ABSTRACT

AMAR MIRZA
Analysis of the evolutionary constraints distinguishing ErbB tyrosine kinases suggests a *cis* regulatory function for the Juxtamembrane and COOH-terminal tail segments
(Under the Direction of Dr. Natarajan Kannan)

The Epidermal Growth Factor Receptor (EGFR) and related ErbB kinases have functionally diverged from other receptor and non-receptor tyrosine kinases to be activated by a unique allosteric mechanism, in which two kinase domains activate each other through coordinated interactions between the catalytic core and the flanking juxtamembrane (JM) and COOH-terminal tail (C-terminal tail) segments. To identify key residues that contribute to the functional divergence of the ErbB family, a Bayseian approach was used to infer the selective constraints distinguishing ErbB from non-ErbB tyrosine kinases. Strong ErbB-specific constraints are imposed on residues that tether the JM and C-terminal tail segments to key functional regions of the kinase core. The strongest constraint is imposed on a glutamine (Q791), which tethers the C-terminal tail to a conserved inter-lobe salt bridge that is critical for inter-lobe movement. The malleability of this inter-lobe tether in the open and closed states of EGFR suggests a mechanism by which inter-lobe movement and ATP binding can be coordinated with conformational changes in the C-terminal tail. Likewise, an ErbB specific [YL]xx[AG] motif in the C-helix appears to play an active role in coupling JM and C-helix movement. We propose that the JM and C-terminal tail segments flanking the ErbB kinase core are *cis* elements that that have co-evolved with the kinase core to tightly regulate activity. Such a view readily explains the activating and inhibitory roles of the JM and C-terminal tail segments in EGFR functions, and provides new hypotheses for experimental studies.

INDEX WORDS: Epidermal Growth Factor Receptor, C-Terminal Tail, evolution, CHAIN Program, *Cis* regulation, Juxtamembrane region

ANALYSIS OF THE EVOLUTIONARY CONSTRAINTS DISTINGUISHING ERBB TYROSINE KINASES SUGGESTS A CIS REGULATORY FUNCTION FOR THE JUXTAMEMBRANE AND COOH-TERMINAL TAIL SEGMENTS

by

AMAR MIRZA

A Thesis Submitted to the Honors Council of the University of Georgia in Partial Fulfillment of the Requirements for the Degree

BACHELOR OF SCIENCE

in BIOCHEMISTRY AND MOLECULAR BIOLOGY

with HIGHEST HONORS

and CURO SCHOLAR DISTINCTION

Athens, Georgia

2009

ANALYSIS OF THE EVOLUTIONARY CONSTRAINTS DISTINGUISHING ERBB TYROSINE KINASES SUGGESTS A CIS REGULATORY FUNCTION FOR THE JUXTAMEMBRANE AND COOH-TERMINAL TAIL SEGMENTS

by

AMAR MIRZA

Approved:

Natarajan Kannan 12/14/2009
Dr. Natarajan Kannan Date
Faculty Research Mentor

Approved:

Morad Mustafa 12/14/2009
Dr. Morad Mustafa Date
Reader

Approved:

David Williams 12/16/2009
Dr. David S. Williams Date
Director, Honors Program, Foundation Fellows and
Center for Undergraduate Research Opportunities

Approved:

Pamela Kleiber 12/16/2009
Dr. Pamela B. Kleiber Date
Associate Director, Honors Program and
Center for Undergraduate Research Opportunities

ACKNOWLEDGEMENTS

I would like to first acknowledge my research mentor Dr. Natarajan Kannan along with the other co-authors of this manuscript Dr. Morad Mustafa, and Eric Talevich. I also acknowledge the CURO Summer Research Fellowship for funding to conduct this study (Summer, 2009).

TABLE OF CONTENTS

LIST OF FIGURES

Page

CHAPTER 1
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 major signaling pathways associated with cell migration, proliferation, and differentiation. ErbBs consist of an extracellular ligand binding domain and an intracellular kinase domain, which is flanked on the N and C-terminal ends by the juxtamembrane (JM) segment and COOH-terminal tail (C-terminal tail), respectively. Because overexpression or abnormal regulation of ErbB kinase activity is implicated in many human cancers (2), extensive studies have been performed on ErbB kinases to understand their mechanisms of action. Studies on the extracellular ligand binding domain of EGFR(3-7) and more recently on the intracellular kinase domains have provided key insights into how EGFR converts an extracellular ligand binding signal to an intracellular response(8-12). Ligand binding to the extracellular receptor domain induces dimerization and activation of the intracellular kinase domain, which, upon activation, autophosphorylates conserved tyrosine residues in the C-terminal tail (13, 14). Autophoshorylation of the tyrosine residues activates downstream signaling pathways by recruiting signaling and docking proteins to the C-terminal tail (2). While this mechanism of signal propagation is generally believed to be conserved within the ErbB family (9), ErbB3 diverges from this canonical mechanism of action because of its inability to catalyze phosphoryl transfer. ErbB3, however, is important for normal cellular functions inasmuch as it *trans* activates other ErbB members by forming heterodimers (15-18).

A key step in ErbB kinase signaling is the activation of the intracellular kinase domain. In EGFR, this is achieved by an allosteric interaction between two kinase domains in an asymmetric dimer. In the asymmetric dimer, the C-lobe of one kinase (activator) functions as an activator by stabilizing the other (receiver) in an active conformation. This asymmetric dimer formation not only involves the kinase core, but also the JM and C-terminal tail flanking the kinase core. In particular, the JM segment plays an active role in the asymmetric dimer formation by docking to the C-lobe of the activating kinase(8). Likewise, the C-terminal tail, in the inactive dimer, prevents the asymmetric dimer formation by shielding the JM docking surface on the C-lobe of the activator(11).

Activation of the EGFR kinase domain by dimerization involves conformational changes in key functional elements of the kinase core. One such functional element is the flexible C-helix, which switches from an inactive "out" to an active "in" conformation, upon formation of the asymmetric dimer (8). Docking of the C-helix to the C-lobe of the activating monomer facilitates this conformational switch and repositions a kinase conserved glutamate to coordinate with a lysine, which is involved in catalysis. This mode of regulating catalytic activity by controlling the conformation of C-helix is also observed in other protein kinases such as Cdk2 (19), Src (20) and PKA (21).

In addition to the precise positioning of the regulatory C-helix, activation of EGFR and ErbB4 kinases also involves movement of the N-terminal ATP binding lobe relative to the C-lobe, and repositioning of the activation loop from an inactive to active conformation (9, 22-28). Presumably these conformational changes are also tightly controlled in ErbB kinases to avoid physiological catastrophes regulate catalytic activity. Although previous crystallographic and computational studies(28, 29) have provided some insights into how these conformational

changes might be regulated, the precise mechanisms by which inter-lobe movement and activation loop movement is coupled to ErbB kinase activation is still not fully understood. Experimental delineating these mechanisms requires, however, that one first formulate the right hypotheses based on available information on ErbB kinases. One valuable source of information in this regard is the evolutionary constraints distinguishing ErbB kinases from other functionally distinct tyrosine kinases.

We have shown using several case studies that Bayesian analysis of the evolutionary constraints distinguishing functionally divergent protein kinases is a viable approach for identifying structural elements that contribute to the allosteric specificity of kinases in signaling pathways (29-32). Using this approach, we recently demonstrated that the conserved C-terminal tail, which wraps around the AGC kinase core, is a cis regulatory module that can allosterically regulate AGC kinase activity (30). This study also allowed us to formulate and test new hypotheses regarding the role of conserved motifs in the C-terminal. In particular, the role of a conserved proline rich (PXXP) motif in the maturation of PKC, and the role of a conserved FDxY motif in PDK1 binding(33) was demonstrated using a combination of mutational analysis and peptide array studies.

In this study, we compare the functional constraints distinguishing ErbB kinases from other tyrosine kinases to identify key residues/motifs that contribute to ErbB kinase functional divergence. Our analysis reveals that strong ErbB kinase specific constraints are imposed on residues that are involved in tethering the JM and the C-terminal tail segments 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 family suggests a model in which the C-terminal tail and JM region can function as *cis* regulators by controlling inter-lobe and C-helix

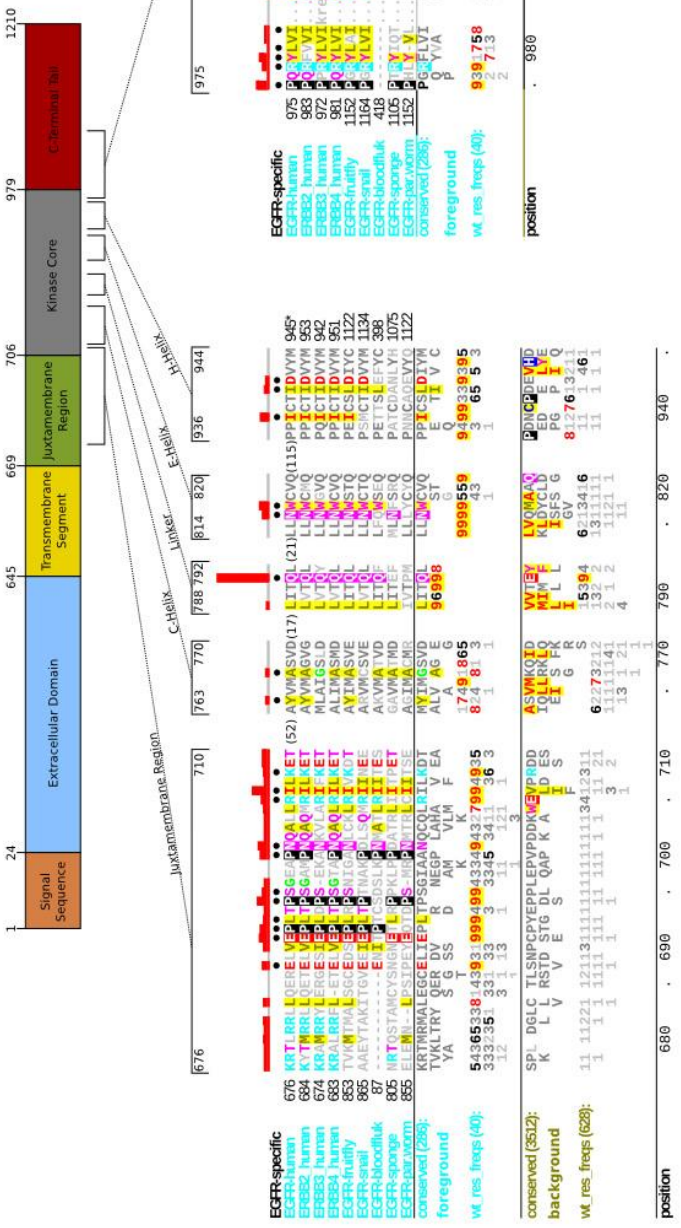
movement, respectively. Such a model readily explains the activating and inhibitory functions observed for the JM and C-terminal segments, and provides mechanistic clues to a frequently occurring lung cancer mutation (L861Q), which appears to activate EGFR by altering the canonical tethering interactions between the JM segment and C-helix.

CHAPTER 2
RESULTS AND DISCUSSION

A co-conserved sequence pattern characteristic of the ErbB family

To identify which sequence features most distinguish ErbB kinases from other receptor and non-receptor tyrosine kinases, we performed statistical analysis of the selective constraints distinguishing ErbB kinases from other tyrosine kinases (see Methods). These constraints generally correspond to residues that are highly conserved in ErbB kinases but strikingly different in tyrosine kinases outside of the ErbB family. Within the catalytic core, these residues correspond to Y/L764 and A/G767 in the C-helix ([YL]xx[AG] motif), Q791 in the inter-lobe linker, N816 and W817 in the E-helix, and I938, IL941 and D942 in the I-helix (Fig 1). These ErbB-specific residues although widely dispersed in sequence spatially interact with conserved motifs in the JM segment and C-terminal tail (Fig 2). To understand why nearly all the ErbB residues interact with the JM and C-terminal tail, we compared the selective constraints imposed on these flanking segments in ErbB and other receptor tyrosine kinases (see Methods). This revealed that the JM and C-terminal tails also conserve sequence motifs that are unique to the ErbB family(Fig 1). Some of the characteristic motifs in the C-terminal tail are PxRYLVI and [DE][AP][DE][ED]Y motifs that are tethered to the C and N-lobe of the kinase core, respectively (Fig 1,2). The JM segment, likewise, contains key sequence motifs that are tethered to the N and C-lobes of the kinase in the asymmetric dimer (Fig 1; Fig 2 and Fig3).

Figure 1: Contrast Hierarchical Alignment showing sequence patterns that distinguish ErbB kinases (foreground) from other tyrosine kinases (background). The residues that contribute to the functional divergence 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 set (ErbB sequences) contrast with residues observed at the corresponding position in the background set (other tyrosine kinase sequences). The secondary structural location of these constraints relative to the overall domain organization of the ErbB family is shown above the alignment. The numbering using in the alignment and in the text is according to the pre-mature EGFR numbering scheme which includes the 24 amino acid signaling sequence. 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-snail : 146217140; EGFR-bloodfluke : 76156675; EGFR-sponge : 18146642; EGFR-par.worm : 170592919



The ErbB prototypic features are generally well conserved across diverse eukaryotic phyla. However, some lower eukaryotes and parasitic organisms diverge from the canonical ErbB features in interesting ways. EGFR orthologues in sponges and parasitic roundworms, for example, lack the canonical C-terminal tail

Figure 1

[DE][AP][DE][ED]Y motif, which is generally tethered to the N-lobe of the kinase through hydrogen bonding interactions with an ErbB conserved glutamine (Q791) in the inter-lobe linker region (Fig 1,2).The concomitant mutation of Q791 to a glutamate and proline in organisms that lack the [DE][AP][DE][ED]Y motif, suggests a co-evolution of the C-terminal tail with the kinase core (Fig 1). Interestingly, *Schistosoma japonicum*, a blood fluke which infects humans and a host of other mammals, conserve Q791 in the inter-lobe linker region, but lack the [DE][AP][DE][ED]Y motif in the C-terminal tail. The structural and functional implication of this variation is discussed in the later sections.

Structural analysis of ErbB-specific sequence patterns

As mentioned above, nearly all the ErbB-specific residues are involved in tethering the JM and C-terminal tail segments to key functional regions of the kinase core (Fig 2). These residues can be broadly classified into four categories based on their structural location (i) Inter-Lobe Tether (ILT): residues that tether the C-terminal tail to the inter-lobe linker and N-lobe (ii) Active Site Tether (AST): residues that tether the C-terminal tail to the ATP binding pocket (iii) C-Helix Tether (CHT): residues that tether the JM segment to the regulatory C-helix, and (iv) C-Lobe Tether (CLT): residues that tether the JM and C-terminal tail to the C-lobe (Fig 2).

The ILT, AST, CHT and CLT residues are highly conserved in ErbB kinases, but strikingly different in non-ErbB kinases. This selective conservation is likely to be important for ErbB-specific functions, rather than for maintaining ErbB kinase structure or fold, because non-ErbB kinases that lack these residues essentially adopt the samefold. Indeed, recent studies on the activation mechanism of EGFR kinases confirm the functional importance of some of the residues identified in our analysis. In particular, the NW motif in the E-helix was shown to play an important role in ErbB functions by providing a malleable docking surface for the JM and C-

terminal tail in the active (11, 12)and inactive states (11), respectively(Fig 3A-B)(11). Likewise, the C-terminal tail PxRYxVI motif, which is part of the CLT, was shown to be important for Erb3-Erb2 hetero-dimerization(10, 34), and asymmetric dimer formation in EGFR(11)(Fig 3C).

Whereas the role of CLT residues is well understood, little is known about the role of the

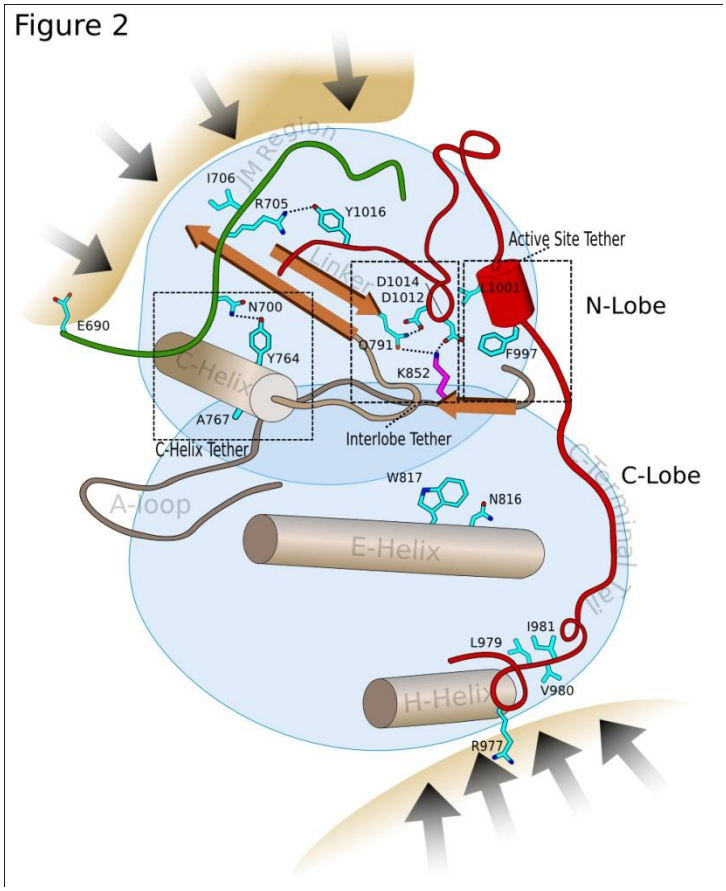


Figure 2: A schematic of the ErbB kinase domain showing the distinctive feature of ErbB kinases. Four major tethering interactions involving the kinase core, the JM segment and C-terminal tails are shown. These four regions correspond to (i) Inter-lobe tether (ILT) (ii) Active site tether (AST) (iii) C-helix tether (CHT) and (iv) C-lobe tether (CLT). The residues that are characteristic of the ErbB family are shown in sticks representation and are colored in cyan. Kinase conserved residues are colored in magenta. The sites of homo/heterodimerization are shown by dark arrows.

ILT and AST residues in ErbB functions. The role of some of the CHT residues have recently been noted (11, 12), but others such as the [YL]xx[AG] motif in the C-helix has not been pointed out before.To better understand the role of these residues in ErbB functions, we analyzed the structural interactions associated with the ILT, AST and CHT residues, as described in the sections below.

Figure 3

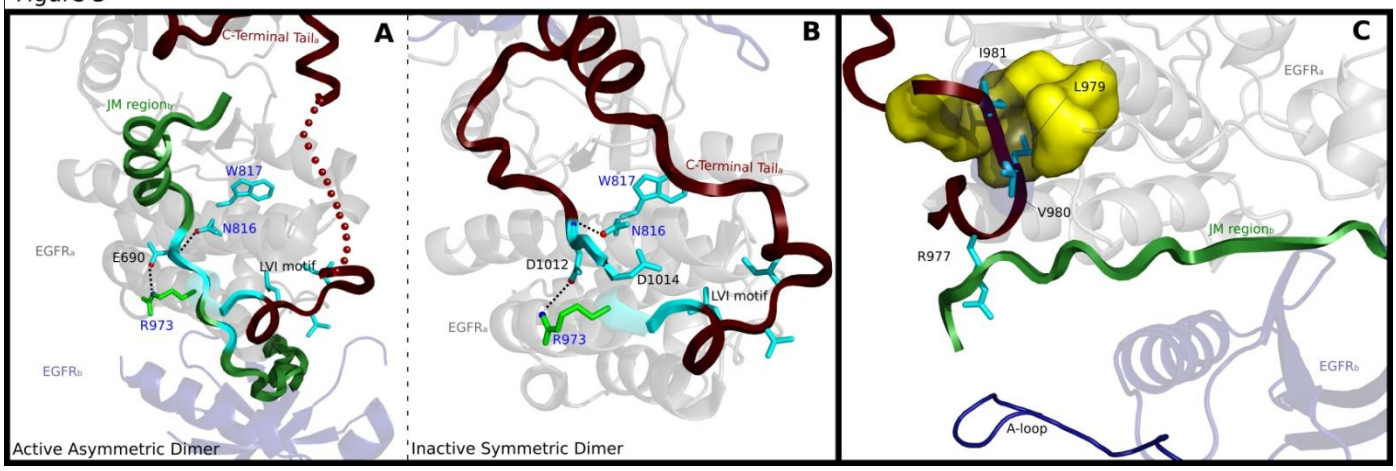


Figure 3. Role of CLT residues in tethering the JM and C-terminal tails. A) Tethering of the JM segment to the C-lobe in the active state. B) Tethering of the C-terminal tail to the C-lobe in the inactive state C) Role of the PxRYLVI motif in tethering the C-terminal tail to the C-lobe. The secondary structural regions are colored as mentioned above. The disordered C-terminal tail in the active state (Fig 3A) is shown in dotted representation. Note that R973 is colored in green in Figures A and B because this residue is shared by some tyrosine kinases outside of the ErbB family.

Structural interactions associated with the ILT: A hypothetical mechanism for coupling C-terminal tail and inter-lobe movements

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 [Taylor;chem. review]. This inter-lobe movement in eukaryotic protein kinases is facilitated by the inter-lobe linker, and other lobe spanning salt bridge interactions that serve as hinge points for domain movments. In particular, a malleable salt bride formed between a kinase conserved glutamate (Q791 in Fig 2) in the N-lobe, and a kinase conserved lysine (K852 in Fig 2) in the C-lobe is known to serve as a pivot for inter-lobe movement.

In ErbB kinases, this pivot is directly tethered to the C-terminal tail through ILT interactions. Specifically, the glutamate, which forms a salt bridge with K852, is conserved in ErbBs as a glutamine (Q791), which by the virtue of having a donor group in its side-chain is able to form a hydrogen bond with the conserved aspartates (D1012 and D1014) in the C-

terminal tail (Fig 4). One of the aspartates (D1014) also hydrogen bonds to K852 in the C-lobe. Together this ErbB specific structural arrangement appears to directly tether the C-terminal tail to the hinge point critical for inter-lobe movement. Why would such an arrangement be important for ErbB functions? One possibility is that this arrangement may facilitate an additional layer of regulation by allowing the C-terminal tail to internally control inter-lobe movement and consequently kinase activity. Consistent with this view, in the inactive structure of EGFR, where the two lobes are in a closed conformation, the canonical salt bridge between the glutamine and the lysine is lost, due, in part, to the movement of C-terminal tail away from the lysine (K852) (Fig 4C). This concomitant change in the C-terminal tail conformation upon closing of the two lobes suggests an active role of the C-terminal tail in inter-lobe movement.

The inter-lobe salt bridge between Q791 and K852 is also lost in the recently determined inactive structure of EGFR, where one monomer inactivates the other monomer by forming a symmetric dimer(11). However, unlike the inactive monomer, where the inter-lobe salt bridge is still tethered to the [DE][AP][DE][ED]Y motif in the C-terminal tail (Fig 4C), in the inactive dimer, Q791 and K852 are engaged in strikingly different interactions with the C-terminal tail (Fig 4D). Specifically, negatively charged residues from the C-terminal tail of the two monomers (E1005 in monomer A and E1004 in monomer B) appear to prevent the formation of the inter-lobe salt bridge between Q791 and K852 by engaging them in different interactions(11) (Fig 4C). Presumably this alternative arrangement facilitates the stabilization of the closed ATP "inaccessible state" of the kinase (11).

The control of inter-lobe movement by the C-terminal tail is likely to differ between metazoan and non-metazoan EGFRs since non-metazoan lack some of the canonical residues involved in the ILT. EGFR orthologue in *Schistosoma japonicum*, for instance, conserve a glutamine at position 791, but lack the canonical motifs in the C-terminal tail (Fig 1). The functional implication of this *S. japonicum* specific change is unclear, but it is likely that *S. japonicum* is either constitutively active or it borrows the C-terminal tail from the host for its functions.

Figure 4

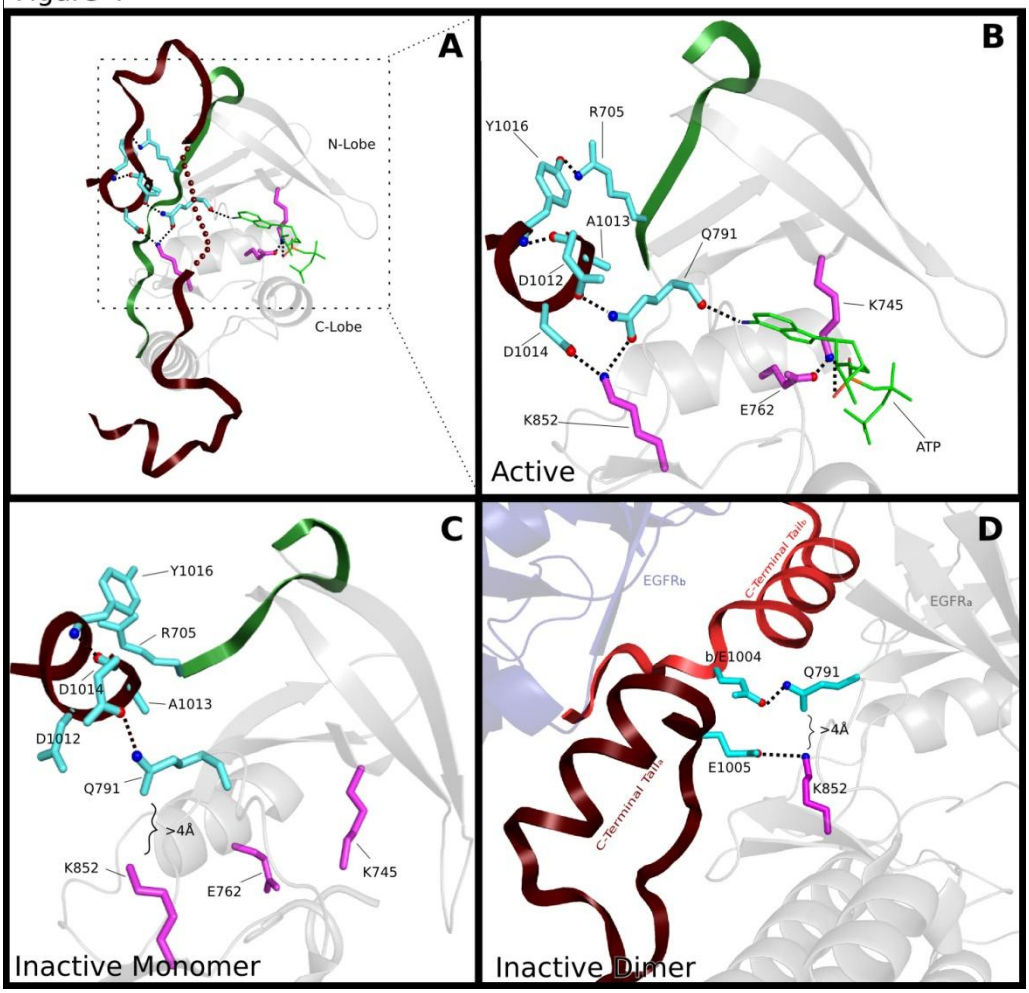


Figure 4: Inter-lobe tether (ILT) interactions showing the modulation of the inter-lobe salt-bridge (between Q791 and K852) by the C-terminal tail. A) Structural location of these interactions relative to the N and C-terminal tails B-D) A close-up view of these interactions in the (B) active dimer (pdb: 2itn(25)) (C) inactive monomer (pdb: 2rfd(45)) (D) inactive dimer

(pdb: 3gt8(11)). The structure images were generated using Pymol (www.pymol.org). The ErbB specific residues are colored cyan, kinase catalytic residues are magenta, and ATP is shown as black sticks. Hydrogen bonds are depicted as black dotted lines, and disordered regions are shown as dots. Some hydrogen bonds have been omitted for clarity.

AST may serve as a gate for ATP binding

The AST is largely formed by a helical segment (residues 997-1001) in the C-terminal tail (Fig 5A), 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) (26, 35), this segment adopts two distinct conformations. In one of the conformation, it protrudes into the ATP binding pocket, thereby tethering the C-terminal tail to the ATP binding pocket (via two hydrophobic residues; L1001 and F997), while in the other conformation this segment swings away from the ATP binding pocket and becomes solvent-exposed (Fig 5A-B). This mode of dynamically tethering the C-terminal tail to the ATP binding pocket is remarkably similar to PKA (Fig 5C), where a conserved phenylalanine (F327) in the C-terminal tail moves in and out of the ATP binding pocket to serve as a gate for nucleotide binding (36). An analogous role for F997 in EGFR would suggest a similar gating mechanism in which the AST can directly control nucleotide binding. Previous studies had suggested an autoinhibitory role for AST since mutation or deletion in the AST region increased catalytic activity (8, 37). However, such a role fails to explain the recently observed impaired kinase activity upon mutations in the AST region (11). We propose a *cis* regulatory role for AST in which the AST can serve both as an autoinhibitor or activator depending on its conformation.

Figure 5

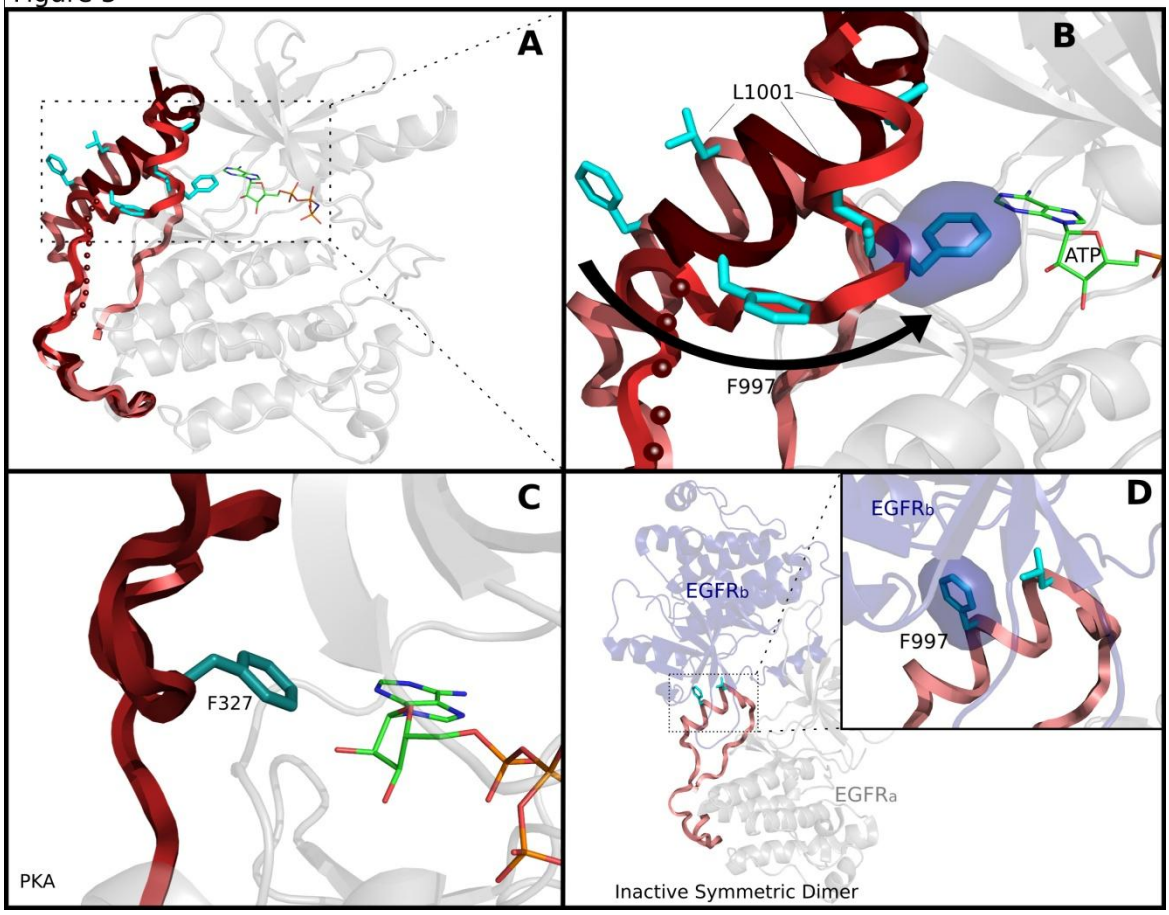


Figure 5: Active site tether (AST) and its role in ATP binding. A) Superimposition of the C-terminal tail in the active (pdb: 2jiu), inactive monomer (pdb: 1xkk) and inactive dime (pdb: 3gt8) showing the conformational flexibility in the AST. B) A close up view of F997 and its proposed role in ATP binding. The bound ATP was modeled to show the proximity of F997 to ATP. C) Conformational changes associated with F997 in the inactive symmetric dimer D) An analogous role of F327 in the C-tail of PKA (pdb: 1atp(46)).

CHT couples the JM segment to the regulatory C-helix

One of the key functional elements of the kinase core is the flexible C-helix, which switches between an inactive "out" conformation to an active "in" conformation upon EGFR activation. Previous structural studies have shown that the conformation of the C-helix is stabilized upon forming the asymmetric dimer in which C-lobe of one monomer (activator) interacts with the C-helix of the other monomer (receiver) to position it in an active

conformation (8). Our analysis suggests that in addition to this *trans* interaction, *cis* interaction between the C-helix and JM region, through CHT interactions described below, may also contribute to JM mediated activation of the kinase domain.

CHT is largely mediated through an ErbB-specific [YL]xx[AG] motif in the C-helix and conserved residues in the JM region. In EGFR, the tyrosine (Y764) within the [YL]xx[AG] motif forms a hydrogen bond with a conserved asparagine (N700) in the JM region (Fig 6A), while the alanine packs up against a hydrophobic residue (not shown) in the JM segment. These tethering interactions are also coupled to the kinase conserved glutamate in the C-helix through a backbone hydrogen bond between the alanine(A767) in the [YL]xx[AG] motif, and an alanine (A763) sequence adjacent to the C-helix glutamate (E762) (Fig 6A). This coupling of the JM segment to the C-helix glutamate may be significant considering that the precise positioning of the C-helix glutamate is critical for the activation of most kinases(38). Consistent with this notion, in the inactive state of EGFR, where the C-helix is in the "out" conformation, some of the canonical CHT interactions are altered. Specifically, the hydrogen bond between N700 and Y764 in the C-helix is lost because of the disordering of this portion of the JM segment, and Y764 and the C-helix glutamate adopt a strikingly different conformation in the inactive state (Fig 6B). The disruption of the CHT interactions in the inactive state, and the reformation of the interactions in the active state, suggests an active role for the CHT residues in coupling JM and C-helix functions. Such a role readily explains the *trans* regulatory interactions involving the JM segment in the active (Fig 6A) and inactive dimer (Fig 6C), and the significant reduction in EGFR activity upon Y764F and N700A mutation(12).

Figure 6

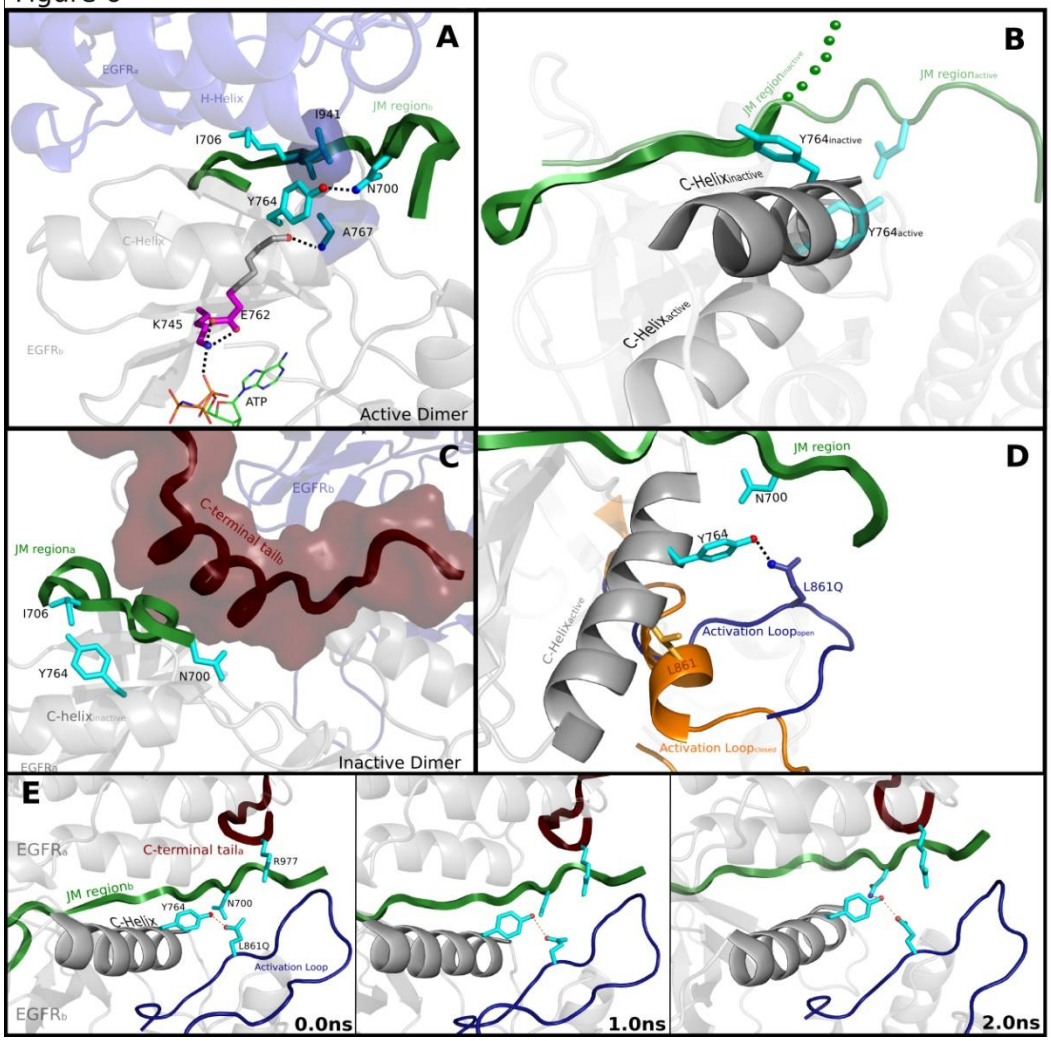


Figure 6: Role of C-helix tether (CHT) in C-helix movement A) CHT interactions in the active state of EGFR (pdb: 2j5f(47)). The van der Walls interaction mediated by I941 is shown in transparent spheres representation. B) Conformational changes associated with the CHT residues in the active (pdb: 2j5f) and inactive state of EGFR (pdb: 2rfe(47)) C) CHT interactions in the inactive dimer (pdb: 3gt8), where the C-terminal tail of one monomer docks on to the JM segment of the other monomer. The C-terminal tail is shown in transparent surface representation. D) L861Q mutation shown in the active state of EGFR. E) Snapshots of the L861Q mutation in the first two nano seconds of the molecular dynamics trajectory.

A frequently occurring lung cancer mutation (L861Q) may activate EGFR by altering the canonical CHT interactions

In addition to tethering the JM region to the C-helix, the tyrosine in the [YL]xx[AG] motif is also involved in packing up against a leucine (L861) in the activation loop (Fig 6D).

L861 is frequently mutated to a glutamine (L861Q) in many lung cancer patients (39). In fact, the L861Q mutation is the third most frequent mutation in EGFR in the COSMIC database (40). Modeling of a glutamine at the L861 position suggests that the glutamine can form two potential hydrogen bonds in the active dimer. It can hydrogen bond to Y764 in the [YL]xx[AG] motif as well to a conserved arginine (R977) in the activating monomer. To better understand the role of these hydrogen bonds in the abnormal activation of EGFR, we performed molecular dynamics simulation studies on the modeled L861Q mutant structure (see Methods).This revealed that, unlike the Q861-R977 hydrogen bond, the hydrogen bond between Q861 and Y764 is stable during the course of simulation and prevents Y764 from switching to the inactive conformation This suggests that L861Q mutation can potentially active EGFR by stabilizing Y764 in an active conformation.. This stabilization of Y764 may also facilitate the formation of the asymmetric dimer and destabilization of the inactive state, as proposed in previous studies (41, 8).

Notably, Erb3 and Erb4 in humans, and EGFR orthologues in lower eukaryotes do not conserve the tyrosine in the [YL]xx[AG] motif (Fig 1), instead they conserve a leucine at the [YL]xx[AG] tyrosine position. This suggests that L861Q mutations in Erb3 and Erb4 may not have the same functional impactas in EGFR or ErbB2.

CHAPTER 3
CONCLUDING REMARKS

Activation of ErbB kinases requires allosteric coupling of the conformational changes associated with catalytic functions of the kinase core, and the regulatory functions of the flanking JM and C-terminal tail segments. Analysis of the selective constraints imposed on ErbB kinases suggests a model in which conformational coupling can be achieved through *cis* interactions between the kinase core and the flanking JM and C-terminal tail segments. A compelling aspect of this model is that it readily explains the inhibitory and activating functions of the C-terminal tail segment in *in vitro* and *in invo* studies(10-12, 37), and provides new hypotheses for experimental studies. In particular, the proposed role of the ILT in coupling inter-lobe movement with C-terminal tail movement can be tested by mutating the ErbB kinase specific glutamine (Q791) to a glutamate. The conformational regulation of ErbB kinase by JM and C-terminal tail segments is also predicted to be specific to metazoan ErbB's since non-metazoans lack the canonical motifs for tethering the JM and C-terminal tail segments to the kinase core. Finally, the proposed role of the AST in ATP binding, provides clues for designing ErbB kinase specific small molecule inhibitors.

CHAPTER 4
MATERIALS AND METHODS

Identification of ErbB-specific selective constraints

The selective constraints imposed on ErbB sequences were measured using the CHAIN program (41), which requires construction of a "Contrast Hierarchical Alignment (CHA)" 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 (~286 sequences) correspond to the foreground set and all tyrosine protein sequence (~3512 sequences) correspond to the background set. The sequence pattern that most optimally distinguishes all ErbB sequences (the foreground set) from other tyrosine kinase sequences (background set) was identified using the Bayesian pattern partitioning procedure as implemented in the CHAIN program (41). For the C-terminal tail, a standard background set was used because the C-terminal tail segment does not share any significant sequence similarity with other receptor or non-receptor tyrosine kinase sequences.

Molecular dynamics simulations

Simulations were performed with NAMD software (42), version 2.7b1, and using all-atom ff03 force fields (43). Trajectory files were written every 2 ps using a time step of 2 fs and the NPT ensemble. A smoothing function was applied to both the electrostatic and the van der Waals forces at a distance of 10 Å with a switching cutoff distance of 12 Å. The non-bonded interactions parameters (switching, spherical cutoff, and pair list distances) were carefully chosen to achieve realistic macromolecular simulations (44). The Particle Mesh Ewald (PME)

algorithm, implemented to add a reciprocal space long-range electrostatic energy to the limited non-bonded component, was used with an interpolation function of order 5, and PME box grid size of 256, 256, 256 in x, y, and z dimensions, respectively.

WORKS CITED

1. Manning G, Whyte DB, Martinez R, Hunter T, & Sudarsanam S (2002) The protein kinase complement of the human genome. (Translated from eng) *Science* 298(5600):1912-1934 (in eng).
2. Yarden Y & Sliwkowski MX (2001) Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2(2):127-137.
3. Cho H-S & Leahy DJ (2002) Structure of the Extracellular Region of HER3 Reveals an Interdomain Tether. *Science* 297(5585):1330-1333.
4. Ferguson KM, *et al.* (2003) EGF Activates Its Receptor by Removing Interactions that Autoinhibit Ectodomain Dimerization. *Molecular Cell* 11(2):507-517.
5. Garrett TPJ, *et al.* (2002) Crystal Structure of a Truncated Epidermal Growth Factor Receptor Extracellular Domain Bound to Transforming Growth Factor [alpha]. *Cell* 110(6):763-773.
6. Ogiso H, *et al.* (2002) Crystal Structure of the Complex of Human Epidermal Growth Factor and Receptor Extracellular Domains. *Cell* 110(6):775-787.
7. Schlessinger J (2002) Ligand-Induced, Receptor-Mediated Dimerization and Activation of EGF Receptor. *Cell* 110(6):669-672.
8. 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(6):1137-1149.
9. Qiu C, *et al.* (2008) Mechanism of Activation and Inhibition of the HER4/ErbB4 Kinase. *Structure* 16(3):460-467.

10. 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(48):46265-46272.

11. Jura N, *et al.* (2009) Mechanism for Activation of the EGF Receptor Catalytic Domain by the Juxtamembrane Segment. *Cell* 137(7):1293-1307.

12. Red Brewer M, *et al.* (2009) The Juxtamembrane Region of the EGF Receptor Functions as an Activation Domain. *Molecular Cell* 34(6):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(2):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(5):1434-1442.

15. Guy PM, Platko JV, Cantley LC, Cerione RA, & Carraway KL (1994) Insect cell-expressed p180erbB3 possesses an impaired tyrosine kinase activity. *Proceedings of the National Academy of Sciences of the United States of America* 91(17):8132-8136.

16. Klapper LN, *et al.* (1999) The ErbB-2/HER2 oncoprotein of human carcinomas may function solely as a shared coreceptor for multiple stroma-derived growth factors. *Proceedings of the National Academy of Sciences of the United States of America* 96(9):4995-5000.

17. Citri A, Skaria KB, & Yarden Y (2003) The deaf and the dumb: the biology of ErbB-2 and ErbB-3. *Experimental Cell Research* 284(1):54-65.

18. Tzahar E, *et al.* (1996) A hierarchical network of interreceptor interactions determines signal transduction by Neu differentiation factor/neuregulin and epidermal growth factor. *Mol. Cell. Biol.* 16(10):5276-5287.

19. Jeffrey PD, *et al.* (1995) Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature* 376(6538):313-320.

20. Sicheri F, Moarefi I, & Kuriyan J (1997) Crystal structure of the Src family tyrosine kinase Hck. *Nature* 385(6617):602-609.

21. 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. (Translated from eng) *Proc Natl Acad Sci U S A* 104(4):1272-1277 (in eng).

22. Johnson LN, Noble MEM, & Owen DJ (1996) Active and Inactive Protein Kinases: Structural Basis for Regulation. *Cell* 85(2):149-158.

23. Nolen B, Taylor S, & Ghosh G (2004) Regulation of Protein Kinases: Controlling Activity through Activation Segment Conformation. *Molecular Cell* 15(5):661-675.

24. Huse M & Kuriyan J (2002) The Conformational Plasticity of Protein Kinases. *Cell* 109(3):275-282.

25. Yun C-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(3):217-227.

26. Wood ER, *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(18):6652-6659.

27. Wood ER, *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(8):2773-2778.

28. Gotoh N, Tojo A, Hino M, Yazaki Y, & Shibuya M (1992) A highly conserved tyrosine residue at codon 845 within the kinase domain is not required for the transforming activity of human epidermal growth factor receptor. *Biochemical and Biophysical Research Communications* 186(2):768-774.

29. Kannan N & Neuwald AF (2005) Did Protein Kinase Regulatory Mechanisms Evolve Through Elaboration of a Simple Structural Component? *Journal of Molecular Biology* 351(5):956-972.

30. 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(4):1272-1277.

31. Natarajan Kannan & Andrew F. Neuwald (2004) Evolutionary constraints associated with functional specificity of the CMGC protein kinases MAPK, CDK, GSK, SRPK, DYRK, and CK2alpha. *Protein Science* 13(8):2059-2077.

32. Torkamani A, Kannan N, Taylor SS, & Schork NJ (2008) Congenital disease SNPs target lineage specific structural elements in protein kinases. (Translated from eng) *Proc Natl Acad Sci U S A* 105(26):9011-9016 (in eng).

33. Robert A. Romano, Natarajan Kannan, Alexandr P. Kornev, Craig J. Allison, & Susan S. Taylor (2009) A chimeric mechanism for polyvalent <I>trans</I>-phosphorylation of PKA by PDK1. *Protein Science* 18(7):1486-1497.

34. 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(2):859-866.

35. Yun C-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(6):2070-2075.

36. 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(7):921-935.

37. Xia L, *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(34):30716-30723.

38. Huse M & Kuriyan J (2002) The conformational plasticity of protein kinases. *Cell* 109(3):275-282.

39. Shigematsu H, *et al.* (2005) Clinical and Biological Features Associated With Epidermal Growth Factor Receptor Gene Mutations in Lung Cancers. *J. Natl. Cancer Inst.* 97(5):339-346.

40. Bamford S, *et al.* (2004) The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website. *Br J Cancer* 91(2):355-358.

41. Neuwald AF (2007) The CHAIN program: forging evolutionary links to underlying mechanisms. *Trends in Biochemical Sciences* 32(11):487-493.

42. James C. Phillips, *et al.* (2005) Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* 26(16):1781-1802.

43. Yong Duan, *et al.* (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry* 24(16):1999-2012.

44. Peter J. Steinbach & Bernard R. Brooks (1994) New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of Computational Chemistry* 15(7):667-683.

45. Zhang X, *et al.* (2007) Inhibition of the EGF receptor by binding of MIG6 to an activating kinase domain interface. *Nature* 450(7170):741-744.

46. Zheng J, *et al.* (1993) 2.2 A refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor. *Acta Crystallographica Section D* 49(3):362-365.

47. Blair JA, *et al.* (2007) Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. *Nat Chem Biol* 3(4):229-238.