Structure, Vol. 12, 2265–2275, December, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.str.2004.10.006

A Putative Mechanism for Downregulation of the Catalytic Activity of the EGF Receptor via Direct Contact between Its Kinase and C-Terminal Domains

Meytal Landau, Sarel J. Fleishman, and Nir Ben-Tal*
Department of Biochemistry
George S. Wise Faculty of Life Sciences
Tel-Aviv University
Ramat-Aviv 69978
Israel

Summary

Tyrosine kinase receptors of the EGFR family play a significant role in vital cellular processes and in various cancers. EGFR members are unique among kinases, as the regulatory elements of their kinase domains are constitutively ready for catalysis. Nevertheless, the receptors are not constantly active. This apparent paradox has prompted us to seek mechanisms of regulation in EGFR's cytoplasmic domain that do not involve conformational changes of the kinase domain. Our computational analyses, based on the three-dimensional structure of EGFR's kinase domain suggest that direct contact between the kinase and a segment from the C-terminal regulatory domains inhibits enzymatic activity. EGFR activation would then involve temporal dissociation of this stable complex, for example, via ligand-induced contact formation between the extracellular domains, leading to the reorientation of the transmembrane and intracellular domains. The model provides an explanation at the molecular level for the effects of several cancer-causing EGFR mutations.

Introduction

The epidermal growth factor receptor (EGFR) family of receptor tyrosine kinases (RTKs), also known as ErbB or HER, consists of four members, ErbB1, -2, -3, and -4 (Schlessinger, 2000). The receptors, which are activated by some dozen ligands, including EGF and TGFα, play an important role in the control of many fundamental cellular processes (Schlessinger, 2000). Mutations and overexpression of the ErbBs have been implicated in malignant diseases such as carcinoma and glioblastoma and are linked with aggressive disease, resistance to chemotherapy, and poor survival (Dancey, 2004). Accordingly, the ErbBs are attractive targets for anticancer drugs (Cho et al., 2003; Yarden and Sliwkowski, 2001). Structurally, the ErbBs consist of an N-terminal, extracellular domain that is connected by a short transmembrane span to a tyrosine kinase domain, which is in turn followed by a C-terminal domain.

In all RTKs, including the ErbBs, the active kinase triggers a wide spectrum of crucial intracellular signaling events (Schlessinger, 2000), and their catalytic activity is encapsulated in multiple layers of regulation (Huse and Kuriyan, 2002). A primary means of regulation in RTKs is ligand binding to the extracellular domain, leading to dimerization or formation of higher-order oligomers of the receptors and enzymatic activation (Schlessinger, 2000, 2003). Similarly, activation of ErbB1, -3, and -4 involves ligand-induced contact formation between the extracellular domains of different members of the ErbB family to form homo- and heterodimers (Schlessinger, 2000). Some studies have shown that, without a ligand, EGFR exists mostly in a monomeric form and that ligands induce its dimerization and activation (Yarden and Schlessinger, 1987). On the other hand, recent studies have demonstrated that while required, dimerization is not sufficient for activation and that in the absence of a ligand, stable, inactive dimers exist in a form in which contact between monomers involves the transmembrane and intracellular domains (Biswas et al., 1985; Gadella and Jovin, 1995; Moriki et al., 2001; Yu et al., 2002). Experimental evidence (Cadena et al., 1994), as well as the computational results presented below, demonstrates that the C-terminal domain plays a role in such contact formation.

In most tyrosine kinases (TKs) excluding the ErbBs, an important means of regulation involves profound structural changes along with transautophosphorylation of the kinase domain (Schlessinger, 2000). In contrast, the ErbB family is unique in that activation is independent of its phosphorylation state (Gotoh et al., 1992). The structure of the apo-EGFR kinase domain demonstrated that its unphosphorylated conformation was, in essence, identical to the phosphorylated conformations of other TKs (Stamos et al., 2002).

Recently, a structure of the kinase domain of the EGFR in complex with the inhibitor GW572016 (Lapatinib) was determined (Wood et al., 2004). This structure shows several differences, including different conformations of the substrate and ATP binding sites (Wood et al., 2004), from either the structure of the apo-EGFR or of EGFR bound to the OSI-774 (Tarceva) inhibitor (Stamos et al., 2002). The authors have suggested that these differences are due to the fact that the conformation seen in the GW572016 bound kinase domain reflects an inactive state that is accessible to the kinase domain under physiological conditions. However, GW572016 is very bulky in comparison to OSI-774. Thus, as the authors indicated, another possibility is that the differences in the structures are due to the inhibitor's large size, which forces a conformation that is far from native. That the apo-EGFR kinase domain is seemingly in a constitutively active conformation (Stamos et al., 2002) leads to an apparent paradox, since it is well established that ErbBs are not constitutively active (Schlessinger, 2000). Hence, our working hypothesis, as presented in Figure 1, was that ErbBs are regulated by another mechanism intrinsic to the intracellular domain; one that is phosphorylation independent.

The orphan receptor ErbB2 presents an even more intriguing case than other members of the EGFR family because its activation is not only phosphorylation inde-

*Correspondence: bental@ashtoret.tau.ac.il

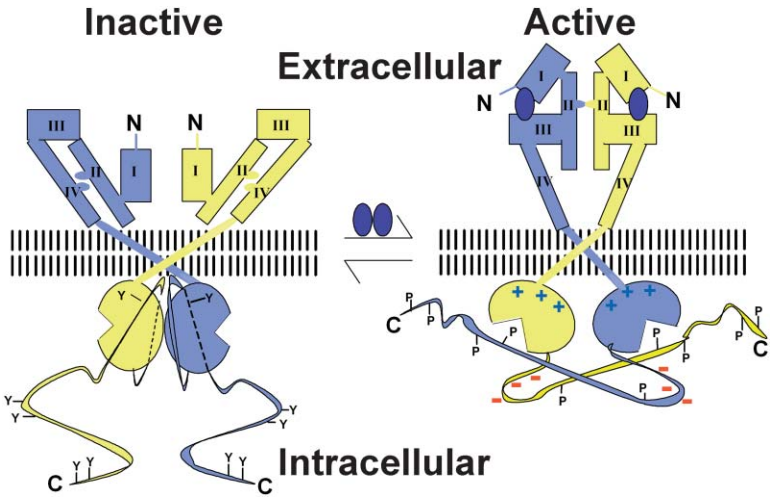

Figure 1. Schematic Diagram Representing the Suggested Model of EGFR Activation

Two EGFR monomers are colored light purple and yellow. The extracellular domain (residues 1–620, labeled I, II, III, and IV according to its subdomains) and the kinase domain (residues 685–957) are connected via a transmembrane helix (residues 621–642) and a short juxtamembrane segment (not shown). The C-terminal domain, comprising 229 amino acids, whose structure has not been determined, follows the kinase domain. Tyrosine residues (Y) known as the autophosphorylation sites in the C-terminal domain are indicated. In the inactive conformation (left), each of the extracellular domains assumes a compact structure, and the intracellular domains contact via the C-terminal fragments, leading to an inactive and stable form. Activation (right) occurs when ligands (purple ovals) bind to the extracellular domains, leading to the formation of a stable extracellular contact, which is followed by the rotation of the transmembrane helices and the subsequent destabilization of the contacts between the C-terminal and kinase domains. The kinase can now transautophosphorylate the tyrosine residues of its own C-terminal domain, as well as tyrosine residues of its protein substrates. The figure displays an illustration of the transmembrane domain; the suggested molecular model for the transmembrane domain in the active and inactive states was presented in (Fleishman et al., 2002). Positive and negative charges are marked in the active conformation on the kinase and C-terminal domains, respectively. In the inactive conformation, they roughly neutralize each other (Figure 2).

pendent, but also ligand independent (Cho et al., 2003). The absence of clear regulation of ErbB2 activation prompted us to propose a molecular mechanism for rotation-coupled activation of this receptor (Fleishman et al., 2002). Specifically, the transmembrane domain of an ErbB2 homodimer may occupy one of two stable conformations, corresponding to the active and inactive states of the receptor. The switch between the two conformations, involving a rotation of the transmembrane domain (Jiang and Hunter, 1999), induces the reorientation of the cytoplasmic domains within receptor dimers, thus leading to transautophosphorylation and stimulation of enzymatic activity. In this paper, we shall analyze the implications of this mode of activation on the conformation of the intracellular kinase domain.

The C-terminal domain plays a role in the internalization and degradation of the EGFR (Chang et al., 1995) and in EGFR's regulation by other molecules (Huse and Kuriyan, 2002). This domain also serves as a docking site for protein modules that bind the phosphotyrosines on the activated receptors (Schlessinger, 2000). In addition to these roles, the importance of the C-terminal domain for proper functioning of the EGFR was previously noted on the basis of studies of viral and other mutant EGFR members (Boerner et al., 2003; Chang et al., 1995; Wedegaertner and Gill, 1992).

Naturally occurring retroviral oncogene variants (v-ErbB) are an extracellular domain-truncated form of the EGFR gene that affects cell growth, motility, and survival (Gamett et al., 1986). These v-ErbB variants share striking homology with mutants of the human EGFR members that have been identified in gliomas and carcinomas (Frederick et al., 2000). Truncation of the extracellular domain is insufficient to manifest the transforming properties of the different v-ErbB variants; these properties are probably related to amino acid replacements, insertions, and deletions in the C-terminal domain (Boerner et al., 2003; Massoglia et al., 1990; Riedel et al., 1987). Variations in the C-terminal domain of ErbB receptors are known to be responsible for the alterations in the transforming potential and type of malignant diseases due to the expression of v-ErbBs in affected cells (Gamett et al., 1986; Pelley et al., 1989; Raines et al., 1988). The increased substrate-phosphorylation capacity of the C-terminally impaired EGFR is not attributed to lesser degradation and internalization, but rather to an enhanced rate of autophosphorylation (Robinson et al., 1992), thus providing direct evidence for a relationship between C-terminal domain impairment and increased catalysis.

Here, we propose a molecular model clarifying some of the ambiguity regarding the role of the C-terminal domain in ErbB regulation. According to the model (Figure 1), contact between the intracellular domains of the EGFR within a dimer leads to receptor inactivation, while ligand-induced contact between the extracellular domains leads to rotation-coupled activation (Fleishman et al., 2002; Jiang and Hunter, 1999; Moriki et al., 2001) by destabilization of the intermonomer contacts in the cytoplasmic domain. According to this scenario, interactions between the intracellular domains regulate activation (Burgess et al., 2003; Chantry, 1995), and the C-terminal domain acts as an inherent negative regulator of the EGFR's activity. This model offers a molecular mechanism that underlies the tumorigenic activity of EGFR mutants.

Results

Geometric Complementarity between the Kinase and C-Terminal Domains

The crystal structure of the EGFR (Stamos et al., 2002) (PDB entries 1m14 and 1m17) includes the kinase domain (residues 685–957) and a segment from the

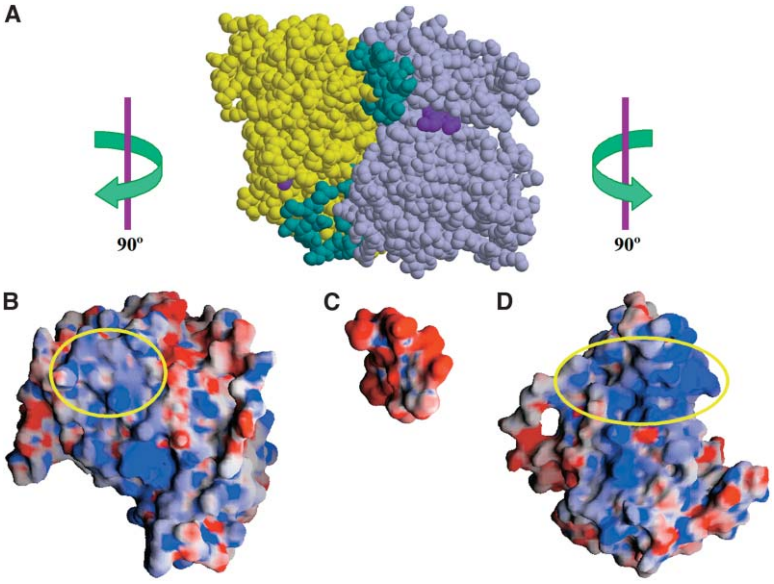

Figure 2. Geometric and Electrostatic Complementarity between the Kinase and C-Terminal Domains

(A) A space-filled model of EGFR's homodimeric complex (Stamos et al., 2002) showing the geometric complementarity within the complex. The kinase domains are colored light purple and yellow, the C-terminal fragments are colored cyan, and the inhibitor is colored purple. The dimer is symmetric, which means that each kinase domain is in contact with both C-terminal fragments, yielding one large and one small interface per monomer. The interactions with the C terminus are identical in both monomers. Figure 2A was made by using MOLSCRIPT (Kraulis, 1991) and Raster3D (Merritt and Bacon, 1997).

(B–D) A projection of the electrostatic potential (ϕ) onto the molecular surface of the kinase domain and the C-terminal fragment that comprise the complex in (A); $\phi > 10$ kT/e is dark blue, $\phi = 0$ is white, and $\phi < -10$ kT/e is dark red. Yellow ellipses mark the interfaces between the kinase domains and the C-terminal fragments. The figures were produced by using GRASP (Nicholls et al., 1991). (B) The left-most kinase domain shown in (A) (yellow) was rotated 90° to the left relative to its orientation in (A). (C) The C-terminal fragment is shown in the same orientation as the upper segment in (A). (D) The right-most kinase domain shown in (A) (light purple) was rotated 90° to the right relative to its orientation in (A). The electrostatic complementarity between the negatively charged C-terminal fragment and the positively charged residues of the kinase domain that interact with it is noticeable.

C-terminal domain (residues 977–995). The crystal structure reveals six putative dimer forms (Stamos et al., 2002). We focus here on the one with the largest intersubunit interface. In this complex, the kinase domain was found as a symmetric homodimer (Figure 2A), in which two copies of the fragment of the C-terminal domain mediate contact between the two kinase domains. This dimer is also the only one in which the kinase domains' N termini are facing in the same direction, in accordance with the physiological requirement that the two domains connect to the membrane bilayer.

We calculated the water-accessible surface area of the kinase domain alone and within the homodimeric complex. Each kinase monomer contacts two C-terminal fragments (Figure 2A). The water-accessible surface areas of these interfaces are 1419 Å^2 and 1048 Å^2. Thus, the total interface between each monomer of the kinase domain and the C-terminal fragments is 2467 Å^2, and the interface within the entire complex is twice as large, constituting a very large interface compared to typical interprotein interfaces (Jones and Thornton, 1996).

Charge Complementarity between the Kinase and C-Terminal Domains

Electrostatic calculations show strong positive potential in the kinase domain at its interface with the C-terminal fragments (Figures 2B and 2D). This potential originates from positively charged residues in both subunits, suggesting that the kinase domains would repel one another in the absence of the C-terminal fragments. Kinase domains from other ErbBs, which were constructed by using homology modeling, displayed similar positive electrostatic potentials in the corresponding regions (data not shown). The C-terminal fragment of all of the ErbBs contained 8–10 acidic and no basic residues (Figure 3). These residues produced a highly negative electrostatic potential (Figure 2C). Thus, the kinase domain and the C-terminal fragments form complementary surfaces in terms of their electrostatic potential. The geometric and charge complementarity (Figures 2B–2D), together with the significant size of the interface (Figure 2A), are indicative of the stability of the complex and suggest that it may be biologically meaningful.

Following the experiments of Chang et al. (1995) discussed below, we simultaneously substituted each of the negatively charged residues 979–982 (DEED, Figure 3) in the C-terminal domain with its polar equivalent, i.e., D→N and E→Q. The mutated C-terminal fragment is much less negatively charged than the native fragment (Figure 4B), and this difference in charge obstructs its electrostatic complementarity with the kinase domain and presumably destabilizes the complex. We further mutated the same positions to four positively charged lysine residues (Figure 4C). Electrostatic analysis of the mutated C-terminal fragment displayed a positive potential at the N-terminal region of the fragment, which would lead to its electrostatic repulsion from the kinase domain. To test whether the charge complementarity is

EGFR	979	D	E	E	D	M	D	D	V	V	D	A	D	E	Y	L	I	P	Q	996
ERB2	986	E	D	D	D	M	G	D	L	V	D	A	E	E	Y	L	V	P	Q	1003
ERB3	977	E	E	V	E	L	E	P	E	L	D	L	D	L	D	L	E	A	E	994
ERB4	985	D	E	E	D	L	E	D	M	M	D	A	E	E	Y	L	V	P	Q	1002

Figure 3. Abundance of Acidic Residues in the Fragment of the C-Terminal Domain

The multiple sequence alignment of the C-terminal segments of the four human members of the ErbB family. Each segment contains between 8 and 10 acidic residues (marked in red).

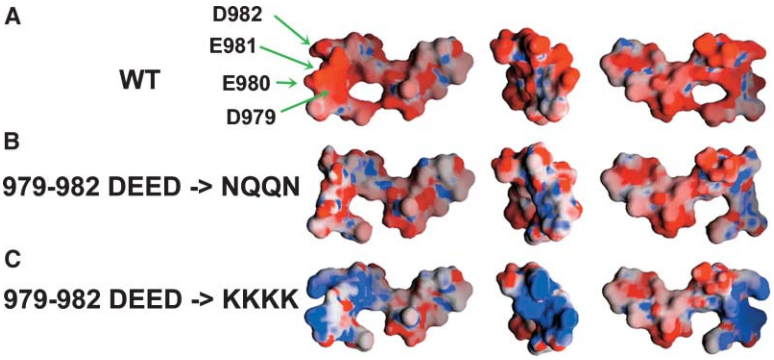

Figure 4. Electrostatic Analysis of Reported and Novel Mutations in the EGFR C-Terminal Fragment

A projection of the electrostatic potential (ϕ) onto the molecular surface of the C-terminal fragment; the color coding is as in Figures 2B–2D.

(A) The C-terminal fragment of the native EGFR, in the same orientation as in Figure 2C (central image), rotated 90° to the right (right image) or left (left image). The location of selected residues is marked.

(B) The C-terminal fragment, in the same orientations as in (A), in which the negatively charged residues in the 979–982 positions (DEED) were mutated to their polar equivalent (NQQN).

(C) The C-terminal fragment, in the same orientations as in (A), in which the same positions were mutated to positively charged lysine residues.

unique to ErbBs among TKs, we examined the electrostatic potentials of a few TKs of known structures as described in the Supplemental Data (available with this article online; Electrostatic calculations). These domains, which were derived from remotely related proteins, display diverse electrostatic characteristics. In particular, they do not share EGFR's strong positive electrostatic potential at the interface with the C-terminal fragments (data not shown), suggesting that such electrostatic interactions between the kinase and the C-terminal domains are specific to the ErbBs.

A Network of Ion Pairs and Hydrogen Bonds at the Interface

Our analysis demonstrated that a network of salt bridges and hydrogen bonds connects the two adjacent kinase domains through the C-terminal fragments (Figure 5B). We identified four charged residues within this network that are involved in several interactions with neighboring residues and are buried at the interface of the EGFR complex. Of these residues, two are positively charged (Lys822 and Lys828 on the kinase domain) and two are negatively charged (Asp988 and Asp990 on the C-terminal fragment) (Figure 5B).

Polar networks, such as the one observed in the EGFR interface (Figure 5B), significantly increase the stability of complexes and contribute to the binding specificity (Sheinerman et al., 2000). Therefore, mutations of charged positions in the network would alter the stability of the complex (Serrano et al., 1990). An even larger effect would be obtained by mutating them in pairs. For example, mutating Lys822 and Lys828 to aspartates or Asp988 and Asp990 to lysines altered the electrostatic surface of the kinase domain and the C-terminal fragment, respectively (Figure 6). Such mutations would impinge on the formation of the EGFR complex and kinase activation.

The importance of the network for the stability of the

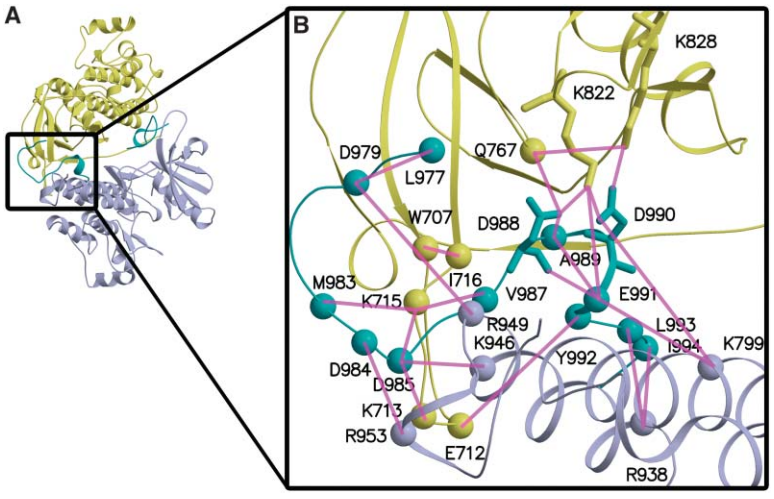

Figure 5. A Network of Ion Pairs and Hydrogen Bonds across the Interface of the EGFR Complex

The kinase domain monomers are displayed as ribbons and colored light purple and yellow. The C-terminal fragment is colored cyan.

(A) The EGFR homodimeric complex (Stamos et al., 2002) as viewed with a clockwise rotation of about 90° compared to Figure 2A.

(B) A close view, in the same orientation as in (A), of the polar network connecting the C-terminal fragment with its two adjacent kinase domains. The Cα atoms of residues comprising the polar network are displayed as spheres. Four selected residues in the network (Lys822, Lys828, Asp988, Asp990; their side chains displayed as sticks) are buried in the core of the kinase/C-terminal fragment interface, suggesting that they play a key role in complex stabilization (Sheinerman et al., 2000). Solid pink lines connect the Cα (or nearest neighbors) atoms of residues that form ion pairs and hydrogen bonds in the network. By symmetry, identical interactions connect residues between the second C-terminal fragment and the kinase domains (not shown). Each residue in the network is involved in a few interactions with neighboring residues. For instance, Asp990, located on the C-terminal domain, interacts with Lys822, located on one kinase domain monomer (yellow), and with Lys799, located on the second kinase domain monomer (light purple), presumably stabilizing the complex.

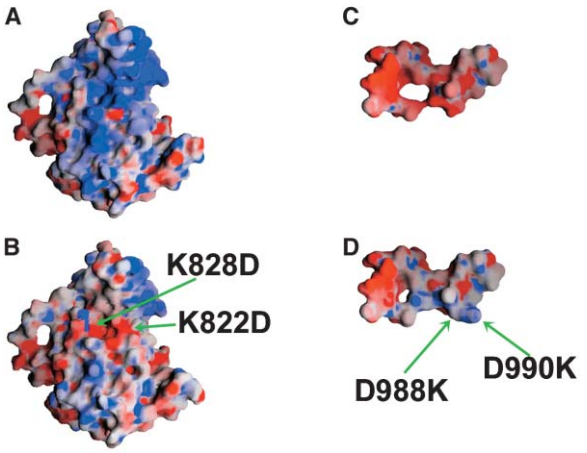

Figure 6. Electrostatic Analysis of Novel Mutations in the EGFR Kinase and C-Terminal Domains

A projection of the electrostatic potential (ϕ) onto the molecular surface of the kinase domain and the C-terminal fragment; the color coding is as in Figures 2B–2D.

(A) The native kinase domain in the same orientation as in Figure 2D.

(B) The R822D/R828D double mutant EGFR kinase domain in the same orientation as in (A).

(C) The native C-terminal fragment of the EGFR rotated 90° to the left relative to its orientation in Figure 2C.

(D) The D988R/D990R double mutant C-terminal domain in the same orientations as in (C).

complex can be tested experimentally by using the double mutant cycle approach (Serrano et al., 1990). Briefly, if the additive effects of mutating two residues separately (e.g., Lys822→Asp and Asp990→Lys) is significantly different from the effect of mutating the same two residues simultaneously, then the two positions are interdependent (Serrano et al., 1990), e.g., are involved in a salt bridge. Based on our analysis of the network, we suggest using a double mutant cycle, in which each step involves mutating a pair of similarly charged residues in the EGFR interface, as specified above.

Model of C-Terminal Domain Regulation of Kinase Activity

One of the phosphorylation sites of the C-terminal domain (Tyr992) is located on the fragment that forms contact with the kinase domain and is therefore inaccessible to phosphorylation in this conformation. The catalytic sites in the kinase domains are facing away from each other in the complex; therefore, transphosphorylation of residues on the kinase domain is improbable. The above two observations suggest that the EGFR crystal structure represents an inactive form of the receptor. The EGFR participates in imperative cell processes and ought to remain inactive under most physiological conditions (Huse and Kuriyan, 2002). Therefore, its inactive state should be very stable. Indeed, the complex in the crystal structure of the EGFR appears to be stable, based on the geometric and charge complementarity, further supporting the notion that this complex is inactive. It has been suggested that ligand-induced contact formation of the extracellular domains would lead to reorientation of the transmembrane domains (Fleishman et al., 2002; Jiang and Hunter, 1999; Moriki et al., 2001) and, subsequently, to rearrangements in the cytoplasmic domains (Figure 1). Any reorganization of the cytoplasmic complex, followed by a change in the position of the negatively charged C-terminal fragment, would lead to electrostatic repulsion between the two positively charged kinase monomers (Figure 2). Hence, this model of conformational changes during receptor activation may constitute a hitherto unknown mode of regulation.

Strong reinforcement of this model of regulation is provided by data on the EGFR analog c-ErbB (Chang et al., 1995). Deletions of a C-terminal fragment of this receptor (corresponding to residues 966–1006 of the EGFR) lead to higher autokinase activity compared to normal c-ErbB and transforming ability in vitro and in vivo. Moreover, a mutant in which the four consecutive acidic residues EEED were replaced by the polar segment QQQN showed higher autokinase activity and a partial transformation phenotype. Since the two mutants and normal receptors have similar rates of degradation, the higher transforming ability of the mutants could not be attributed to a longer half-life of the mutant receptor (Chang et al., 1995). These data are consistent with our results. The four acidic residues, which correspond to the DEED segment (Asp979–Asp982) in the EGFR, are located on the C-terminal fragment (Figure 3) that forms contact with the kinase domain. Our analysis showed that these positions contribute significantly to the negative electrostatic potential of the fragment (Figure 2), and their substitution with polar residues reduces the complementarity between the kinase and C-terminal domains (Figure 4B), presumably destabilizing the inactive complex.

Internal deletions of segments in the C-terminal domain of the EGFR have also been detected in naturally occurring EGFR mutants, which display tumorigenic properties. For example, an internal deletion of residues 959–1030 has been detected in EGFRs sequenced from human glioblastomas (Boerner et al., 2003; Chang et al., 1995; Frederick et al., 2000). Some viral ErbBs contain an in-frame deletion of 139 residues within the intracellular region, immediately following the kinase domain (Boerner et al., 2003; Chang et al., 1995; Frederick et al., 2000). This region contains the C-terminal fragment contacting the kinase domain according to the X-ray structure (Stamos et al., 2002). Our model suggests that the internal deletions in the C-terminal domain yield constitutively active forms of EGFR by means of destabilization of the inactive complex.

Evolutionary Conservation Analysis

The kinase domain of ErbB3 has no catalytic activity, yet it dimerizes with other members of the ErbB family to produce heterodimers with highly efficient catalytic activity (Schlessinger, 2000). These distinct features are manifested in the evolutionary-conservation analysis. ErbB3's kinase domain displays variations in the catalytic site in comparison to other members of the ErbB family, thus rendering it inactive. However, the interface between the kinase domain and the C-terminal fragment

is highly conserved within the ErbBs and their orthologs, including ErbB3. As a reference, an analysis of 121 kinase domains from various TKs showed that the catalytic site, including the ATP and substrate binding loop, was highly conserved, whereas the interface between the kinase domain and the C-terminal fragment was highly variable (data not shown).

Overall, the conservation analysis provides further support for the suggestion that the dimeric complex observed in the crystal structure is not common to all the TKs. However, the contact area between the kinase and C-terminal domains in this complex is common to the ErbBs, which thus maintain the ability to produce homo- and heterodimers through the same interface.

A Network of Correlated Amino Acid Substitutions between Regulatory Elements

By and large, all TKs carry out the same catalytic process. Thus, key residues in the kinase domain, which are responsible for catalysis of phosphotransfer, are under strong evolutionary constraint, as mentioned above. However, in order for the kinases to be involved in numerous and distinct signal transduction pathways, each kinase family exhibits variations in its amino acid sequence that are necessary for the modification of the mode of regulation. Since multiple positions are involved in determining these traits, these sequence variations should occur concomitantly in relevant regulatory elements. In other words, during evolution, substitutions of one residue in regulatory elements may be compensated by a concurrent change in another residue, in order to maintain the structural or functional relationship between these positions (Fleishman et al., 2004b).

In order to look for particular positions that could play a role in regulation, we analyzed the set of 121 multiply aligned TKs of diverse families in search of pairs of amino acid positions that might be evolutionarily correlated (Fleishman et al., 2004b). The analysis revealed 152 pairs of correlated residues, among which we identified a network of 14 highly intercorrelated positions (Figure 7A and Table 1).

The kinase domain includes several regulatory elements, such as the αC helix and activation loop, which play a role in allosteric regulation and are responsible for conformational changes. These elements function together to control activation, i.e., their movements are concurrent and their conformations are mutually dependent (Huse and Kuriyan, 2002). Our analysis showed pairs of evolutionarily correlated positions in these known regulatory elements. For example, Ala743, which is located on the αC helix, is correlated with Gly849 of the activation loop (Figure 7A).

The LVI segment (residues 955–957) of EGFR and its equivalent segments in other ErbBs are necessary for ligand-independent dimerization of the EGFR intracellular domains and for transphosphorylation in ErbB2 and ErbB3 heterodimers through allosteric regulation (Stamos et al., 2002). Leu955 in this LVI segment is correlated with Tyr740, which is located on the αC helix (Figure 7A). The association of the αC helix with a known dimerization motif exemplifies interdomain relationships between regulatory elements in the ErbBs. Both of these

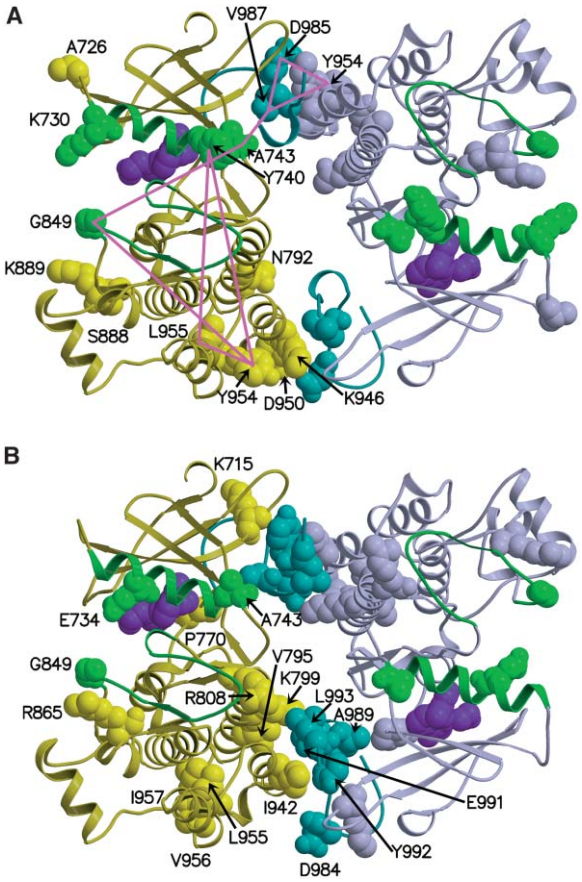

Figure 7. Evolutionarily Correlated and Specificity-Determining Amino Acid Sites

The EGFR homodimeric complex (Stamos et al., 2002) as viewed with an upward rotation of about 180° compared to Figure 2A. The kinase domains, presented by using trace models, are colored light purple and yellow, the C-terminal fragments are colored cyan, and the inhibitor is shown as a purple space-filled model. The αC helix (residues 729–744) and the activation loop (residues 831–852) are colored green.

(A) The residues in the cluster of the most significant pairs of correlated amino acid sites are displayed as space-filled models. Solid pink lines connect a few of the pairs of correlated residues (highlighted in Table 1) in the EGFR homodimer. Correlations within the kinase domain are demonstrated only on the left monomer, and correlations between the kinase and the C-terminal domains are demonstrated only on the upper interface. The correlations between known regulatory elements, such as the αC helix and the activation loop and the interface between the kinase domain and C-terminal fragment, suggest that the latter may also be involved in regulation.

(B) The main specificity-determining residues are located on the αC helix, the activation loop, the C-terminal fragment, and its interfaces on the kinase. This suggests that the regulatory elements in the EGFR had evolved specifically to stabilize the active conformation. Concurrently, an alternative negative regulatory mechanism had evolved in the form of the inactive complex between the kinase and the C-terminal domains. The figures were made by using MOLSCRIPT (Kraulis, 1991) and Raster3D (Merritt and Bacon, 1997).

domains are important for regulation (Huse and Kuriyan, 2002; Stamos et al., 2002); for example, mutations in Leu955 or Tyr740 severely impaired the kinase activity of the EGFR (Stamos et al., 2002; Walker et al., 1998).

Based on these results, we concluded that this net-

Table 1. Correlated Pairs in the TK Family

Pairs of Correlated Positions	Correlation Coefficients
Ala726-Lys730	0.65 (0.49, 0.78)
Ala726-Ser888	0.60 (0.34, 0.77)
Ala726-Lys946	0.45 (0.20, 0.64)
Lys730-Ser888	0.59 (0.37, 0.76)
Tyr740-Tyr954	**0.45 (0.19, 0.67)**
Tyr740-Leu955	**0.48 (0.21, 0.69)**
Ala743-Asn792	0.50 (0.19, 0.76)
Ala743-Gly849	**0.47 (0.18, 0.72)**
Ala743-Ser888	0.45 (0.21, 0.69)
Ala743-Lys889	0.51 (0.14, 0.73)
Ala743-Val987	0.54 (0.34, 0.72)
Asn792-Gly849	0.48 (0.26, 0.70)
Asn792-Ser888	0.50 (0.25, 0.72)
Gly849-Tyr954	**0.56 (0.29, 0.76)**
Ser888-Lys946	0.61 (0.40, 0.78)
Lys946-Asp950	0.52 (0.31, 0.69)
Asp950-Tyr954	0.48 (0.31, 0.65)
Asp950-Leu955	0.50 (0.26, 0.67)
Tyr954-Leu955	**0.54 (0.32, 0.72)**
Tyr954-Asp985	**0.50 (0.30, 0.67)**
Tyr954-Val987	**0.52 (0.36, 0.67)**
Asp985-Val987	**0.52 (0.36, 0.67)**

A list of 22 pairwise correlations between positions comprising the most significant cluster of correlated residues. The trimmed means in the 95% confidence interval of correlations (r), which were calculated from 400 bootstrapping samples, are indicated, and the 95% confidence interval is shown in parentheses (see the Supplemental Data). The numbering of the positions is done according to the EGFR sequence. The pairs of positions that are located on known regulatory regions are highlighted in bold and are connected by solid pink lines in Figure 7A.

work of correlations identified amino acids playing a role in regulation. Interestingly, the same cluster also displays correlations between residues mediating contact between the kinase and C-terminal domains. Tyr954 is located on the kinase domain and contacts the C-terminal fragment. This residue is in close proximity to, and is highly correlated with, residues Asp985 and Val987 of the C-terminal fragment (Figure 7A). Taken together, these correlations consolidate our hypothesis that the contact between the kinase and C-terminal domains is biologically meaningful.

The same cluster of 14 highly intercorrelated positions also includes correlations between positions at the interface of the kinase domain and the C-terminal fragment and known regulatory elements. For instance, Tyr954 located on this interface is correlated with Gly849 of the activation loop, with Leu955 of the LVI segment, and with Tyr740 of the αC helix (Figure 7A). This network of correlations suggests that this interface is also involved in regulation.

Val987 of the C-terminal fragment is correlated with Ala743, which is located on the αC helix (Figure 7A). In this context, it is important to note that the C-terminal domain is a vital modulator of TKs' activity (Jorissen et al., 2003; Schlessinger, 2000), as was elaborated above. For example, structure determination and mutagenesis experiments have shown that the kinase domains of the insulin, the Tie2, and the platelet-derived growth factor β receptor (PDGFR) TKs are autoinhibited by their C-terminal domains through direct contacts with the kinase domain (Chiara et al., 2004; Noelle et al., 2000; Shewchuk et al., 2000). Accordingly, evolutionary correlation between the kinase and C-terminal domains is expected to be general. The mechanism by which direct contacts control activation may vary between the kinases and could not be inferred from the evolutionary-correlation analysis. We anticipate that in the ErbBs, the direct contact between the kinase and C-terminal domains regulates catalysis by the formation of the inactive dimer shown in Figure 2A.

Specificity Determinants in Regulatory Regions

Although TKs share an identical catalytic mechanism, each kinase family is regulated by various means, responds to different ligands, and activates diverse substrates. It is anticipated that certain positions would be responsible for these different traits, and would be reflected in their patterns of substitution (Fleishman et al., 2004a). Due to such differences in functions, these positions are not expected to be strictly conserved in evolution. Rather, they should be conserved among kinases of similar functions in different species (orthologs), and would differ in paralogs. Substitutions involving these residues are presumably responsible for certain alterations in the functions of the various families of the TK superfamily.

We have identified some of these specificity-determining amino acid positions in a set of 121 multiply aligned TKs. The main specificity-determining residues are presented in Figure 7B, and their locations are indicated in Table 2. The list includes residues from the known regulatory regions, as well as residues that connect the kinase and the C-terminal domains and participate in the polar network across the interface (Figure 5B).

Discussion

ErbBs are structurally unique among TKs, as all of the catalytic elements in their kinase domains are ready for phosphotransfer at all times (Stamos et al., 2002). Yet, various functional assays show them not to be constitutively active (Schlessinger, 2000). The absence of a central regulatory module raises a fundamental dilemma, namely, what prevents the receptors from being spuriously activated? One possible mechanism is that changes in the relative orientation of the subunits within a dimer control activation, as suggested by the model of rotation-coupled activation (Jiang and Hunter, 1999). According to this view, contact formation between the extracellular domains leads to reorientation in the transmembrane domain, which is propagated into the cytoplasm (Fleishman et al., 2002; Jiang and Hunter, 1999; Moriki et al., 2001). Thus, the reorientation of the kinase domains vis-à-vis each other serves as a molecular switch that turns the kinase domains "on." What might be the mechanism by which this reorientation is translated into kinase activation is not yet clear.

Understanding the molecular details of how the ErbB proteins are regulated will most probably have to await the emergence of a structure of the full-length receptor in oligomeric complexes. The structures of parts of the kinase and the extracellular domains available today only provide a fragmentary view of the regulatory ele-

Table 2. Specificity Determinants in the TK Family

Position	Correlation Coefficients	Location and Putative Functional Role in the EGFR
Lys715	0.23 (0.11, 0.36)	Located on the kinase at the large interface with the C-terminal fragment; participates in the polar network across the interface (Figure 5).
Pro770	0.24 (0.07, 0.39	Located on the kinase at the large interface with the C-terminal fragment.
Val795	0.29 (0.17, 0.40)	Located on the kinase at the small interface with the C-terminal fragment.
Ile942	0.24 (0.13, 0.36)	Located on the kinase at the small interface with the C-terminal fragment.
Lys799	0.24 (0.10, 0.36)	Located on the kinase at the small interface with the C-terminal fragment; participates in the polar network across the interface (Figure 5).
Glu734	0.27 (0.12, 0.42)	Located on the αC helix of the kinase domain; involved in regulation.
Ala743	0.29 (0.14, 0.42)	Located on the αC helix of the kinase domain; involved in regulation.
Arg808	0.23 (0.12, 0.35)	Located on the kinase domain, close to the activation loop. Involved in hydrogen bonds that stabilize the activation loop (Stamos et al., 2002).
Arg865	0.26 (0.10, 0.41)	Located on the kinase domain, close to the activation loop. Involved in hydrogen bonds that stabilize the activation loop (Stamos et al., 2002).
Gly849	0.28 (0.18, 0.37)	Located on the activation loop of the kinase domain; involved in regulation.
Leu955	0.29 (0.13, 0.42)	A part of the "LVI motif". Important for dimerization of the kinases.
Val956	0.29 (0.17, 0.42)	A part of the "LVI motif". Important for dimerization of the kinases.
Ile957	0.25 (0.08, 0.39)	A part of the "LVI motif". Important for dimerization of the kinases.
His964	0.29 (0.15, 0.43)	A putative negative regulator of EGFR's activity; located on the C-terminal domain.
Leu965	0.24 (0.10, 0.39)	A putative negative regulator of EGFR's activity; located on the C-terminal domain.
Ser967	0.23 (0.10, 0.35)	A putative negative regulator of EGFR's activity; located on the C-terminal domain.
Pro968	0.26 (0.13, 0.40)	A putative negative regulator of EGFR's activity; located on the C-terminal domain.
Ser971	0.27 (0.13, 0.38)	A putative negative regulator of EGFR's activity; located on the C-terminal domain.
Tyr974	0.30 (0.16, 0.43)	A putative negative regulator of EGFR's activity; located on the C-terminal domain.
Asp984	0.31 (0.19, 0.44)	A putative negative regulator of EGFR's activity; located on the C-terminal fragment; participates in the polar network across the interface (Figure 5).
Ala989	0.25 (0.13, 0.38)	A putative negative regulator of EGFR's activity; located on the C-terminal fragment; participates in the polar network across the interface (Figure 5).
Glu991	0.25 (0.11, 0.41)	A putative negative regulator of EGFR's activity; located on the C-terminal fragment; participates in the polar network across the interface (Figure 5).
Leu993	0.29 (0.17, 0.41)	A putative negative regulator of EGFR's activity; located on the C-terminal fragment; participates in the polar network across the interface (Figure 5).
Tyr992	0.24 (0.11, 0.38)	An autophosphorylation site, located on the C-terminal fragment; participates in the polar network across the interface (Figure 5).

A list of 24 out of 47 residues that were identified as specificity determinants (Fleishman et al., 2004a). The location of each residue in the EGFR sequence and its functional role are indicated. The trimmed means in the 95% confidence interval of correlations (r), which were calculated from 400 bootstrapping samples, are indicated, and the 95% confidence interval is shown in parentheses (see the Supplemental Data). In addition to the residues presented above, the list of specificity determinant includes the following residues: V750, Q763, L775, E780, D783, N792, V821, Q825, T830, S888, K889, I899, S901, I902, P910, K925, S933, D950, Q952, Q958, G959, D960, and E961. Their putative roles remain to be tested experimentally.

ments in the structure. In Figure 1, we suggest a model for such regulation in the ErbB family; this model is based on the available structures and is supported by a large body of biochemical and physiological data.

The role of the C-terminal domain as a modulator of kinase activity has been discussed extensively (Cadena et al., 1994; Jorissen et al., 2003), especially in the v-ErbB products (Boerner et al., 2003). Our results offer a model of the molecular mechanism for this modulation (Figure 1). In the inactive state (Figure 1, left), the EGFR extracellular domains assume a tethered structure (Ferguson et al., 2003) that hinders contact formation between the two subunits (Burgess et al., 2003). In this conformation, the extracellular domains are connected to the transmembrane helices in their inactive state (Fleishman et al., 2002), thereby maintaining the intracellular domains as a stable, inactive dimer (Figure 2A). In this state, the C-terminal domain is in contact with the kinase domain and is inaccessible to downstream substrates (Cadena et al., 1994). Ligand-induced activation of the EGFR (Figure 1, right) leads to conformational changes in the extracellular domains, allowing contact formation between the two subunits (Ogiso et al., 2002), followed by a rotation of the transmembrane helices toward their active state (Fleishman et al., 2002; Jiang and Hunter, 1999; Moriki et al., 2001). This switch in the orientation of the transmembrane helices leads to the destabilization of the inactive intracellular dimer. The C-terminal domain detaches from the kinase domain and may undergo phosphorylation, making the kinase accessible to its substrates (Moriki et al., 2001).

The structure of the GW572016 bound EGFR comprises the kinase domain and part of the C-terminal domain that is packed along the kinase domain. In this structure, the C-terminal domain partly blocks the ATP binding site (Wood et al., 2004), as in the inactive forms of the myosin light chain kinase of the Ser/Thr kinase family (Huse and Kuriyan, 2002) and the Tie2 RTK (Shewchuk et al., 2000). That the GW572016 bound EGFR structure shows an inactive conformation that is not primed for catalysis suggests that activation of the EGFR involves conformational changes within the kinase domain, in contrast to the view that the kinase domain is constitutively ready for phosphotransfer (Stamos et al., 2002). We note, however, that the new structure suggests an important role for the C-terminal domain in stabilizing an inactive conformation of the kinase domain (Wood et al., 2004); this finding is in harmony with the model of activation suggested here.

The proposed molecular model may explain the un-

derlying molecular causes of malignancy mediated by EGFRs that contain mutations in their C-terminal domain. According to the model, the transforming properties of these mutations (Boerner et al., 2003; Chang et al., 1995; Frederick et al., 2000) are due to destabilization of the inactive EGFR.

All TKs catalyze the same reaction, which is the transfer of the γ-phosphate of ATP to the hydroxyl group of tyrosine. Indeed, the active conformation of the kinase domain of most TKs is nearly identical. In contrast to the uniform active conformation, TKs differ from each other in their inactive conformations (Huse and Kuriyan, 2002). In some RTKs, as in the PDGFR family, the juxtamembrane domain serves to block the active conformation. Autophosphorylation of tyrosine residues in highly conserved juxtamembrane motifs, specific to each family, relieves autoinhibition (Griffith et al., 2004). In the case of the EGFR family, inhibition by the juxtamembrane domain is less likely, since there are no tyrosine residues in the juxtamembrane segment that can be phosphorylated.

Various regulatory mechanisms could play an important role in ensuring the signaling specificity in the TK superfamily. Accordingly, we suggest that certain amino acid substitutions in regulatory elements were sustained during evolution, leading to alterations in the regulatory mechanisms. This hypothesis is supported by the analysis of specificity determinants (Figure 7B). In the vast majority of the TKs, kinase activity is regulated through a change in the conformation of the activation loop and αC helix. Nevertheless, these regulatory regions undergo different conformational changes in different isoforms, and their inactive conformations are stabilized by fastidious means specific to each kinase family (Huse and Kuriyan, 2002). The ErbBs are further exceptional among TKs, in that the activation loop and αC helix are constitutively stable in the active conformation (Stamos et al., 2002). Our analysis of correlated mutations (Figure 7A) suggests that in order to complement the role of these known regulatory elements in maintaining an active conformation, other residues in ErbBs have evolved to keep the enzyme dormant, as in the "inactive" complex shown in Figure 2A.

We propose that members of the EGFR family utilize the unique regulatory mechanism that is presented in Figure 1. These receptors contain a long C-terminal domain that is involved in signal transmission inside the cell and is also an inherent regulator of kinase activity (Chang et al., 1995). Our results suggest that the complex between the kinase and C-terminal domains of Figure 2A is stable and biologically significant, as indicated by the large intersubunit interface, the electrostatic and geometric complementarity between the C-terminal segments and the kinases (Figures 2B–2D and 5), as well as the evolutionary correlation between specified amino acid sites (Figure 7A). This complex appears to correspond to the basal, inactive form of the receptor, as delineated above and in accordance with previous experimental data (Boerner et al., 2003; Chang et al., 1995). Although our computational analysis and the experimental data support the presence of an inactive dimer (Yu et al., 2002) and the necessity of contact between the kinase and C-terminal domains (Chang et al., 1995), the biological relevance of the crystal dimer has yet to be determined. The importance of the interface between the kinase domain and the C-terminal fragment for the regulation of EGFR activity can be tested experimentally, as delineated in the section entitled "A Network of Ion Pairs and Hydrogen Bonds at the Interface."

Our model of EGFR's regulation (Figure 1) and its relevance to cancer could be further tested by examining the properties of a short peptide analog to the C-terminal fragment. Such a peptide may have a regulatory effect on EGFR activation. For instance, in tumorigenic cells, the short peptide may associate with the kinase domain instead of the truncated C-terminal domain. This would stabilize the inactive configuration and thereby thwart the constitutive activation of the mutant receptor. Interestingly, a similar approach was applied successfully in a recent study on the PDGFR, which is also selfinhibited by direct contact with its C-terminal domain (Chiara et al., 2004). In this work, the authors showed that a soluble peptide, corresponding to the inhibitory fragment in the PDGFR C-terminal domain, delayed the activation of the receptor and inhibited the enhanced kinase activity of a C-terminal truncated PDGFR. Hence, the small peptide mimicked the role of the C-terminal fragment in regulating kinase activity (Chiara et al., 2004). It will be interesting to examine the therapeutic properties of such a peptide in the case of the EGFR.

Experimental Procedures

Biophysical and Structural Analysis

Electrostatic, solvent-accessible surface area calculations and homology modeling were carried out as described in the Supplemental Data.

Collection of Sequence Homologs and Their Alignment

A multiple-sequence alignment (MSA) of homologous kinase domains was produced by combining multiple-structure and sequence alignments to obtain high-quality alignments as described by Al-Lazikani et al. (2001) and in the Supplemental Data. This resulted in an MSA of 121 homologous sequences comprising the kinase domain and about 50 positions C-terminal to it (corresponding to positions 683–998 of the EGFR). The MSA is shown in Supplemental Figure S1 in the Supplemental Data.

Evolutionary Conservation

Evolutionary conservation scores were calculated by using the MSA and *Rate4Site*'s maximum-likelihood algorithm (Pupko et al., 2002), as implemented in the ConSurf web server (Glaser et al., 2003) (http://consurf.tau.ac.il/).

Correlated Amino Acid Substitutions

Pairs of amino acids that appear to change concomitantly during evolution within the TKs were detected by using the MSA and the *CorrMut* algorithm (Fleishman et al., 2004b). The methodological details are provided as Supplemental Data.

Specificity Determinants

Residues in the TK superfamily, which may be responsible for determining specific characteristics in different kinase families, were detected by using the MSA and the *SpecDet* algorithm (Fleishman et al., 2004a). A description of the algorithm is provided as Supplemental Data.

Supplemental Data

Supplemental Data including analysis of the electrostatic potential of representative TKs of known structure; solvent-accessible surface area calculations and homology modeling of selected TKs;

the MSA of the TK family; methodological details of the correlated mutations analysis; and a description of the algorithm used for detecting the specificity-determining residues are available at http://www.structure.org/cgi/content/full/12/12/2265/DC1/.

Acknowledgments

We thank Tony Hunter, Antony Burgess, Joseph Schlessinger, Idit Kopatz, Amit Kessel, and Saul Yankofsky for their critical comments on the manuscript, Miriam Eisenstein for her help in the identification of the dimeric conformation of the EGFR kinase, and Lisa Shewchuk for sharing the coordinates of the EGFR/GW572016 structure before their release. This study was supported by a grant from the Israel Cancer Association (ICA) and by a Research Career Development Award from the Israel Cancer Research Fund (ICRF) to N.B.-T. S.J.F. was supported by a doctoral fellowship from the Clore Israel Foundation. M.L. was supported by a Travel Scholarship from the Constantiner Institute for Molecular Genetics.

Received: June 16, 2004
Revised: September 22, 2004
Accepted: October 8, 2004
Published: December 7, 2004

References

Al-Lazikani, B., Sheinerman, F.B., and Honig, B. (2001). Combining multiple structure and sequence alignments to improve sequence detection and alignment: application to the SH2 domains of Janus kinases. Proc. Natl. Acad. Sci. USA *98*, 14796–14801.

Biswas, R., Basu, M., Sen-Majumdar, A., and Das, M. (1985). Intrapeptide autophosphorylation of the epidermal growth factor receptor: regulation of kinase catalytic function by receptor dimerization. Biochemistry *24*, 3795–3802.

Boerner, J.L., Danielsen, A., and Maihle, N.J. (2003). Ligand-independent oncogenic signaling by the epidermal growth factor receptor: v-ErbB as a paradigm. Exp. Cell Res. *284*, 111–121.

Burgess, A.W., Cho, H.S., Eigenbrot, C., Ferguson, K.M., Garrett, T.P., Leahy, D.J., Lemmon, M.A., Sliwkowski, M.X., Ward, C.W., and Yokoyama, S. (2003). An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol. Cell *12*, 541–552.

Cadena, D.L., Chan, C.L., and Gill, G.N. (1994). The intracellular tyrosine kinase domain of the epidermal growth factor receptor undergoes a conformational change upon autophosphorylation. J. Biol. Chem. *269*, 260–265.

Chang, C.M., Shu, H.K., Ravi, L., Pelley, R.J., Shu, H., and Kung, H.J. (1995). A minor tyrosine phosphorylation site located within the CAIN domain plays a critical role in regulating tissue-specific transformation by erbB kinase. J. Virol. *69*, 1172–1180.

Chantry, A. (1995). The kinase domain and membrane localization determine intracellular interactions between epidermal growth factor receptors. J. Biol. Chem. *270*, 3068–3073.

Chiara, F., Bishayee, S., Heldin, C.H., and Demoulin, J.B. (2004). Autoinhibition of the platelet-derived growth factor beta-receptor tyrosine kinase by its C-terminal tail. J. Biol. Chem. *279*, 19732–19738.

Cho, H.S., Mason, K., Ramyar, K.X., Stanley, A.M., Gabelli, S.B., Denney, D.W., Jr., and Leahy, D.J. (2003). Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature *421*, 756–760.

Dancey, J.E. (2004). Predictive factors for epidermal growth factor receptor inhibitors—The bull's-eye hits the arrow. Cancer Cell *5*, 411–415.

Ferguson, K.M., Berger, M.B., Mendrola, J.M., Cho, H.S., Leahy, D.J., and Lemmon, M.A. (2003). EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol. Cell *11*, 507–517.

Fleishman, S.J., Schlessinger, J., and Ben-Tal, N. (2002). A putative molecular-activation switch in the transmembrane domain of erbB2. Proc. Natl. Acad. Sci. USA *99*, 15937–15940.

Fleishman, S.J., Unger, V.M., Yeager, M., and Ben-Tal, N. (2004a). A C-alpha model for the transmembrane alpha-helices of gap-junction intercellular channels. Mol. Cell *15*, 879–888.

Fleishman, S.J., Yifrach, O., and Ben-Tal, N. (2004b). An evolutionarily conserved network of amino acids mediates gating in voltage-dependent potassium channels. J. Mol. Biol. *340*, 307–318.

Frederick, L., Wang, X.-Y., Eley, G., and James, C.D. (2000). Diversity and frequency of epidermal growth factor receptor mutations in human glioblastomas. Cancer Res. *60*, 1383–1387.

Gadella, T.W., Jr., and Jovin, T.M. (1995). Oligomerization of epidermal growth factor receptors on A431 cells studied by time-resolved fluorescence imaging microscopy. A stereochemical model for tyrosine kinase receptor activation. J. Cell Biol. *129*, 1543–1558.

Gamett, D.C., Tracy, S.E., and Robinson, H.L. (1986). Differences in sequences encoding the carboxyl-terminal domain of the epidermal growth factor receptor correlate with differences in the disease potential of viral erbB genes. Proc. Natl. Acad. Sci. USA *83*, 6053–6057.

Glaser, F., Pupko, T., Paz, I., Bell, R.E., Bechor-Shental, D., Martz, E., and Ben-Tal, N. (2003). ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. Bioinformatics *19*, 163–164.

Gotoh, N., Tojo, A., Hino, M., Yazaki, Y., and 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. Biochem. Biophys. Res. Commun. *186*, 768–774.

Griffith, J., Black, J., Faerman, C., Swenson, L., Wynn, M., Lu, F., Lippke, J., and Saxena, K. (2004). The structural basis for autoinhibition of FLT3 by the juxtamembrane domain. Mol. Cell *13*, 169–178.

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

Jiang, G., and Hunter, T. (1999). Receptor signaling: when dimerization is not enough. Curr. Biol. *9*, R568–R571.

Jones, S., and Thornton, J.M. (1996). Principles of protein-protein interactions. Proc. Natl. Acad. Sci. USA *93*, 13–20.

Jorissen, R.N., Walker, F., Pouliot, N., Garrett, T.P., Ward, C.W., and Burgess, A.W. (2003). Epidermal growth factor receptor: mechanisms of activation and signalling. Exp. Cell Res. *284*, 31–53.

Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. J. Appl. Crystallogr. *24*, 946–950.

Massoglia, S., Gray, A., Dull, T.J., Munemitsu, S., Kun, H.J., Schlessinger, J., and Ullrich, A. (1990). Epidermal growth factor receptor cytoplasmic domain mutations trigger ligand-independent transformation. Mol. Cell. Biol. *10*, 3048–3055.

Merritt, E.A., and Bacon, D.J. (1997). Raster3D photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

Moriki, T., Maruyama, H., and Maruyama, I.N. (2001). Activation of preformed EGF receptor dimers by ligand-induced rotation of the transmembrane domain. J. Mol. Biol. *311*, 1011–1026.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Noelle, V., Tennagels, N., and Klein, H.W. (2000). A single substitution of the insulin receptor kinase inhibits serine autophosphorylation in vitro: evidence for an interaction between the C-terminus and the activation loop. Biochemistry *39*, 7170–7177.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J.H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., and Yokoyama, S. (2002). Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell *110*, 775–787.

Pelley, R.J., Maihle, N.J., Boerkoel, C., Shu, H.K., Carter, T.H., Moscovici, C., and Kung, H.J. (1989). Disease tropism of c-erbB: effects of carboxyl-terminal tyrosine and internal mutations on tissue-specific transformation. Proc. Natl. Acad. Sci. USA *86*, 7164–7168.

Pupko, T., Bell, R.E., Mayrose, I., Glaser, F., and Ben-Tal, N. (2002). Rate4Site: an algorithmic tool for the identification of functional

regions in proteins by surface mapping of evolutionary determinants within their homologues. Bioinformatics *18*, S71–S77.

Raines, M.A., Maihle, N.J., Moscovici, C., Moscovici, M.G., and Kung, H.J. (1988). Molecular characterization of three erbB transducing viruses generated during avian leukosis virus-induced erythroleukemia: extensive internal deletion near the kinase domain activates the fibrosarcoma- and hemangioma-inducing potentials of erbB. J. Virol. *62*, 2444–2452.

Riedel, H., Schlessinger, J., and Ullrich, A. (1987). A chimeric, ligand-binding v-erbB/EGF receptor retains transforming potential. Science *236*, 197–200.

Robinson, H.L., Tracy, S.E., Nair, N., Taglienti-Sian, C., and Gamett, D.C. (1992). Characterization of an angiosarcoma-inducing mutation in the erbB oncogene. Oncogene *7*, 2025–2030.

Schlessinger, J. (2000). Cell signaling by receptor tyrosine kinases. Cell *103*, 211–225.

Schlessinger, J. (2003). SIGNAL TRANSDUCTION: autoinhibition control. Science *300*, 750–752.

Serrano, L., Horovitz, A., Avron, B., Bycroft, M., and Fersht, A.R. (1990). Estimating the contribution of engineered surface electrostatic interactions to protein stability by using double-mutant cycles. Biochemistry *29*, 9343–9352.

Sheinerman, F.B., Norel, R., and Honig, B. (2000). Electrostatic aspects of protein-protein interactions. Curr. Opin. Struct. Biol. *10*, 153–159.

Shewchuk, L.M., Hassell, A.M., Ellis, B., Holmes, W.D., Davis, R., Horne, E.L., Kadwell, S.H., McKee, D.D., and Moore, J.T. (2000). Structure of the Tie2 RTK domain: self-inhibition by the nucleotide binding loop, activation loop, and C-terminal tail. Structure *8*, 1105–1113.

Stamos, J., Sliwkowski, M.X., and 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.

Walker, F., Kato, A., Gonez, L.J., Hibbs, M.L., Pouliot, N., Levitzki, A., and Burgess, A.W. (1998). Activation of the Ras/mitogen-activated protein kinase pathway by kinase-defective epidermal growth factor receptors results in cell survival but not proliferation. Mol. Cell. Biol. *18*, 7192–7204.

Wedegaertner, P.B., and Gill, G.N. (1992). Effect of carboxyl terminal truncation on the tyrosine kinase activity of the epidermal growth factor receptor. Arch. Biochem. Biophys. *292*, 273–280.

Wood, E.R., Truesdale, A.T., McDonald, O.B., Yuan, D., Hassell, A., Dickerson, S.H., Ellis, B., Pennisi, C., Horne, E., Lackey, K., 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.

Yarden, Y., and Schlessinger, J. (1987). Epidermal growth factor induces rapid, reversible aggregation of the purified epidermal growth factor receptor. Biochemistry *26*, 1443–1451.

Yarden, Y., and Sliwkowski, M.X. (2001). Untangling the ErbB signalling network. Nat. Rev. Mol. Cell Biol. *2*, 127–137.

Yu, X., Sharma, K.D., Takahashi, T., Iwamoto, R., and Mekada, E. (2002). Ligand-independent dimer formation of epidermal growth factor receptor (EGFR) is a step separable from ligand-induced EGFR signaling. Mol. Biol. Cell *13*, 2547–2557.