosteric networks governing regulation and catalysis of Src-family protein tyrosine kinases: implications for disease-associated kinases. Clin Exp Pharmacol Physiol 37: 93–101.

27. Landau M, Fleishman SJ, Ben-Tal N (2004) A Putative Mechanism for Downregulation of the Catalytic Activity of the EGF Receptor via Direct Contact between Its Kinase and C-Terminal Domains. Structure 12: 2265–2275.

28. Walton GM, Chen WS, Rosenfeld MG, Gill GN (1990) Analysis of deletions of the carboxyl terminus of the epidermal growth factor receptor reveals self-phosphorylation at tyrosine 992 and enhanced in vivo tyrosine phosphorylation of cell substrates. J Biol Chem 265: 1750–1754.

29. Mol CD, Dougan DR, Schneider TR, Skene RJ, Kraus ML, et al. (2004) Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. J Biol Chem 279: 31655–31663.

30. Wybenga-Groot LE, Baskin B, Ong SH, Tong J, Pawson T, et al. (2001) Structural basis for autoinhibition of the Ephb2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region. Cell 106: 745–757.

31. Shewchuk LM, Hassell AM, Ellis B, Holmes WD, Davis R, et al. (2000) Structure of the Tie2 RTK domain: self-inhibition by the nucleotide binding loop, activation loop, and C-terminal tail. Structure 8: 1105–1113.

32. Kannan N, Neuwald AF (2004) Evolutionary constraints associated with functional specificity of the CMGC protein kinases MAPK, CDK, GSK, SRPK, DYRK, and CK2alpha. Protein Sci 13: 2059–2077.

33. Kannan N, Neuwald AF (2005) Did protein kinase regulatory mechanisms evolve through elaboration of a simple structural component? J Mol Biol 351: 956–972.

34. Kannan N, Haste N, Taylor SS, Neuwald AF (2007) The hallmark of AGC kinase functional divergence is its C-terminal tail, a cis-acting regulatory module. Proceedings of the National Academy of Sciences 104: 1272–1277.

35. Gould CM, Kannan N, Taylor SS, Newton AC (2009) The Chaperones Hsp90 and Cdc37 Mediate the Maturation and Stabilization of Protein Kinase C through a Conserved PXXP Motif in the C-terminal Tail. J Biol Chem 284: 4921–4935.

36. Robert ARomano, NatarajanKannan, Alexandr PKornev, Craig JAllison, Susan STaylor (2009) A chimeric mechanism for polyvalent trans-phosphorylation of PKA by PDK1. Protein Science 18: 1486–1497.

37. Hubbard SR (2004) Juxtamembrane autoinhibition in receptor tyrosine kinases. Nat Rev Mol Cell Biol 5: 464–471.

38. Hubbard SR, Wei L, Ellis L, Hendrickson WA (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. Nature 372: 746–754.

39. Schaefer G, Akita RW, Sliwkowski MX (1999) A Discrete Three-amino Acid Segment (LVI) at the C-terminal End of Kinase-impaired ErbB3 Is Required for Transactivation of ErbB2. J Biol Chem 274: 859–866.

40. Yang KS, Macdonald-Obermann JL, Piwnica-Worms D, Pike LJ (2010) Asp-960/Glu-961 controls the movement of the C-terminal tail of the epidermal growth factor receptor to regulate asymmetric dimer formation. J Biol Chem 285: 24014–24022.

41. Wu J, Yang J, Kannan N, Madhusudan, Xuong NH, et al. (2005) Crystal structure of the E230Q mutant of cAMP-dependent protein kinase reveals an unexpected apoenzyme conformation and an extended N-terminal A helix. Protein Sci 14: 2871–2879.

42. Zhang X, Pickin KA, Bose R, Jura N, Cole PA, et al. (2007) Inhibition of the EGF receptor by binding of MIG6 to an activating kinase domain interface. Nature 450: 741–744.

43. Walker F, Hibbs ML, Zhang HH, Gonez LJ, Burgess AW (1998) Biochemical characterization of mutant EGF receptors expressed in the hemopoietic cell line BaF/3. Growth Factors 16: 53–67.

44. Chen YR, Fu YN, Lin CH, Yang ST, Hu SF, et al. (2006) Distinctive activation patterns in constitutively active and gefitinib-sensitive EGFR mutants. Oncogene 25: 1205–1215.

45. Feinmesser RL, Wicks SJ, Taverner CJ, Chantry A (1999) Ca2+/calmodulin-dependent kinase II phosphorylates the epidermal growth factor receptor on multiple sites in the cytoplasmic tail and serine 744 within the kinase domain to regulate signal generation. J Biol Chem 274: 16168–16173.

46. Landau M, Ben-Tal N (2008) Dynamic equilibrium between multiple active and inactive conformations explains regulation and oncogenic mutations in ErbB receptors. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer 1785: 12–31.

47. Kumar A, Petri ET, Halmos B, Boggon TJ (2008) Structure and clinical relevance of the epidermal growth factor receptor in human cancer. J Clin Oncol 26: 1742–1751.

48. Papakyriakou A, Vourloumis D, Tzortzatou-Stathopoulou F, Karpusas M (2009) Conformational dynamics of the EGFR kinase domain reveals structural features involved in activation. Proteins 76: 375–386.

49. Padyana AK, Qiu H, Roll-Mecak A, Hinnebusch AG, Burley SK (2005) Structural Basis for Autoinhibition and Mutational Activation of Eukaryotic Initiation Factor 2{alpha} Protein Kinase GCN2. J Biol Chem 280: 29289–29299.

50. Lamers MBAC, Antson AA, Hubbard RE, Scott RK, Williams DH (1999) Structure of the protein tyrosine kinase domain of C-terminal Src kinase (CSK) in complex with staurosporine. Journal of Molecular Biology 285: 713–725.

51. Gonfloni S, Williams JC, Hattula K, Weijland A, Wierenga RK, et al. (1997) The role of the linker between the SH2 domain and catalytic domain in the regulation and function of Src. EMBO J 16: 7261–7271.

52. Young MA, Gonfloni S, Superti-Furga G, Roux B, Kuriyan J (2001) Dynamic coupling between the SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation. Cell 105: 115–126.

53. Yun C-H, Mengwasser KE, Toms AV, Woo MS, Greulich H, et al. (2008) The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. Proceedings of the National Academy of Sciences 105: 2070–2075.

54. Yang J, Kennedy EJ, Wu J, Deal MS, Pennypacker J, et al. (2009) Contribution of non-catalytic core residues to activity and regulation in protein kinase A. J Biol Chem 284: 6241–6248.

55. Xia L, Wang L, Chung AS, Ivanov SS, Ling MY, et al. (2002) Identification of Both Positive and Negative Domains within the Epidermal Growth Factor Receptor COOH-terminal Region for Signal Transducer and Activator of Transcription (STAT) Activation. J Biol Chem 277: 30716–30723.

56. Neuwald AF (2009) Rapid detection, classification and accurate alignment of up to a million or more related protein sequences. Bioinformatics 25: 1869–1875.

57. Neuwald AF, Kannan N, Poleksic A, Hata N, Liu JS (2003) Ran's C-terminal, basic patch, and nucleotide exchange mechanisms in light of a canonical structure for Rab, Rho, Ras, and Ran GTPases. Genome Res 13: 673–692.

58. Neuwald AF (2007) The CHAIN program: forging evolutionary links to underlying mechanisms. Trends in Biochemical Sciences 32: 487–493.

59. Word JM, Lovell SC, Richardson JS, Richardson DC (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J Mol Biol 285: 1735–1747.

60. Yun C-H, Boggon TJ, Li Y, Woo MS, Greulich H, et al. (2007) Structures of Lung Cancer-Derived EGFR Mutants and Inhibitor Complexes: Mechanism of Activation and Insights into Differential Inhibitor Sensitivity. Cancer Cell 11: 217–227.

61. Blair JA, Rauh D, Kung C, Yun C-H, Fan Q-W, et al. (2007) Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. Nat Chem Biol 3: 229–238.

62. Xu G, Abad MC, Connolly PJ, Neeper MP, Struble GT, et al. (2008) 4-Amino-6-arylamino-pyrimidine-5-carbaldehyde hydrazones as potent ErbB-2/EGFR

dual kinase inhibitors. Bioorganic & Medicinal Chemistry Letters 18: 4615–4619.

63. Xu G, Searle LL, Hughes TV, Beck AK, Connolly PJ, et al. (2008) Discovery of novel 4-amino-6-arylaminopyrimidine-5-carbaldehyde oximes as dual inhibitors of EGFR and ErbB-2 protein tyrosine kinases. Bioorganic & Medicinal Chemistry Letters 18: 3495–3499.

64. Bamford S, Dawson E, Forbes S, Clements J, Pettett R, et al. (2004) The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. Br J Cancer 91: 355–358.

65. Sali A (1995) Comparative protein modeling by satisfaction of spatial restraints. Mol Med Today 1: 270–277.

66. Case DA, Darden TA, Cheatham TEI, Simmerling CL, Wang J, et al. (2008) AMBER 10. San Francisco, California, USA: University of California.

67. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys 79: 926–935.

68. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, et al. (2005) Scalable Molecular Dynamics with NAMD. J Comput Chem 26: 1781–1802.