A Model for the Activation of the Epidermal Growth Factor Receptor Kinase: Involvement of an Asymmetric Dimer? †

Leo C. Groenen,* Francesca Walker, Antony W. Burgess, and Herbert R. Treutlein

Ludwig Institute for Cancer Research, P.O. Box 2008, Royal Melbourne Hospital, and Cooperative Research Centre for Cellular Growth Factors, PO Royal Melbourne Hospital, Victoria 3050, Australia

Received June 13, 1996; Revised Manuscript Received January 22, 1997[⊗]

ABSTRACT: The binding of epidermal growth factor (EGF) to its receptor leads to receptor dimerization, which activates the intracellular kinase domain. Homology models of the inactive and active forms of the EGF-receptor kinase domains have been derived, and these models suggest that the active form can be stabilized by the interaction of helix C and the surrounding area in one receptor monomer with one of two possible complementary surfaces on a second receptor monomer. Both hydrophobic interaction sites are strongly conserved within the EGF-receptor family but not in other tyrosine kinases. Two of the three predicted kinase dimers are symmetric; the other is asymmetric and is predicted to contain only one active kinase. One of the symmetric models and the asymmetric model would account for the effects of two mutations in helix C (Y740F and V741G) on kinase activity. They also provide an explanation for previously reported dominant negative mutants of the EGF receptor and have interesting implications for the signaling through homo- and heterodimers of the family members: EGF receptor, erbB2, erbB3, and erbB4.

The epidermal growth factor receptor (EGFR)[1] and its family members erbB2 (HER2 or neu) and erbB4 (HER4) are ligand-activated tyrosine kinase receptors (Ullrich & Schlessinger, 1990; Carraway & Cantley, 1994). The fourth known family member, erbB3 or HER3, has seriously impaired tyrosine kinase activity (Guy et al., 1994). The EGF receptor consists of an extracellular ligand-binding domain, a transmembrane region, and a cytoplasmic domain containing a juxtamembrane region, the catalytic domain, and a C-terminal tail with at least five tyrosine phosphorylation sites [reviewed in Carraway and Cantley (1994), Carter and Kung (1994), and Boonstra et al. (1995)]. Epidermal growth factor (EGF) can bind with both low and high affinity to the extracellular domain of the EGF receptor (Livneh et al., 1986). Binding of EGF leads to receptor dimerization (Cochet et al., 1988) and simultaneous kinase activation (Yarden & Schlessinger, 1987; Canals, 1992). The first substrate of the EGFR kinase appears to be the EGF receptor itself, resulting in the autophosphorylation of a number of tyrosine residues in the C-terminal tail (Carter & Kung, 1994; Boonstra et al., 1995). Subsequently, many signaling events are triggered inside the cell, partly through the binding of cytoplasmic proteins to the phosphorylated tyrosine residues (Carter & Kung, 1994; Boonstra et al., 1995). Overexpression of the EGF receptor is associated with tumorigenic transformation when the cells are stimulated with EGF or transforming growth factor α (Carter & Kung, 1994). Several EGFR mutants are known that cause transformation without dependence on EGF in certain cell types (Carter & Kung, 1994). Interestingly, while mice can tolerate an EGF receptor with significantly impaired kinase activity (Luetteke et al., 1994; Fowler et al., 1995), loss of the receptor is lethal for most mouse embryos (Miettinen et al., 1995; Sibilia & Wagner, 1995; Threadgill et al., 1995).

Despite considerable investigation it is still not clear how the binding of EGF to the extracellular domain and subsequent dimerization of the EGF receptors activate the kinase domain inside the cell. The dimerization of the EGF receptor generally coincides with an increase of kinase activity (often only indirectly inferred as an increased tyrosine phosphorylation on the EGF receptor), irrespective of whether dimerization is induced by binding of EGF to the full-length receptor on cells (Cochet et al., 1988) or in solution (Boni-Schnetzler & Pilch, 1987; Canals, 1992), binding of bivalent antibodies to full-length receptors (Spaargaren et al., 1991), or formation of a non-native disulfide bond between two receptors (Sorokin et al., 1994). EGF receptors missing the extracellular domain form dimers spontaneously and are constitutively autophosphorylated (Chantry, 1995), and the intracellular domain in solution can be activated by the addition of protein-precipitating agents, which promote dimerization or oligomerization (Mohammadi et al., 1993). As it had been previously shown that EGF receptors truncated at residue 972, missing most of the C-terminal tail, are activated normally by EGF and phosphorylate intracellular substrates (Walton et al., 1990), it appears that the C-terminal domain does not play a role in the activation of the EGF receptor. We propose that the interaction between the two kinase domains in the EGF-receptor dimer is

† This work was supported in part by Grant 950826 from the National Health and Medical Research Council of Australia to F.W. and A.W.B.

* Author to whom correspondence should be addressed at the Ludwig Institute for Cancer Research. FAX: +61 (3) 9347 1938. E-mail: groenen@ludwig.edu.au.

[⊗] Abstract published in *Advance ACS Abstracts,* March 1, 1997.

[1] Abbreviations: AMP, adenosine monophosphate; ATP, adenosine triphosphate; bFGFR, basic fibroblast growth factor receptor; cAPK, cyclic AMP-dependent protein kinase; CDK2, cyclin-dependent kinase 2; CK1, casein kinase 1; EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; HGFR, hepatocyte growth factor receptor; IRK, insulin receptor kinase; MAP kinase, mitogen-activated protein kinase; MNDO, modified neglect of diatomic overlap; NGFR, nerve growth factor receptor; PDB, Protein Data Bank; PDGFR-B, platelet-derived growth factor receptor B; PhK, phosphorylase kinase.

S0006-2960(96)01414-6 CCC: $14.00

FIGURE 1: Amino acid sequence alignment of the kinase domains of the EGF receptor (EGFR), insulin receptor (IRK) and cyclic AMP-dependent kinase (cAPK). Residues identical with EGFR are boldfaced. Above the sequences of IRK and cAPK the α-helices (letters) and β-strands (numbers) are indicated. Residues used for coordinate assignment for the active form are underlined, and the six newly built loops L1−L6 in the active form are indicated above the EGFR sequence. The residues forming the extension of helix I in EGFR are indicated by an α above the EGFR sequence.

responsible for activation of the receptor. We have used homology modeling and molecular dynamics to develop three-dimensional models for the EGFR kinase, which has led to three models of the active EGFR kinase dimer, both symmetric and asymmetric.

MATERIALS AND METHODS

Homology Models of Inactive and Active EGF-Receptor Kinase Domains. The model of the inactive tyrosine kinase domain is based on the 2.1 Å resolution X-ray structure of the kinase domain of the insulin receptor (IRK; Hubbard et al., 1994), with which it shares 34% sequence identity [entry 1irk in the Brookhaven Protein Data Bank (PDB); Bernstein et al., 1977]. The alignment by visual inspection of the amino acid sequences of both kinase domains is straightforward and requires four deletions and one insertion in the human EGFR sequence (Figure 1). The model of the active form was built on the basis of the X-ray structures of IRK and of cyclic AMP-dependent protein kinase (cAPK) in a complex with ATP and a peptide inhibitor (PDB entry 1atp, 2.2 Å resolution, 21% sequence identity with EGFR; Zheng et al., 1993a,b). The sequences of cAPK and IRK were aligned on the basis of their three-dimensional structures (Figure 1). The N- and C-terminal lobes of IRK were superimposed separately onto the respective lobes of cAPK, using the core β-sheet in the N-terminal lobes of IRK and cAPK (residues equivalent to 690−695, 701−705, 718−722, 752−757, and 763−767 in EGFR) and the residues equivalent to 774−780, 788−805, 871−884, 896−903, and 917−926 in EGFR in the C-terminal lobe, corresponding to α-helices D, E, F, G, and H (Figure 1).

For the initial model building the programs Homology and InsightII (Molecular Simulations Inc., San Diego, CA) were used. Atomic coordinates for residues of the EGF-receptor kinase were transferred from IRK, except for residues in the interface between the two lobes in the active model, for which the cAPK coordinates were used (Figure 1), and certain loops. In both models helix I was extended to residue 950 to reflect the local sequence similarity between EGFR and casein kinase 1, in which this helix is also extended (Xu et al., 1995). Loops with insertions or deletions and loops with very unfavorable side-chain contacts or proline conformations were built on the basis of templates identified from a loop search procedure in Homology. This procedure searches a subset of the PDB for stretches of amino acids of identical length to the loop, for which the conformation of the anchor residues most closely resembles that of the loop anchor residues. The default subset supplied with Homology (version 2.3.5) was supplemented with a list of structures in the PDB that are less than 30% identical in sequence (Hobohm et al., 1992). Equivalent structures and duplicate chains were removed. The ten templates found in each search were examined for appropriate ϕ, ψ and peptide bond dihedral angles, amino acid composition, and steric clashes with the rest of the structure. A suitable template was normally found after repeated searches, in which the length of the loop and the number of anchor residues at each end were varied. The loops in the active structure for which templates were used are indicated in Figure 1. For the inactive form these loops are L1, L2, and L4, as in the active form, and residues 838−841 and 848−851. After assignment of the coordinates the structure was visually checked for steric clashes involving side chains, and where necessary the conformations of side chains were adjusted to better fitting rotamers using an automated procedure in Homology. One ATP molecule, two Mg^{2+} ions, and three water molecules coordinated to the ions were included in the active model. Furthermore, 19 and 27 crystal water molecules were included in the active and inactive models, respectively, which either resided in cavities inside the protein or were associated with conserved loops on the surface of the protein.

Refinement of Kinase Models. The initial models were refined by minimization and dynamics runs using the program X-PLOR (Brünger, 1992). The models were surrounded by a 5 Å layer of TIP3P water containing *ca.* 1600 molecules. For all calculations the all-atom CHARMm22 force field was used (Brooks et al., 1983; MacKerell et al., 1995), which is distributed with X-PLOR. Missing bond, angle, and dihedral parameters for the triphosphate group were estimated from the X-ray structure and similar parameters in the force field. Atomic charges for the triphosphate group were approximated from a one-point MNDO calculation (using the program MOPAC; Stewart, 1990) on the conformation of ATP as found in the cAPK structure (Zheng et al., 1993b). The energy of each system was minimized in four stages using the conjugate gradient method (Brünger, 1992). First, the protein structure was fixed and only the solvent water molecules were allowed to move in 500 minimization steps. Then, for another 500 steps, all loops and all side chains were released and the backbone atoms of the two residues on either side of the loops were harmonically tethered to their initial positions with a force constant of 1.0 kcal mol^{-1} $Å^{-2}$. In this stage and in the next, dihedral restraints were used to restrict the ϕ and/or ψ angles of non-glycine residues with positive ϕ angles in the template structures ($\phi = +60 \pm 30°$ and $\psi = +40 \pm 50°$), of glycine residues in loops ($\phi = +90 \pm 60°$), and of residues that tended to move to positive ϕ angles but are not expected to ($\phi = -90 \pm 60°$). A square-well function (Brünger, 1992) was used to restrain hydrogen bonds in β-sheets, α-helices, and 3_{10}-helices and conserved hydrogen bonds involving side-chain atoms (e.g., the salt bridge between E860 and R934, the hydrogen bonds between the side chain of D872 and the amide hydrogen atoms of H811 and R812 and between the protein and Mg_2ATP in the active model). In the third and fourth stages all atoms were allowed to move, and minimization was continued until the energy gradient was less than 0.3 and 0.1 kcal mol^{-1} $Å^{-1}$, respectively. In these stages and in the dynamics runs for the active kinase, Mg_2ATP and coordinated water molecules were harmonically restrained to their initial positions with a force constant of 50.0 kcal mol^{-1} $Å^{-2}$, mainly because no well-tested force field parameters are available for Mg^{2+} ions and the triphosphate group.

In the dynamics runs the geometry of individual solvent molecules was kept fixed using the SHAKE algorithm (Brünger, 1992), as were the covalent bonds to hydrogen atoms of the protein. The dynamics runs used a time step of 2 fs, heat bath coupling constant of 100 ps^{-1}, and a group-based nonbonded cutoff distance of 7.5 Å with a switch function for the van der Waals interactions and a shift function for the electrostatic interaction between 7.0 and 7.5 Å (Brünger, 1992). The systems were heated in steps of 50 K to a final temperature of 300 K in 0.5 ps, after which dynamics was continued for 50 ps at 300 K. After each interval of 5 ps the structure was minimized and stored. The stereochemical quality of each structure was checked with the program Procheck (Laskowski et al., 1993), and the quality of side-chain packing was evaluated with ProsaII (Sippl, 1993) and the 3D-profiles method (Lüthy et al., 1992). The models presented here are the minimized structures after 35 ps of dynamics, which are representative for the fluctuating structures after the initial 10−20 ps of dynamics.

Models for Kinase Dimers. The initial models for the kinase dimer were constructed by manual docking of the unsolvated monomers (35 ps structures), using the InsightII program. For both the symmetric and asymmetric dimer only four to six orientations were considered because of the elongated nature of the hydrophobic patches around helix C and around helix H. Each structure was first optimized by rigid body minimization, in which side chains and loop backbones at the interface were partitioned into small nonlinear rigid bodies and the backbone of helix C and the remaining bulk of each monomer were treated as three separate rigid bodies. These calculations were executed in vacuum (dielectric constant of four) with the united-atom CHARMm19 force field (Brooks et al., 1983), using the X-PLOR program. The two monomers were then replaced by the 35 ps structures of the active or inactive kinase domain with their water shell, and overlapping water molecules within 3 Å of a heavy atom of the other monomer plus water shell were removed. This complex was optimized by an initial energy minimization and a 25 ps dynamics run with minimization every 5 ps using the same parameters, force field, and restraints as for the monomer models. The presented dimers have the highest buried surface areas after 20 of the 25 ps of dynamics. Solvent-accessible surface areas were calculated with X-PLOR using a 1.4 Å solvent probe and only the non-hydrogen atoms of the proteins.

The coordinates of the active and inactive kinase models and of the three dimer models can be obtained from the corresponding author and will be posted on the World Wide Web (URL: http://liba.ludwig.edu.au/).

^{125}I-Labeled EGF (^{125}I-EGF) Binding. Iodination of murine EGF and equilibrium binding studies were performed as described (Walker et al., 1990). The data were plotted and equilibrium binding constants were derived using the LIGAND program (Munson & Rodford, 1980).

In Vitro Protein Kinase Assays. Receptor protein was immunopurified from Baf/BO_3 cells transfected with the wild-type, Y740F, or V741G EGF-receptor constructs with an antibody to the extracellular domain of the EGF receptor (mAb 528). Autokinase activity was measured as described (Fowler et al., 1995). Quantitation of the radioactivity incorporated into the receptor proteins, and of the amounts of receptors, was carried out on a Molecular Dynamics Phosphoimager.

RESULTS

Three-dimensional models of the kinase domain of the human EGF receptor (residues 687−951) in both the inactive and active forms were derived by homology-based modeling. Currently, crystal structures of eight different protein kinases have been solved and published (Zheng et al., 1993a; De Bondt et al., 1993; Zhang et al., 1994; Hu et al., 1994; Hubbard et al., 1994; Owen et al., 1995; Xu et al., 1995; Goldberg et al., 1996). The EGF-receptor kinase domain shares the highest sequence identity (34%) with the only tyrosine kinase among these, the insulin receptor kinase (IRK; Hubbard et al., 1994). The insulin receptor kinase was crystallized in the inactive, unphosphorylated form and has no bound ATP. This structure served as the template for the model of the inactive form of the EGFR kinase, based on the sequence alignment shown in Figure 1. The model of the active form was based on the three-dimensional

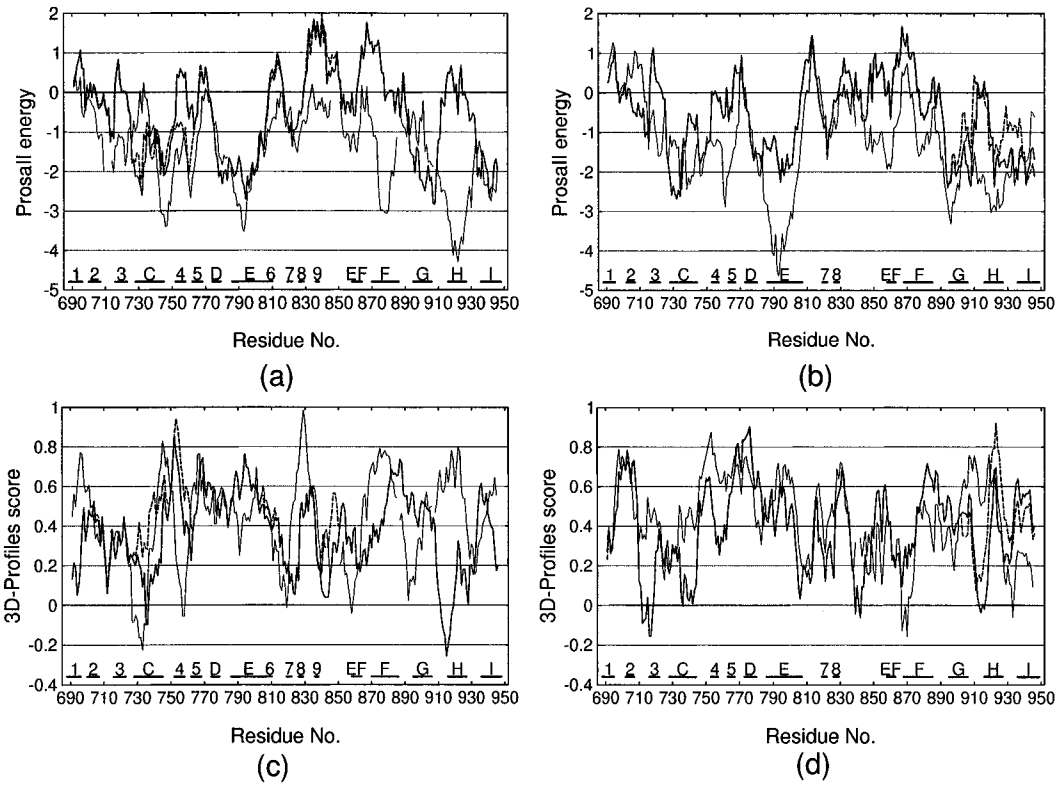

FIGURE 2: Quality of chain fold and side-chain packing. The ProsaII energy profiles shown in panels a and b and the 3D-Profiles score profiles in panels c and d are averaged over a window of ten residues. A low ProsaII energy or a high 3D-Profiles score indicates a correctly folded amino acid sequence. Panels a and c show the models of the active form of the EGFR-kinase domain as a monomer (thick line) and in the asymmetric dimer (dashed thick line) in comparison with cAPK (thin line). In panels b and d the inactive models as a monomer (thick) and in the asymmetric dimer (dashed) are shown in comparison with IRK (thin line).

structures of both insulin receptor kinase and cyclic AMP-dependent kinase (cAPK). The X-ray structure of cAPK is one of three crystal structures of an active kinase, the others being phosphorylase kinase (PhK; Owen et al., 1995), and casein kinase 1 (CK1; Xu et al., 1995). It was chosen as a template because it has higher sequence identity (21%) with EGFR than either PhK or CK1 (19% and 16%, respectively) and is the only structure that contains both ATP and a substrate-like molecule, the inhibitor peptide (Zheng et al., 1993a,b). For the model building the N- and C-terminal lobes of IRK were superimposed separately onto the structure of cAPK and used as templates as such, but the loops that are in contact with ATP (residues 694−702, 767−773, and 808−854) or are in close contact with these loops (residues 744−752 and 863−869) were derived from the cAPK structure (Figure 1).

The stereochemical quality of the models for the inactive and active EGFR kinase domains after refinement by molecular dynamics and energy minimization was assessed with the program Procheck (Laskowski et al., 1993). Both models have 84% of the residues in the most favored region of the Ramachandran plot, as compared with 91% for the equivalent fragment of IRK and 90% for cAPK. There are three to four residues with ϕ,ψ angles outside the "most favored and additionally allowed" regions in the Ramachandran plot (three for IRK and none for cAPK) and one (R812, which has departed from its initial left-handed α-helical conformation) in a disallowed region. The overall *G*-factors calculated using Procheck, which is a combined measure for the quality of main-chain and side-chain dihedral angles, distortion around the C_α atom, and the energy of main-chain hydrogen bonds, are −0.20 for the active model and −0.17 for the inactive model. These values are slightly lower (worse) than for the template structures (0.31 for IRK, −0.16 for cAPK). A small problem is found with the N−C_α−C bond angles, of which around 30 are more than 10° less than the "ideal" values used by Procheck. These ideal values, 112.5° for glycine and 111.2° for non-glycine residues, are 4.2−5.5° higher than the equilibrium value of 107° defined in the CHARMm22 force field.

Several methods have been proposed to evaluate the quality of the packing of residues in a protein [reviewed in Sali (1995)]. The results of an analysis with two of these methods, ProsaII (Sippl, 1993), which calculates an energy profile using empirical surface and C_β−C_β interaction potentials, and 3D-Profiles (Lüthy et al., 1992), which takes into account the type of secondary structure element a residue is part of, its solvent accessibility, and the number of polar atoms surrounding its side chain, are shown in Figure 2. Considering that X-ray structures and modeled structures alike contain a number of regions with unfavorable energies (higher than one energy unit in the ProsaII plots, lower than zero in the 3D-Profiles plots), the models of both the active and inactive EGFR kinase seem to be packed reasonably well. In some areas both methods agree that the fold is less favorable in the models than in the template structures. These are the loop between β-strands 4 and 5 (possibly because hydrophobic L758 is on the surface), the loop after β-strand 8, the beginning of helix F, and especially the sequence around helix H. In this last area five surface residues are hydrophobic in EGFR but polar in IRK (I914, I917, M921, V924, and M928; Figure 3). These residues form a continuous hydrophobic patch on the surface that might form an interaction site either with another part of

```
          730       740       755       835       920       930
          .+....:....*....:  ..:..*.  ....:.*.  ..*:.*..+*...*:..*.+
EGFR      PKANKEILDEAYVMASV  LLGICLT  DFGLAKLL  PPICTIDVYMIMVKCWMID
erbB2     PKANKEILDEAYVMAGV  LLGICLT  DFGLARLL  PPICTIDVYMIMVKCWMID
erbB3     RQSFQAVTDHMLAIGSL  LLGLCPG  DFGVADLL  PQICTIDVYMVMVKCWMID
erbB4     PKANVEFMDEALIMASM  LLGVCLS  DFGLARLL  PPICTIDVYMVMVKCWMID
          cccccccccccccc     44444              HHHHHHHHHH
IRK       LRERIEFLNEASVMKGF  LLGVVSK  DFGMTRDI  PDNCPERVTDLMRMCWQFN
PDGFR-B   SSEKQALMSELKIMSHL  LLGACTK  DFGLARDI  PAHASDEIYEIMQKCWEEK
bFGFR     EKDLSDLISEMEMMKMI  LLGACTQ  DFGLARDI  PSNCTNELYMMMRDCWHAV
NGFR      -SARQDFQREVELLTML  FFGVCTE  DFGMSRDI  PRACPPEVYAIMRGCWQRE
HGFR      IGEVSQFLTEGIIMKDF  LLGICLR  DFGLARDM  PEYCPDPLYEVMLKCWHPK
EPH       GGQWWNFLREATIMGQF  LEGVVTK  DFGLTRLL  PVDCPAPLYELMKNCWAYD
            ccccccccccccc    444444     99       HHHHHHHHHH
cAPK      LKQIEHTLNEKRILQAV  LEFSFKD  DFGFAKRV  PSHFSSDLKDLLRNLLQVD
PhK       QELREATLKEVDILRKV  LKDTYET  DFGFSCQL  WDDYSDTVKDLVSRFLVVQ
CDK2      EGVPSTAIREISLLKEL  LLDVIHT  DFGLARAF  VPPLDEDGRSLLSQMLHYD
ERK2      QTYCQRTLREIKILLRF  INDIIRA  DFGLARVA  FPNADSKALDLLDKMLTFN
CK1       RSDAPQLRDEYRTYKLL  VYYFGQE  DFGMVKFY  CAGFPEEFYKYMHYARNLA
twitchin  ESDKETVRKEIQTMSVL  LHDAFED  DFGLTAHL  FSGISEDGKDFIRKLLLAD
```

FIGURE 3: Amino acid sequence alignment of helix C, β-strands 4 and 9, and helix H of the kinase domains of human EGFR, erbB2, erbB3, and erbB4, selected tyrosine kinase receptors, and serine/threonine kinases of which the X-ray structure has been solved. The positions of the helices and strands in IRK and cAPK are indicated by letters and numbers above their sequences. The highlighted residues are conserved in tyrosine kinases (Hanks & Quinn, 1991). The hydrophobic residues that are conserved in the EGFR family only are indicated with an asterisk. The numbering is for EGFR.

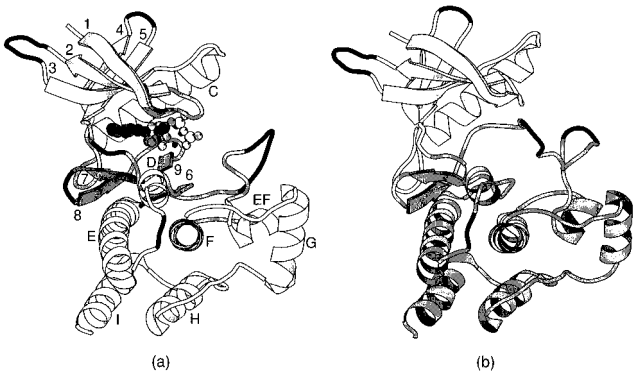

FIGURE 4: Ribbon diagrams of the models for the active (a) and inactive form (b) of the kinase domain of the EGF receptor. In the inactive form (b) the N- and C-terminal lobes are colored white and gray. In the active form (a) the residues for which the coordinates were transferred from IRK are colored white and those from cAPK are in gray. Also shown is the bound ATP molecule in a ball-and-stick representation. Letters and numbers indicate α-helices and β-strands, respectively. In both diagrams the loops obtained from template searches are black. The drawing was prepared with Molscript (Kraulis, 1991).

the EGF receptor or with another protein. It is unlikely that this helix is built in a wrong orientation because the hydrophobic residues that form the interface of helix H (V919, I922, M923, C926, and W927) with the rest of the structure are strongly conserved between the EGFR kinase and IRK and indeed other tyrosine kinases (Figure 3).

The ribbon diagrams of the models for the inactive and active EGFR kinase are shown in Figure 4. In both diagrams the C-terminal lobes are shown in the same orientation. Comparison of the inactive and active kinase shows that the orientation of the N-terminal lobe relative to the C-terminal lobe is not very different in the two models, in contrast to the differences between IRK and cAPK (Hubbard et al., 1994). This is due to a rearrangement of the N-terminal lobe in the inactive model during the molecular dynamics calculations. The major difference between the active and inactive EGFR kinases is found in the conformation of the long regulatory loop (residues 830−855), which is highlighted in Figure 5. In the inactive kinase this loop traverses both the ATP binding site, with the hydrophobic side chain of F832 partly filling the hydrophobic cavity required for the

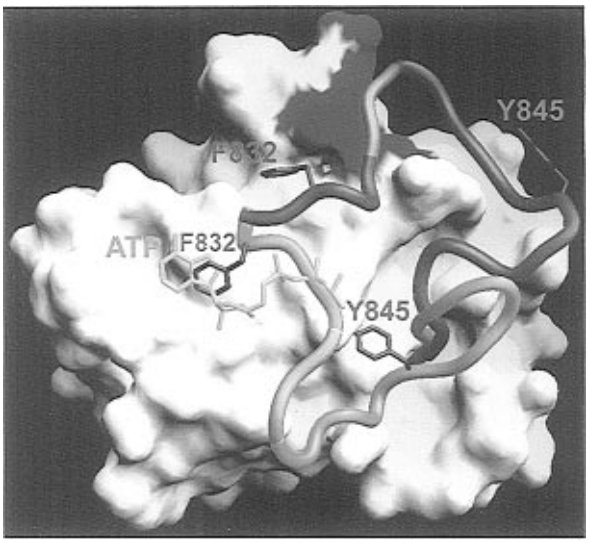

FIGURE 5: Conformational change of the regulatory loop (residues 830−855). The C-terminal lobes of the active and inactive kinases are superimposed, and the molecular surface of residues 770−829 and 856−951 of the active kinase is shown, together with a worm presentation of the regulatory loop (residues 830−855) of active (magenta) and inactive (cyan) kinase, in a view perpendicular to the bound ATP (in yellow) in the active form. Residues F832 and Y845 in the active (red) and inactive (blue) forms are highlighted. The green patch on the surface shows the location of residues 808−810 of β-strand 6, and the orange patches in the worms indicate residues 836−838 of β-strand 9. The drawing was prepared with Grasp (Nicholls et al., 1991).

adenine moiety, and the substrate site where Y845 occupies the site for the tyrosine residue to be phosphorylated. This is similar to the structure of the inactive form of IRK (Hubbard et al., 1994). In the active form Y845 is on the surface of the protein, and the side chain of F832 is predicted to occupy a different hydrophobic cavity flanked by residues V741, V750, L804, and H811. Glutamate 842 of the regulatory loop forms an ion pair with R812, and a short β-sheet is formed consisting of strands 6 and 9 (Figures 4a and 5). As a result of the reorientation of the regulatory loop both the ATP binding site and the substrate binding site are accessible in the active form of the kinase.

The three-dimensional models of the inactive and active kinase domains point to a mechanism for the activation of the EGF-receptor kinase through dimerization. In quiescent cells EGF receptors have a basal kinase activity (Livneh et al., 1987). This implies that some of the EGF receptors have an active kinase domain and that there is an equilibrium between inactive and active EGF-receptor kinase domains. The binding of EGF to the extracellular domain results in dimerization of EGF receptors (Cochet et al., 1988) and a concomitant shift in the equilibrium toward receptors with an active kinase conformation. We suggest that the EGF-receptor dimer involves contacts not only between the two extracellular domains but also between the kinase domains (Chantry, 1995) and, further, that the active conformation of the regulatory loop of one kinase domain is stabilized by interaction with the second kinase domain.

Recently, a mutant mouse EGF receptor was described, in which valine 743 in helix C is substituted by glycine (Luetteke et al., 1994; Fowler et al., 1995). We have characterized the equivalent mutation in the human EGF receptor, V741G, and also a new mutation, Y740F. The V741G mutant has only low-affinity binding sites for EGF

Table 1: ^{125}I-EGF Binding and in Vitro Autophosphorylation of EGF-Receptor Mutants[a]

receptor[b]	K_d1 (pM)[c]	K_d2 (nM)[c]	R1 (10^3/cell)[d]	R2 (10^4/cell)[d]	fold stimulation by EGF[e]	specific ^{32}P incorporation (relative to WT)[f]
WT	8.9	1.3	4.1	4.3	1.9	1
Y740F	12	1.1	1.3	3.5	1.8	0.11
V741G		1.3		6.8	1.5	0.003

[a] Wild-type and mutant receptors were transfected in Baf/BO_3 cells. Equilibrium dissociation constants and receptor numbers were determined by ^{125}I-EGF binding followed by analysis with the LIGAND program. Autokinase activity was determined with or without EGF in immunoprecipitates in the presence of [γ-^{32}P]ATP and $MnCl_2$. Specific ^{32}P incorporation was calculated as the ratio between radioactivity incorporated into the receptors and the amount of receptor protein, as assessed by Western blotting and densitometry. [b] WT is the wild-type EGF receptor; Y740F and V741G are two mutant EGF receptors. [c] K_d1 and K_d2 are the high- and low-affinity binding constants. [d] R1 and R2 are the number of high- and low-affinity receptors per cell. [e] Ratio of ^{32}P incorporation in the presence and absence of EGF. [f] Measured in the presence of EGF.

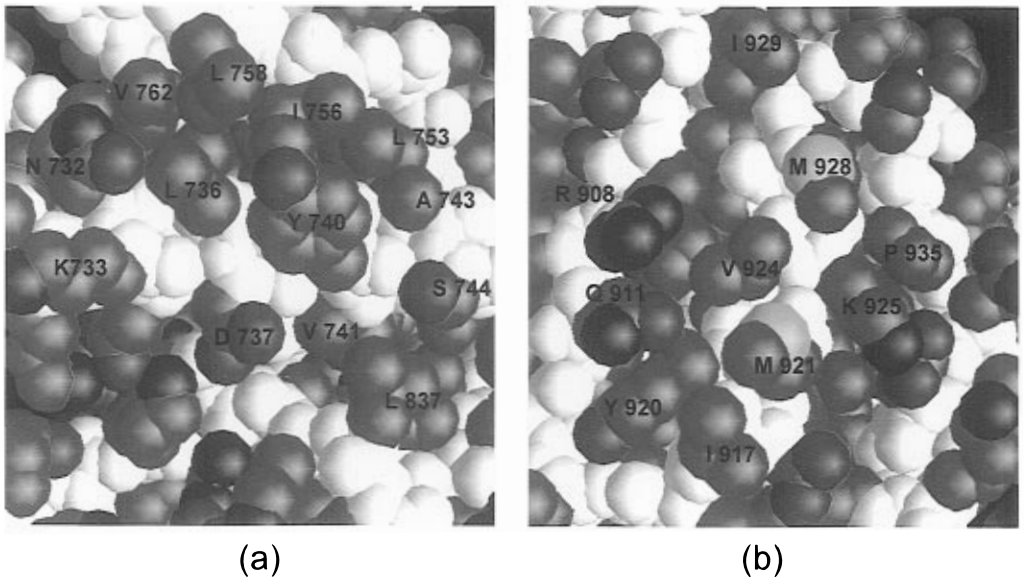

FIGURE 6: Van der Waals representation of the surface residues around helix C (a) and helix H (b) of the active and inactive EGFR kinase, respectively. The side-chain atoms are colored according to atom type: C, green; O, red; N, blue; and S, yellow. Backbone atoms are in white. Drawing prepared with Grasp (Nicholls et al., 1991).

and a 300-fold reduced EGF-dependent autophosphorylation in vitro (Table 1). For the Y740F mutant the high-affinity binding is still intact, albeit a smaller number of binding sites is present, and the in vitro autophosphorylation is reduced *ca.* 9-fold (Table 1). Thus, both mutations significantly reduce the kinase activity of the EGF receptor.

In the model of the active kinase, residues F832 and L837 of the regulatory loop are packed against residue V741 of helix C. Leucine 837 lies on the surface of the kinase domain and forms a hydrophobic patch with other residues from helix C and β-strand 4 (Figure 6a). Some of the hydrophobic residues in this patch, L736 and I756, are conserved in most kinases and form part of the interfaces of helix C with β-strand 4 and with the C-terminal lobe (Figure 3). Other hydrophobic residues in this patch, L758, L837, and also Y740, are strongly conserved within the EGF-receptor family but not in most other kinases (Figure 3). This hydrophobic patch seemed a conspicuous site for interaction between two EGF-receptor kinases. Therefore, we tried to manually dock two active kinase domains together to form a symmetric dimer, in which the areas around helix C in both kinases interact with each other. Several alternatives were found in which the kinases were shifted relative to each other, requiring different orientations of tyrosine residue 740 in the middle of the surface to prevent severe steric clashes. We also considered the hydrophobic surface patch around helix H as the second interaction site (Figure 6b). The five hydrophobic residues in this patch are completely conserved in the EGF-receptor family, whereas in other classes of kinases hydrophobic residues are only found sporadically in these positions (Figure 3). Attempts to manually dock the surface around helix H of an inactive kinase (this kinase is not interacting via its helix C, so that the active conformation of its regulatory loop is not being stabilized and is therefore inactive) to the surface around helix C of an active kinase showed that the two surfaces are surprisingly complementary, and it proved straightforward to find orientations with few side-chain clashes and with the two hydrophobic surfaces closely packed together.

The complexes were refined by rigid-body energy minimization, molecular dynamics, and energy minimization, and the two symmetric dimers and one asymmetrical dimer with the highest buried solvent-accessible surface area are presented (Figure 7). The symmetric dimers are not completely symmetric because no symmetry restraints were imposed on the two kinase molecules during the refinement. The total surface area that is buried upon formation of the two symmetric dimers is 2073 and 2083 Å^2, and 2087 Å^2 for the asymmetric dimer, considerably more than found on average (1685 Å^2) in protein homodimers (Jones & Thornton, 1996). In all three dimers the ATP binding sites and the substrate binding sites are fully accessible. Expanded views of the dimer interfaces are shown in Figure 8. In the first symmetric dimer (Figure 8a) the hydrophobic patches around

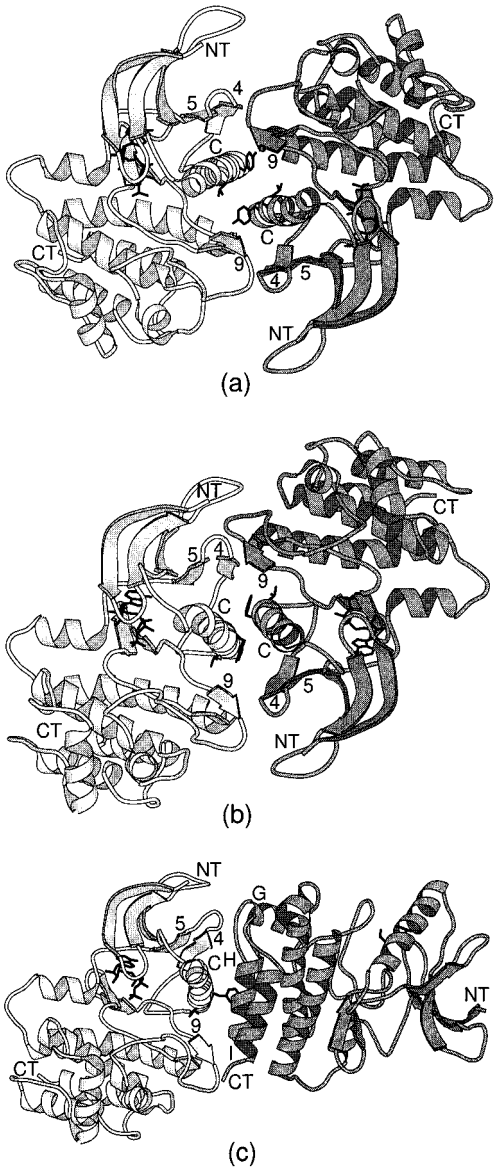

FIGURE 7: Ribbon diagrams of two symmetric (a, b) and an asymmetric kinase dimer (c). ATP and the side chains of residues Y740 and V741 are shown as thick lines. The kinase on the left is approximately 90° rotated around the vertical axis running in the paper relative to the orientation in Figure 4. Letters and numbers indicate α-helices and β-strands, respectively. The N- and C-termini are indicated by NT and CT. In the asymmetric kinase dimer the inactive kinase is shown on the right. Drawing prepared with Molscript (Kraulis, 1991).

helix C of either kinase only partly interact with each other. Leucine 758 is closest to R836 of the regulatory loop, and L837 lies between Y740 and S744. The hydroxyl group of Y740 is not involved in a hydrogen bond but is *ca*. 4.5 Å away from the carboxylate of D737. Arginine 808 of either kinase forms an ion pair with D746 of the other kinase. In the second symmetric dimer (Figure 8b) the hydrophobic patches around helix C on each kinase are well aligned, with L736, I756, and L758 forming a new hydrophobic cluster with Y740 and L837 of the other kinase. There is no obvious hydrogen bond partner for the hydroxyl group of Y740. Residues R808 and E710 form ion pairs with each other. In the asymmetric dimer (Figure 8c) the hydrophobic residues in both surface patches that are conserved within the EGFR family (Figure 3), with the exception of I914, are all in contact with hydrophobic residues of the complementary kinase domain. Tyrosine 740 in helix C of the active kinase is part of the hydrophobic interface but can also form hydrogen bonds through its hydroxyl group with the backbone oxygen atoms of residues 909 and 910 of the second kinase. Aspartate 737 in helix C forms an ion pair with lysine 925 of helix H of the second kinase, and at least two other ion pairs have formed, namely, between K733 of helix C and E939 of helix I and between K836 of the regulatory loop and D918 of helix H. The quality of the side-chain packing of residues around helix C in the active structure considerably improves upon formation of this dimer, but this is less obvious for helix H in the inactive structure (Figure 2). In both symmetric dimers the packing of the surface residues around helix C seems to improve, especially for the β4,β5 turn, but only small changes are observed for the regulatory loop (data not shown).

DISCUSSION

Three other models of the EGFR kinase domain in the active form, based solely on cAPK, have been published earlier (Knighton et al., 1993; Singh, 1994; Timms et al., 1995). These models were used mainly to discuss the active site and ATP and substrate binding, and in that respect they are similar to our model of the active kinase. Outside the active site, the three models deviate considerably more, mainly because of differences in the sequence alignment between EGFR and cAPK. With the IRK structure available, this alignment can be improved both because of structural similarities between IRK and cAPK and because of the higher sequence identity of the EGFR kinase with IRK than with cAPK (34% and 21%, respectively). Models based on IRK will also be more accurate than models based on cAPK alone, because a sequence identity of 34%, in contrast to 21%, is high enough to expect the assignment of secondary structure elements and their packing to be correct (Sander & Schneider, 1991; Sali, 1995), although some errors might be present in the longer loops (Sali, 1995). This is borne out by the results of the analyses of the two models.

The published X-ray structures of kinases have revealed that there exists a wide variety of activation mechanisms for kinases [reviewed in Johnson et al. (1996)]. The structures of three fully active kinases, cAPK (Zheng et al., 1993a), phosphorylase kinase (Owen et al., 1995), and casein kinase 1 (Xu et al., 1995), are very similar due to the requirements to bind ATP and substrate and to correctly position the γ-phosphate, catalytic residues, and substrate. The inactive structures have in common that the substrate binding site and/or the ATP binding site are (partially or fully) blocked or occupied either by the regulatory loop, as in IRK (Hubbard et al., 1994), cyclin-dependent kinase 2 (De Bondt et al., 1993), and the MAP kinase ERK2 (Zhang et al., 1994), or by an autoinhibitory loop from beyond the core kinase domain, as observed for twitchin kinase (Hu et al., 1994) and Ca^{2+}/calmodulin-dependent kinase 1 (Goldberg et al., 1996). It has recently become clear that the binding of cyclin A to the PSTAIRE helix (helix C) and surrounding surface of cyclin-dependent kinase 2 (CDK2) stabilizes a conformation of the regulatory loop in CDK2 that allows access to the substrate binding site (Jeffrey et al., 1995). The interaction is partly hydrophobic in nature, as in the EGFR kinase dimers, but the total buried surface area in this heterodimer is approximately 3550 $Å^2$, considerably more than in the EGFR kinase dimers. Binding of cyclin A is

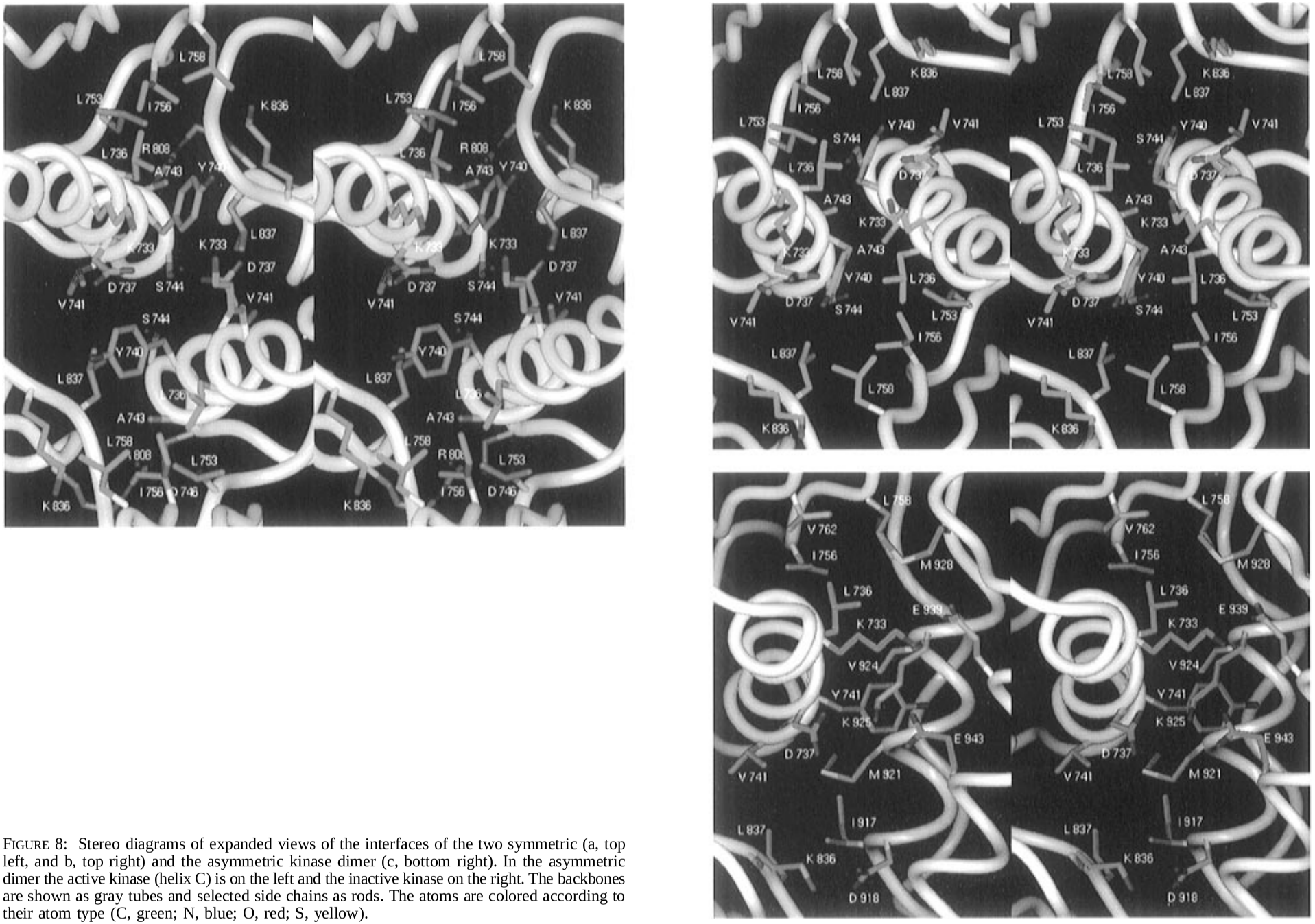

FIGURE 8: Stereo diagrams of expanded views of the interfaces of the two symmetric (a, top left, and b, top right) and the asymmetric kinase dimer (c, bottom right). In the asymmetric dimer the active kinase (helix C) is on the left and the inactive kinase on the right. The backbones are shown as gray tubes and selected side chains as rods. The atoms are colored according to their atom type (C, green; N, blue; O, red; S, yellow).

further accompanied by a rotation of 14° of the N-terminal lobe relative to the C-terminal lobe and the formation of a small β-sheet (Jeffrey et al., 1995), equivalent to β-strands 6 and 9 in EGFR. Full activation of CDK2 requires phosphorylation of T160 in the regulatory loop (Connell-Crowley et al., 1993; Russo et al., 1996), similar to the requirement for phosphorylation of T197 in cAPK (Steinberg et al., 1993), or the phosphorylation of tyrosines 1159, 1162, and 1163 in IRK (Ellis et al., 1986; Wei et al., 1995). The EGF receptor also contains a phosphorylatable residue in the regulatory loop (Y845), but phosphorylation of this residue does not appear to be required for activation of the EGF receptor by EGF. Mutation of Y845 to phenylalanine has no appreciable effect on EGF-dependent kinase activation (in vivo autophosphorylation) and [^{3}H]thymidine intake (Gotoh et al., 1992). Activation of the EGF receptor does not depend on removal of a possible autoinhibitory loop in the C-terminal tail, as deletion of this tail down from residue 973 does not lead to a constitutively active kinase (Walton et al., 1990).

We postulate that activation of the EGFR kinase involves rearrangement of the regulatory loop, characterized by association of β-strands 6 and 9, and surfacing of part of the regulatory loop (most notably L837), leading to extension of the hydrophobic surface around helix C. This extended hydrophobic surface in the active kinase can interact more favorably with a hydrophobic surface of a second kinase. In the case of the asymmetric dimer for instance, if the interaction would take place between helix C of an inactive kinase and helix H of a second kinase, the total buried surface area would only be 1475 Å^2, as opposed to 2087 Å^2 in the activated kinase dimer, and hydrophobic interactions involving L837 and the ion pair formed by K836 would be lost (data not shown). Thus, the active conformation of the regulatory loop is expected to be significantly stabilized relative to the inactive conformation by interaction with a second kinase. This would also hold true if the exact details of the conformation of the inactive loop were incorrect. The X-ray structure of the kinase domain of FGF receptor 1 in the inactive form, which appeared after completion of this work, suggests another possible conformation for the inactive regulatory loop (Mohammadi et al., 1996). However, also in this case the regulatory loop blocks the substrate binding site, and the short β6,β9 sheet, which would cause residue L837 to come to the surface, is not present, so that the proposed activation mechanism would still be perfectly feasible. It is further worthwhile to note that the regulatory loops of the insulin receptor and FGF-receptor kinases are of the same length and that their amino acid sequences can be aligned without any gaps and yet have very different three-dimensional structures. The regulatory loop of the EGF-receptor kinase, although of the same length as well, cannot be aligned without an insertion before and a deletion after the tyrosine residue (Figure 1) and might as likely as not have a three-dimensional structure that is very different from the other two. The crystal structure of the FGF-receptor kinase also contains a crystallographically related dimer that is very similar to the symmetric dimer of Figure 7a, albeit much less closely packed.

Another source of uncertainty in modeling kinases is the conformation of the residues outside the core kinase domain. In the X-ray structures of both cAPK and IRK, and of most of the other kinases for that matter, the sequences extending from either end of the conserved core kinase domain, residues 43−297 of cAPK and 996−1263 of IRK (Hanks & Quinn, 1991), are closely packed onto the core and form an integral part of the kinase domain. In the EGF receptor the juxtamembrane region (residues 645−687) and the C-terminal residues 945−957 after the core are strongly conserved among all known family members, but not with other tyrosine kinases, and could be part of the kinase domain. Residues 951−957 after helix I, which were not modeled due to the lack of suitable templates, might be involved in the asymmetric dimer interface. The residues in the juxtamembrane region are important for determining the position of the kinase domain relative to the membrane. The formation of a disulfide bond in a mutant of EGFR with an extra cysteine residue just outside the transmembrane region (Sorokin et al., 1994) suggests that in the dimer the transmembrane helices are less than 10 Å apart. In the asymmetric dimer the N-terminal ends of the kinase domains are 63 Å apart, and in the symmetric dimers these distances are 53 and 43 Å, respectively. Thus, in the proposed symmetric dimers the 43-residue-long juxtamembrane region of each receptor would be required to span 22 or 17 Å, respectively, to allow correct orientation of the two kinase domains relative to each other. In the case of an asymmetric dimer, the juxtamembrane regions have to span at least 27 Å and are further required to be able to exist in at least two conformations in order to allow for the different orientations of the two kinase domains relative to the membrane.

The mutation of valine 741 to glycine in the EGF receptor reduces the EGF-stimulatable kinase activity of the EGF receptor 300-fold and apparently destabilizes the active form of the kinase, either by destabilizing the active conformation of the regulatory loop or by disrupting the dimerization interface. The V741G mutation is expected to change the interaction of helix C with the regulatory loop in the active state, as V741 makes contact with residues F832 and L837 of this loop. Alternatively, glycine is often found at the end of helices (Aurora et al., 1994), and it is not inconceivable that this mutation prematurely terminates helix C and completely changes the local conformation around helix C. This would affect the stability of the active kinase in both symmetric and asymmetric dimers alike. The Y740F mutation, which has nearly 10-fold reduced receptor autophosphorylation, might reduce the interaction between two kinase domains due to the loss of a hydrogen bond formed by the hydroxyl group of the tyrosine residue, as the hydrophobic contributions of tyrosine or phenylalanine to the interaction energy would be expected to be of similar magnitude. In the first symmetric dimer and in the asymmetric dimer such a hydrogen bond might be present. The predicted effects of the mutations on EGF binding will be discussed more thoroughly in a future paper.

The model also explains the dominant negative effect of EGFR mutants missing the intracellular domain in cells coexpressing normal and mutant receptors (Kashles et al., 1991), as the formation of heterodimers through the extracellular domains cannot lead to the formation of activated kinase dimers. The model predicts that a normal EGF receptor could be activated by forming a heterodimer with a kinase-negative mutant, e.g., K721A (Honegger et al., 1987). But if both wild-type and K721A EGFR are expressed in the same cell, one would still expect only half of the asymmetric or two-thirds of the symmetric dimers

formed to be active, assuming that both receptors are expressed at equal levels and there is no preference for homo- or heterodimerization or for the orientation of the two receptors in the case of an asymmetric dimer.

It is tempting to speculate which tyrosine residues can be phosphorylated in an activated ligand/receptor complex. It has been shown that the autophosphorylation of EGF receptors is predominantly an intermolecular reaction; i.e., one receptor chain is phosphorylated by the kinase of the second receptor chain in the complex (Honegger et al., 1990). Unfortunately, it has not been proven that intramolecular phosphorylation cannot also take place, nor has it been ruled out that intermolecular phosphorylation between complexes can occur. If one assumes that intermolecular phosphorylation is the main mechanism and that phosphorylation takes place predominantly within the dimer, as might be expected under physiological conditions of low ligand and receptor concentrations, in which extensive receptor dimerization and clustering are unlikely to occur, then the proposed model has some interesting implications. In the models the C-terminal tail extends from the end of helix I of one kinase and has to interact with the active site of the other kinase domain (Figure 7). This distance might be too long to allow the autophosphorylation site of the EGF receptor closest in sequence to the kinase domain (Y992) to occupy the active site in the correct orientation and might explain why it is relatively sparsely phosphorylated, as compared with three other phosphorylation sites (Y1068, Y1148, and Y1173) at the C-terminus (Walton et al., 1990).

Furthermore, most residues involved in the kinase dimer interfaces are conserved among the EGF-receptor family members, implying that all might be activated by a similar mechanism, via the formation of either homodimers or heterodimers. Several homologues of EGF, like the heregulins and betacellulin, have been found to stimulate the formation of homo- and/or heterodimers of different family members (Carraway & Cantley, 1994). The possibility of forming heterodimers greatly increases the number of signals that can be initiated by the binding of a ligand, as the tyrosine residues in the C-terminal tail of two different receptors can be phosphorylated and propagate further signals in the cell. The heterodimerization is especially important for erbB3 because this receptor has no (or strongly impaired) intrinsic kinase activity (Guy et al., 1994) and cannot be activated by homodimerization. It is possibly significant that most residues of helix C in erbB3 are mutated relative to EGFR, erbB2, and erbB4, whereas the residues of helix H are conserved (Figure 3). This might imply that erbB3 preferentially (or only) interacts with other kinases via its helix H in an asymmetric dimer, rather than forming a symmetric heterodimer, and in this way acts as the inactive partner that activates the other kinase in the receptor heterodimer. This other family member would then phosphorylate the tyrosine residues in the C-terminal tail of erbB3, and signals specific to erbB3 could be transduced. When symmetric or asymmetric heterodimers of two receptors that both have an intrinsically active kinase are formed, phosphorylation can take place in the C-terminal tails of both receptors, and signals specific for either receptor are activated.

Finally, it should be noted that in erbB4 the residues equivalent to Y740 and V741 are leucine and isoleucine (Figure 3). Little effect of the valine to isoleucine substitution is expected, but from the loss of the hydroxyl group in substituting leucine for tyrosine it would follow in the present model that the kinase activity of erbB4 is reduced [as recently suggested by Tzahar et al. (1996)] similarly to that of theY740F EGF-receptor mutant and/or that erbB4 preferentially interacts via its helix H in an asymmetric dimer, similar to erbB3.

The model presented for the activation of the EGF receptor, formation of a specific dimer of two intracellular kinase domains upon binding of its ligand to the extracellular domain, provides a new framework for discussing the results of experiments on ligand binding and signaling by the EGF receptor and its family members erbB2, erbB3, and erbB4, especially now that the situation has become more complicated with the realization that signaling can take place through both receptor homodimers and heterodimers. At this stage it is not possible to decide which dimer is the preferred model. All three dimers fit the postulated activation model and provide an explanation for the effect of the V741G mutation. Only the asymmetric dimer and possibly the first symmetric dimer show hydrogen bond interactions involving the hydroxyl group of Y740 that might explain the effect of the Y740F mutation. The asymmetric dimer requires the juxtamembrane regions of the receptors to span a large gap of *ca*. 50 Å to the membrane, 10−20 Å more than in the symmetric dimers, and to be able to adopt at least two different conformations. The asymmetric dimer model points to a possible explanation for the existence of kinase-impaired erbB3, with its surface around helix C that is very different from the other family members. To test some of the hypotheses underlying the model and to investigate the intriguing possibility that the dimer could be asymmetric, we have started a number of new mutation experiments. We are currently trying to engineer a constitutively active monomeric EGF receptor and to define the intermolecular interaction sites between two kinase domains.

ACKNOWLEDGMENT

We thank Dr. Michael Veron (Institut Pasteur, Paris, France) for suggesting to us that the point mutation V743G in the mouse EGF receptor might influence the binding to another protein, which led us to the model presented here. We thank Dr. Stevan Hubbard (New York University Medical Center, New York) for making available to us the coordinates of the crystal structure of the insulin receptor kinase domain before they were available through the Brookhaven Protein Data Bank.

REFERENCES

Aurora, R., Srinivasan, R., & Rose, G. D. (1994) *Science 264*, 1126−1130.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. D., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanochi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535−542.

Boni-Schnetzler, M., & Pilch, P. F. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 7832−7836.

Boonstra, J., Rijken, P., Humbel, B., Cremers, F., Verkleij, A., & van Bergen en Henegouwen, P. (1995) *Cell Biol. Int. 19*, 413−430.

Brooks, B. R., Buccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Brünger, A. T. (1992) in *X-PLOR Version 3.1. A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.

Canals, F. (1992) *Biochemistry 31*, 4493−4501.

Carraway, K. L., III, & Cantley, L. C. (1994) *Cell 78*, 5−8.
Carter, T. H., & Kung, H. J. (1994) *Crit. Rev. Oncogen. 5*, 389−428.
Chantry, A. (1995) *J. Biol. Chem. 270*, 3068−3073.
Cochet, C., Kashles, O., Chambaz, E. M., Borrello, I., King, C. R., & Schlessinger, J. (1988) *J. Biol. Chem. 263*, 3290−3295.
Connell-Crowley, L., Solomon, M. J., Wei, N., & Harper, J. W. (1993) *Mol. Biol. Cell 4*, 79−92.
De Bondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O., & Kim, S.-H. (1993) *Nature 363*, 595−602.
Ellis, L., Clauser, E., Morgan, D. O., Edery, M., Roth, R. A., & Rutter, W. J. (1986) *Cell 45*, 721−732.
Fowler, K. J., Walker, F., Alexander, W., Hibbs, M. L., Nice, E. C., Bohmer, R. M., Mann, G. B., Thumwood, C., Maglitto, R., Danks, J. A., Chetty, R., Burgess, A. W., & Dunn, A. R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 1465−1469.
Goldberg, J., Nairn, A. C., & Kuriyan, J. (1996) *Cell 84*, 875−887.
Gotoh, N., Tojo, A., Hino, M., Yazaki, Y., & Shibuya, M. (1992) *Biochem. Biophys. Res. Commun. 186*, 768−774.
Guy, P. M., Platko, J. V., Cantley, L. C., Cerione, R. A., & Carraway, K. L., III (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 8132−8136.
Hanks, S. K., & Quinn, A. M. (1991) *Methods Enzymol. 200*, 38−62.
Hobohm, U., Scharf, M., Schneider, R., & Sander, C. (1992) *Protein Sci. 1*, 409−417.
Honegger, A. M., Dull, T. J., Felder, S., Van Obberghen, E., Bellot, F., Szapary, D., Schmidt, A., Ullrich, A., & Schlessinger, J. (1987) *Cell 51*, 199−209.
Honegger, A. M., Schmidt, A., Ullrich, A., & Schlessinger, J. (1990) *Mol. Cell. Biol. 10*, 4035−4044.
Hu, S.-H., Parker, M. W., Lei, J. Y., Wilce, M. C. J., Benian, G. M., & Kemp, B. E. (1994) *Nature 369*, 581−584.
Hubbard, S. R., Wei, L., Ellis, L., & Hendrickson, W. A. (1994) *Nature 372*, 746−754.
Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J., & Pavletich, N. P. (1995) *Nature 376*, 313−320.
Johnson, L. N., Noble, M. E. M., & Owen, D. J. (1996) *Cell 85*, 149−158.
Jones, S., & Thornton, J. M. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 13−20.
Kashles, O., Yarden, Y., Fischer, R., Ullrich, A., & Schlessinger, J. (1991) *Mol. Cell. Biol. 11*, 1454−1463.
Knighton, D. R., Cadena, D. L., Zheng, J., Ten Eyck, L. F., Taylor, S. S., Sowadski, J. M., & Gill, G. N. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5001−5005.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Livneh, E., Prywes, R., Kashles, O., Reiss, N., Sasson, I., Mory, Y., Ullrich, A., & Schlessinger, J. (1986) *J. Biol. Chem. 261*, 12490−12497.
Livneh, E., Reiss, N., Berent, E., Ullrich, A., & Schlessinger, J. (1987) *EMBO J. 6*, 2669−2676.
Luetteke, N. C., Phillips, H. K., Qiu, T. H., Copeland, N. G., Earp, H. S., Jenkins, N. A., & Lee, D. C. (1994) *Genes Dev. 8*, 399−413.
Lüthy, R., Bowie, J. U., & Eisenberg, D. (1992) *Nature 356*, 83−85.
MacKerell, A. D., Jr., Wiorkiewicz-Kuczera, J., & Karplus, M. (1995) *J. Am. Chem. Soc. 117*, 11946−11975.
Miettinen, P. J., Berger, J. E., Meneses, J., Phung, Y., Pedersen, R. A., Werb, Z., & Derynck, R. (1995) *Nature 376*, 337−341.
Mohammadi, M., Honegger, A., Sorokin, A., Ullrich, A., Schlessinger, J., & Hurwitz, D. R. (1993) *Biochemistry 32*, 8742−8748.
Mohammadi, M., Schlessinger, J., & Hubbard, S. R. (1996) *Cell 86*, 577−587.
Munson, P. J., & Rodford, D. (1980) *Anal. Biochem. 107*, 220−223.
Nicholls, A., Sharp, K. A., & Honig, B. (1991) *Proteins 11*, 281−296.
Owen, D. J., Noble, M. E. M., Garman, E. F., Papageorgiou, A. C., & Johnson, L. N. (1995) *Structure 3*, 467−482.
Russo, A. A., Jeffrey, P. D., Patten, A. K., Massagué, J., & Pavletich, N. P. (1996) *Nature 382*, 325−331.
Sali, A. (1995) *Curr. Opin. Biotechnol. Protein Eng.* 437−451.
Sander, C., & Schneider, R. (1991) *Proteins 9*, 56−68.
Sibilia, M., & Wagner, E. F. (1995) *Science 269*, 234−238.
Singh, J. (1994) *Protein Eng. 7*, 849−858.
Sippl, M. J. (1993) *Proteins 17*, 355−362.
Sorokin, A., Lemmon, M. A., Ullrich, A., & Schlessinger, J. (1994) *J. Biol. Chem. 13*, 9752−9759.
Spaargaren, M., Defize, L. H. K., Boonstra, J., & de Laat, S. W. (1991) *J. Biol. Chem. 266*, 1733−1739.
Steinberg, R. A., Cauthron, R. D., Symcox, M. M., & Shuntoh, H. (1993) *Mol. Cell. Biol. 13*, 2332−2341.
Stewart, J. J. P. (1990) *Quantum Chem. Program Exchange 10*, 86.
Threadgill, D. W., Dlugosz, A. A., Hansen, L. A., Tennenbaum, T., Lichti, U., Yee, D., LaMantia, C., Mourton, T., Herrup, K., Harris, R. C., Barnard, J. A., Yuspa, S. H., Coffey, R. J., & Magnuson, T. (1995) *Science 269*, 230−234.
Timms, J. F., Noble, M. E. M., & Gregoriou, M. (1995) *Biochem. J. 308*, 219−229.
Tzahar, E., Waterman, H., Chen, X., Levkowitz, G., Karunagaran, D., Lavi, S., Ratzkin, B. J., & Yarden, Y. (1996) *Mol. Cell. Biol. 16*, 5276−5287.
Ullrich, A., & Schlessinger, J. (1990) *Cell 61*, 203−212.
Walker, F., Nice, E., Fabri, L., Moy, F. J., Liu, J.-F., Wu, R., Scheraga, H. A., & Burgess, A. W. (1990) *Biochemistry 29*, 10635−10640.
Walton, G. M., Chen, W. S., Rosenfeld, M. G., & Gill, G. N. (1990) *J. Biol. Chem. 265*, 1750−1754.
Wei, L., Hubbard, S. R., Hendrickson, W. A., & Ellis, L. (1995) *J. Biol. Chem. 270*, 8122−8130.
Xu, R.-M., Carmel, G., Sweet, R. M., Kuret, J., & Cheng, X. (1995) *EMBO J. 14*, 1015−1023.
Yarden, Y., & Schlessinger, J. (1987) *Biochemistry 26*, 1434−1442.
Zhang, F., Strand, A., Robbins, D., Cobb, M. H., & Goldsmith, E. J. (1994) *Nature 367*, 704−711.
Zheng, J., Knighton, D. R., Ten Eyck, L. F., Karlsson, R., Xuong, N., Taylor, S. S., & Sowadski, J. M. (1993a) *Biochemistry 32*, 2154−2161.
Zheng, J., Trafny, E. A., Knighton, D. R., Xuong, N., Taylor, S. S., Ten Eyck, L. F., & Sowadski, J. M. (1993b) *Acta Crystallogr. D 49*, 362−365.

BI9614141