Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 5001–5005, June 1993
Biochemistry

Structural features that specify tyrosine kinase activity deduced from homology modeling of the epidermal growth factor receptor

(cAMP-dependent protein kinase/enzyme structure–function/phosphotransferase)

DANIEL R. KNIGHTON*†, DEBORAH L. CADENA‡, JIANHUA ZHENG*, LYNN F. TEN EYCK*§, SUSAN S. TAYLOR*§, JANUSZ M. SOWADSKI¶, AND GORDON N. GILL‡||

Departments of *Chemistry, ¶Biology, and ‡Medicine, University of California at San Diego, La Jolla, CA 92093; and §San Diego SuperComputer Center, P.O. Box 85608, San Diego, CA 92186-0784

Communicated by Helen M. Ranney, February 22, 1993 (received for review October 21, 1992)

ABSTRACT **To identify structural features that distinguish protein-tyrosine kinases from protein-serine kinases, a molecular model of the kinase domain of epidermal growth factor receptor was constructed by substituting its amino acid sequence for the amino acid sequence of the catalytic subunit of cAMP-dependent protein kinase in a 2.7-Å refined crystallographic model. General folding was conserved as was the configuration of invariant residues at the active site. Two sequence motifs that distinguish the two families correspond to loops that converge at the active site of the enzyme. A conserved arginine in the catalytic loop is proposed to interact with the γ phosphate of ATP. The second loop provides a binding surface that positions the tyrosine of the substrate. A positively charged surface provides additional sites for substrate recognition.**

A large and diverse family of protein kinases (PKs) catalyze the transfer of the γ phosphate of ATP to protein substrates. Although various regulatory controls, subcellular localizations, and substrate specificities have evolved, all eukaryotic PKs share a homologous catalytic core (1). Two major groups of PKs have been identified, those that catalyze phosphorylation of serine and threonine [protein-serine(threonine) kinase (PSK); ATP:protein phosphotransferase EC 2.7.1.37] and those that catalyze phosphorylation of tyrosine [protein-tyrosine kinase (PTK); ATP:protein-tyrosine O-phosphotransferase EC 2.7.1.112]; a small number of PKs may phosphorylate both serine and threonine as well as tyrosine (2). The identification of PTKs among retroviral oncogenes and as a necessary functional feature of many growth factor receptors clearly established that, like PSKs, PTKs serve an essential role in growth and development. Because no crystallographic structure of a PTK has been determined, the structural basis of the specificity for tyrosine is not well understood.

Recently, the first crystallographic molecular model of a PK was reported—that of the catalytic subunit of cAMP-dependent PK (cAPK, a PSK) complexed with an inhibitor peptide fragment PKI-(5–24) (3, 4). Alignment of the amino acid sequences of the catalytic domains of 65 PKs showed that the minimal PK catalytic domain is ≈260 residues long (5). The typical sequence identity is 20–60% between two PKs within this catalytic domain; nine invariant and five nearly invariant amino acids are interspersed throughout the domain and serve as fixed sequence alignment points. In the crystallographic model of cAPK, these hallmark amino acids are nearly all associated with essential, conserved functions such as ATP-binding and catalysis (3, 4). When these amino acids are used as points of reference to guide local alignments, insertions and deletions in other PKs coincide with surface loops in cAPK between secondary structural elements, where variation would not disrupt the core structure necessary for enzymatic function.

The cAPK structure thus appears to be a suitable basis for constructing models of other PKs by homology (6). To understand the differences between PTK and PSK, including differences in the catalytic site and in the specificity for the amino acid accepting the phosphate, a model of the core tyrosine kinase domain of epidermal growth factor receptor (EGFR) was constructed. Instead of the binary PKI-(5–24)·cAPK structure, a newly refined 2.7-Å crystallographic molecular model of cAPK complexed with PKI-(5–24) and MgATP was used as the template (7). Although this structure is of an inhibited cAPK, the inhibitor peptide contains all but the hydroxyl group of the serine of a cAPK substrate, and the position of the ATP γ phosphate and the configuration of amino acids at the active site reasonably approximate the geometric requirements for cAPK activity (3, 4). Both enzymes require metal ions for catalysis, although PTK functions better *in vitro* with Mn^{2+} in contrast to PSK preference for Mg^{2+} (8, 9). The Mg^{2+} binding sites in the cAPK crystallographic molecular model were used in the present homology model.

MATERIALS AND METHODS

The model of EGFR was prepared by replacing side chains of a cAPK·MgATP·PKI-(5–24) crystal structure model (7) with EGFR side chains according to the alignment of Fig. 1. The relative insertions and deletions were modeled initially by using backbone fragments of correct length extracted from structures in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory (10, 13–17) and given EGFR sequence, with the exception of the first insertion (*ia*), where a fragment-generating algorithm was used (11, 12). Two residues preceding and following these segments were fit to the insertion/deletion border to determine orientation. The atomic positions were unchanged in the replacement procedure for conserved backbone, and EGFR side chains followed a cAPK side-chain position where possible.

A series of energy minimization steps was carried out to eliminate improper atomic overlaps and poor geometry; also simulated annealing was performed to give the regions of insertion and deletion minima of lower energy. The atoms of insertions and borders of deletions were first energy-minimized against the remainder of the model while holding

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: EGFR, epidermal growth factor receptor; P, phosphorylation site; PK, protein kinase; cAPK, cAMP-dependent PK; PTK, protein-tyrosine kinase; PSK, protein-serine (threonine) kinase; PKI, PK inhibitor.

†Present address: Agouron Pharmaceuticals, Inc., 3565 General Atomics Court, San Diego, CA 92121.
||To whom reprint requests should be addressed.

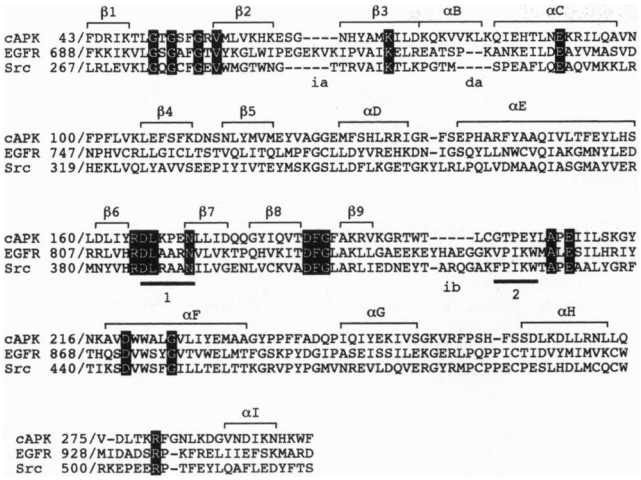

FIG. 1. Sequence alignment of cAPK, EGFR, and Src core PK domain. Primary sequences of cAPK, EGFR, and Src PKs were aligned by using invariant residues and the structural model of cAPK. Brackets above the sequence indicate the β strands and α helices of cAPK that are conserved in EGFR and Src. Highly conserved residues are shaded in black. Regions 1 and 2, which distinguish PTKs from PSKs, are underlined. Insertions *ia* and *ib* and deletion *da* are indicated below the sequence.

that remainder fixed. The side chains of conserved regions and all atoms of insertion and deletion borders were then minimized while holding the conserved α carbon position fixed. An additional subset of atoms from invariant or highly conserved residues was fixed or harmonically restrained because of side-chain or backbone interactions that were expected to be conserved in the PK family. These included ATP; 2 Mg^{2+} ions; residues 698 and 699 on the glycine-rich loop near the ATP phosphates; Lys-721, Glu-738, Arg-812, Asp-813, Asn-818, and Asp-813 to Leu-834, which form interactions with and around the ATP and Mg^{2+} ions; and Lys-851 to Pro-853 near the site of catalysis. The insertion and deletion regions were then subjected to simulated annealing from 2000 to 300 K while holding fixed the remainder of the model, followed by additional energy minimization of the same. This assembly was subjected to 1200 conjugate gradient cycles of energy minimization with α carbon atoms in the conserved regions and a subset near the sites of nucleotide binding and catalysis fixed or restrained. Finally, additional conjugate gradient cycles were run with the conserved α carbon atoms harmonically restrained and the nucleotide/catalytic subset restrained or fixed until the rms gradient was 0.001 kcal/(mol·Å) (1 cal = 4.18 J). The tyrosine-containing substrate pentapeptide (Ala-Glu-Tyr-Leu-Arg) was added, and the phenolic oxygen was positioned manually with respect to the ATP γ phosphorus and Asp-813 side chain, and the side chain of Arg-817 was manually adjusted to place the guanidinium near the position of cAPK Lys-168 to interact with the γ phosphate. The peptide energy, with phenolic oxygen fixed, was then minimized while the EGFR model was held fixed. The program X-PLOR (18) was used for all energy minimization and simulated annealing steps.

The resulting EGFR molecular model has rms deviations from ideal bond lengths of 0.010 Å and from ideal bond angles of 2.9°. The modeling procedure changed the positions of conserved α carbons by an rms distance of 0.19 Å. This difference is smaller than the expected global backbone coordinate differences between homologous proteins of the same 20% sequence identity as exists between cAPK and EGFR (18). However, differences of these magnitudes are more reasonable for the highly conserved constellation of PK active site residues, and energy minimization without restraining the EGFR model to the cAPK coordinates would likely not demonstrably improve the resemblance of the EGFR model to its true structure.

In modeling the EGFR substrate peptide, the tyrosine was placed so as to approximate the requirements of in-line phosphotransfer with inversion of configuration as found for cAPK and expected for EGFR (19, 20). In this location the phenolic oxygen is positioned ≈3 Å from both the γ phosphorus of MgATP and the side chain of invariant Asp-813. This placement, required for catalysis, caused the backbone of a 13-amino acid substrate peptide based on the self-phosphorylation site at Tyr-1173 in EGFR to assume a conformation different from that of PKI-(5–24) found in the crystallographic molecular model of cAPK; therefore only a short PTK peptide substrate consisting of positions P−2 through P+2 from the phosphorylation site P (Ala-Glu-Tyr-Leu-Arg) was included in the model.

RESULTS

Homology Modeling of the EGFR Kinase Domain on cAPK. The amino acid sequence of the EGFR kinase domain was substituted for that of cAPK by using the alignment of the amino acid sequences of the PK family. Amino acids 688–950 from EGFR were aligned with amino acids 43–297 of cAPK. The amino acid identity is 20%, which is similar to the 19–24% range found for other PTKs with cAPK. Six insertions and deletions in the alignment were positioned in the molecular model after consideration both of sequence conservation over the entire family of PKs and of the integrity of the catalytic core of the cAPK structure (Fig. 1). The placement of insertion *ib* in PTK was a critical issue, for although the sequence and length are not conserved in PTKs, insertion *ib* lies between the hallmark sequence motifs Asp-Phe-Gly and Ala-Xaa-Glu (Fig. 1). Within these same boundaries in the cAPK structure are amino acids adjacent to the sites at which the peptide substrate binds and at which catalysis occurs. Thr-197, site of a structurally important phosphorylation in cAPK whose charge may be conserved in many PKs, is linked to the catalytic loop through hydrogen bonding to the highly conserved Arg-165. The least disruptive choice was to make Tyr-845 through Gly-849 of EGFR the insertion and to place it as a short loop after Glu-844, which is equivalent to Thr-197 in the molecular model of cAPK, but before the peptide binding site. This allowed the highly conserved PTK sequence corresponding to region 2 of Fig. 1 to substitute for the highly conserved PSK sequence adjacent to the site of catalysis in the cAPK structure. This choice also allows Glu-842 and Glu-844 to interact with conserved Arg-812 to provide a balance of charge. This region in other tyrosine kinases, such as src and the insulin receptor, contains sites of autophosphorylation that could function similarly to Thr-197 in cAPK.

Because the amino acid sequences of the PK cores of EGFR and cAPK are of similar length, the molecular model of EGFR has the same bilobal architecture as that of cAPK (Fig. 2). The secondary structure of the molecular model necessarily follows closely that of cAPK, beginning with β strand 1 (β1) and ending with α helix I (αI). The insertions and deletions needed to produce the molecular model of EGFR relative to the parent model cAPK were placed between secondary structural elements on surface loops, with the exception of deletion *da*, which combined with a proline substitution disrupts the short α helix B (Fig. 1). The α helix B is also disrupted in the molecular model of myosin light chain kinase (6). The configuration of the essential, conserved residues near the peptide binding site and site of catalysis is thus undisturbed, as is the ATP binding site. The conclusion that the ATP sites in EGFR and cAPK are structurally similar is consistent with their common inacti-

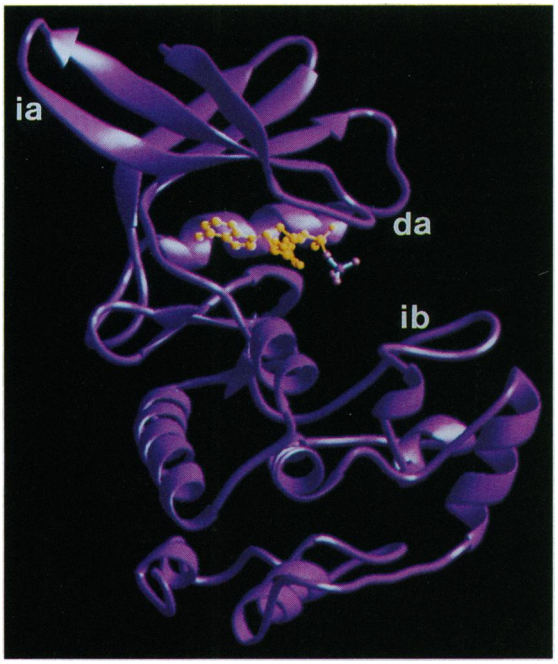

FIG. 2. Ribbon diagram of the model of the EGFR kinase domain based on the crystallographic coordinates of the catalytic subunit of cAPK ternary complex. The bound ATP molecule is shown in yellow with the γ phosphate oxygens in red. Positions of inserts (*ia* and *ib*) and deletion (*da*) are marked with circles and correspond to the following regions: *ia*, 712–715; *ib*, 845–849; and *da*, 729–730.

vation by affinity labeling with fluorosulfonylbenzoyl-5′-adenosine (21, 22). The adenine-ribose binding pocket in the model of EGFR shows homologous hydrophobic amino acids and homologous donors and acceptors of hydrogen bonds as those found in cAPK. In the cAPK structure, ATP is nearly completely buried, in part because of the amino acid sequence beyond the conserved PK core. Because only the core kinase domain of EGFR was modeled, the ATP is more exposed, although in the intact receptor it may be more buried. The phosphates of the ATP bind similarly to those in cAPK in the cleft below the conserved glycine-rich loop and within reach of Lys-721.

Hanks *et al.* (5) identified two regions of amino acid sequence that characteristically differ between PSK and PTK. Region 1 begins with Asp-813 and corresponds to Asp-Leu-Ala-Ala-Arg-Asn in the EGFR, where Asp-813, Leu-814, and Asn-818 are completely conserved in all PKs (Fig. 1). This region also distinguishes the two subfamilies of tyrosine kinases, the src family which always contains Asp-Leu-Arg-Ala-Ala-Asn and the remaining members that contain Asp-Leu-Ala-Ala-Arg-Asn. The analogous position in cAPK begins with Asp-166, has the sequence Asp-Leu-Lys-Pro-Glu-Asn, and corresponds in the cAPK structure to a loop containing amino acids important for catalysis and recognition of the peptide substrate (3, 4). Region 2 begins with Val-852 and corresponds to Val-Pro-Ile-Lys-Trp in EGFR, where Trp-856 is invariant, Pro-853 is nearly so, Lys-855 is always positively charged and Val-852 and Ile-854 are always hydrophobic in the PTK family. The corresponding region in the amino acid sequence of cAPK, beginning at Gly-200 and corresponding to Gly-Thr-Pro-Glu-Tyr, lies adjacent to the active site and forms essential interactions with the substrate peptide. The very strong conservation of sterically similar hydrogen bond donors at the position equivalent to Thr-201 suggests an important role for this amino acid in the PSK family. In the crystallographic molecular model of cAPK, the hydroxyl group of Thr-201 is within hydrogen-bonding distance of invariant Asp-166, the most likely candidate for the required catalytic base in phosphotransfer and thus may contribute to the correct configuration for catalysis.

Active Site of EGFR Tyrosine Kinase. The active site is defined by two β-sheets that form an interface at the cleft between the two lobes. The disposition of the critical residues at the active site in the molecular model of EGFR is shown in Fig. 3. The ATP-binding site is characteristic for the entire PK family (7). The glycine-rich loop anchoring the phosphate and conserved amino acids Lys-721 and Glu-738, located in the small lobe, participate in binding the nucleotide, while the catalytic loop between β strand 6 and β strand 7 in the large lobe functions in both peptide binding and catalysis. In the catalytic loop for PSK, Lys-168 is almost completely conserved, the exception being *Drosophila* ninaC, where arginine is substituted (1, 5). The analogous sequence in PTK, corresponding to region 1, is represented by Asp-Leu-Ala-Ala-Arg-Asn or, in the src family, Asp-Leu-Arg-Ala-Ala-Asn. Although the interesting interchange of two alanines and an arginine distinguishes the two subfamilies of PTKs, the same volume is accessible to the arginine side chains at either position in the molecular model so that the guanidinium groups from arginines at either position can be superposed. Thus, the Ala-Ala-Arg motif of the EGFR and the Arg-Ala-Ala motif of $p60^{src}$ are at either end of a short loop, and the arginine can readily reach the catalytic region from either end of this loop. In the catalytic loop of cAPK, Glu-170 interacts with the substrate. It is possible that the arginine in PTK is involved in substrate recognition, but the complete conser-

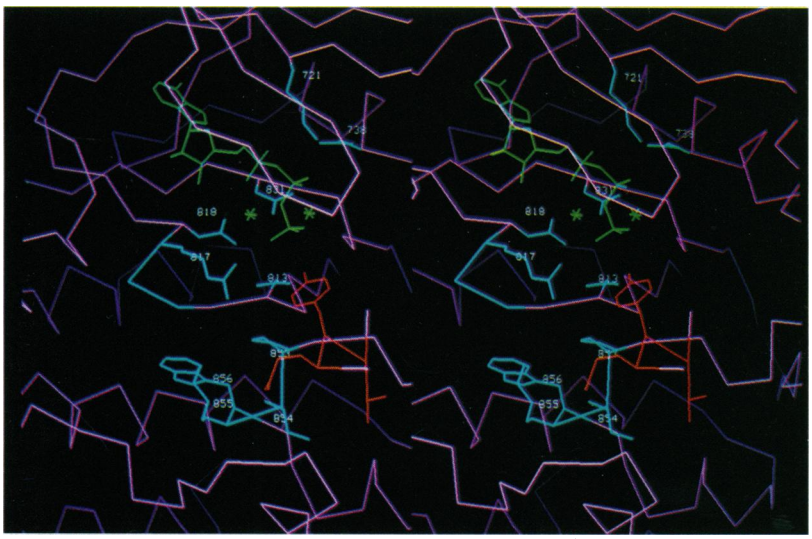

FIG. 3. Stereoview of a model of the active site of EGFR PTK. The ATP molecule and two metal ions (shown as asterisks) interacting with ATP phosphates are green. The peptide substrate is red. The backbone regions unique to tyrosine kinases, Ala-Ala-Arg beginning at position 815 and Pro-Ile-Lys-Trp beginning at position 853, are blue. Substrate tyrosine positioned in part by Pro-853 is in-line to the γ phosphate. Asp-813 is in position to accept the proton from the substrate tyrosyl, and Arg-817 is located on the catalytic loop in close vicinity to the γ phosphate. Glu-738 hydrogen-bonds to the tip of Lys-721, which in turn interacts with the α and β phosphates. Asp-831 and Asn-818 coordinate the two metal ions. Trp-856, conserved in the PTK family, contributes to the overall hydrophobicity of the active site.

vation and similar reach of arginine residues in PTK in either of two positions on this loop, one of which is homologous to the highly conserved Lys-168 of cAPK, implies that the PSK function of Lys-168 in cAPK has been assumed by arginine in PTK. Therefore, Arg-817 is proposed to interact with ATP and likely forms an ion pair with the γ phosphate analogous to cAPK Lys-168.

The tyrosine in the substrate polypeptide must be aligned in the complex between enzyme and substrate so that the phenolic oxygen is positioned for an in-line attack on the γ phosphate of ATP for phosphotransfer to occur (Fig. 3). Asp-813 in EGFR is analogous to Asp-166 in cAPK and is proposed to be the required catalytic base that removes the proton from the phenol. The region Val-Pro-Ile-Lys-Trp (region 2 in Fig. 1) specific to tyrosine kinases provides some of the determinants for proper orientation of the tyrosine. The aromatic ring lies above the conserved Pro-853, and this arrangement provides a hydrophobic interaction. In cAPK, a hydrogen bond likely exists between the acyl oxygen of Gly-200 and the amide proton of the isoleucine at P+1 of PKI-(5–24) (3, 4). The substrate peptide in the complex with EGFR may be positioned by an analogous interaction between the amide proton of the P+1 amino acid and the acyl oxygen of the homologous Val-852. The amide proton of Ile-854 is also available to the acyl oxygen of position P−1. However, such interactions could not be incorporated into the present model while maintaining the proper position of the tyrosine phenolic oxygen and the position of the homologous EGFR region 2, which is superposed on the crystallographic coordinates of cAPK. This suggests that a real difference in location relative to the γ phosphate probably exists between PTK and PSK for these residues. The required small movement for this region of no more than 2 or 3 Å outward from the γ phosphate might be accommodated by small changes in the conserved core. The local sequence Thr-Pro-Glu-Tyr of region 2 in the PSK family (Fig. 1) is conformationally restricted by forming a β-turn in the cAPK structure, and the analogous PTK residues Pro-Ile-Lys-Trp can form a similar turn that restricts the conformation of nearby residues. Conservation of the hydrophobic character and conformation of Pro-853 and the conformation of the adjacent peptide backbone appear to be needed for the proper orientation of the substrate peptide and the tyrosine side chain on it. The bulky side chain of PTK-invariant Trp-856 occupies a nearby hydrophobic cavity, and its interactions must further immobilize the region 2 locale.

Substrate Recognition. PTKs show a preference for acidic residues amino-terminal to the phosphorylated tyrosine (23, 24). However, the low positional specificity has precluded the determination of a consensus sequence for PTK phosphorylation analogous to the Arg-Arg-Xaa-Ser specificity for cAPK (25). Instead, it has been suggested that secondary and tertiary structure are important factors in substrate recognition (23, 24). A segment of EGFR, containing its major autophosphorylation site at Tyr-1173 (26) is an efficient substrate peptide and contains glutamates at the P−1 and P−4 positions. A 13-amino acid peptide corresponding to amino acids 1164–1176 and containing the Tyr-1173 self-phosphorylation site of EGFR is phosphorylated by the purified catalytic domain of EGFR with a K_m of 50 μM (data not shown). Substituting this peptide for the PKI-(5–24) in the molecular model allows an examination of some potential interactions with the substrate. PKI-(5–24) binds to several acidic residues in cAPK that complement the basic determinants proximal to the P site. The analogous surface in the molecular model of EGFR has basic residues that are available for interaction with acidic residues in the amino acid sequence containing Tyr-1173 (Fig. 4). Specifically, Lys-855, part of region 2 and completely conserved as a basic residue in all PTKs, is positioned for possible interaction with glutamic acid at P−1. Glu-203, the analogous residue in cAPK, establishes a precedent for involvement of this position in peptide substrate binding. Although it forms a hydrogen bond with the arginine at P−6 of PKI-(5–24), Glu-203 can readily reach the arginine at P−2 in the absence of an arginine at P−6. Arg-779, conserved as a basic residue in 90% of PTKs, could form a hydrogen bond with the glutamic acid at P−4. Basic residues with the length and flexibility to interact with varying sites in substrates may account for the absence of positional specificity of acidic determinants. It is likely that regions of EGFR outside the core kinase domain are also involved in substrate recognition and provide interactions that contribute to correct placement of substrate tyrosine.

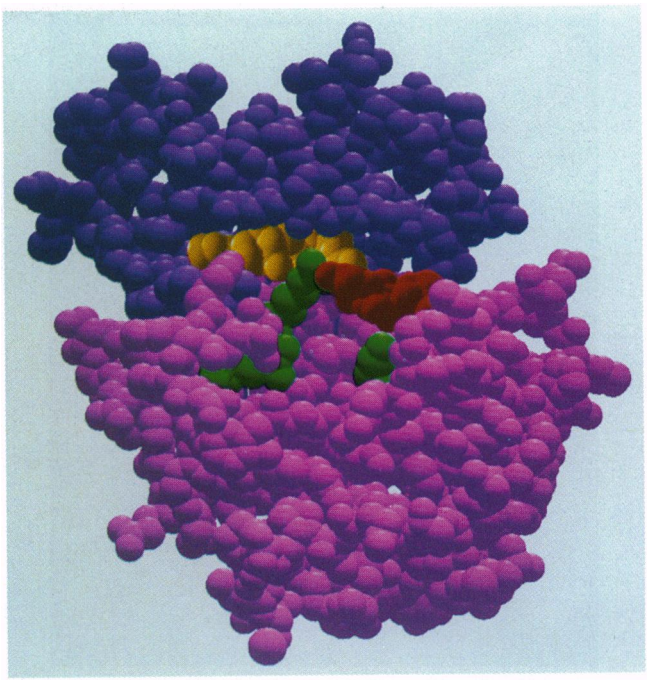

FIG. 4. Space-filling model of the EGFR kinase domain. The smaller upper lobe is dark purple, and the larger lower lobe is light purple. The ATP molecule is yellow, and bound substrate tyrosine and adjacent P−1 and P+1 sites are red. Positive surface charges (green) correspond to the following residues: Lys-855 (right), Arg-779 (lower left), and Arg-817 (upper left).

DISCUSSION

A credible model of a PTK must be able to accommodate tyrosine and correctly position it for phosphotransfer. In the crystallographic molecular model of cAPK, the phosphorylation site is sterically constrained, and the hydroxyl group of serine is positioned for proton abstraction by Asp-166 and by hydrogen bonding between region 2 of cAPK and the backbone of the substrate peptide. For efficient phosphotransfer, the phenolic hydroxyl group of the tyrosine of the substrate must be in the same position in EGFR relative to the γ phosphate of ATP and is likely to interact with the equivalent of Asp-166, which is Asp-813 in EGFR. Inspection of the active site reveals that region 2, corresponding to Val-Pro-Ile-Lys-Trp in EGFR, serves as a binding surface that provides the hydrophobic interactions and hydrogen bonds necessary to orient the phenolic hydroxyl for catalysis.

Analogous to the function of Asp-166 in cAPK, Asp-813 in EGFR is proposed to be the general base removing the phenolic proton. Mutation of the homologous aspartate residue to asparagine in the c-kit and fps proteins abolished tyrosine kinase activity (27, 28). We note that ERBB3, identified by cDNA cloning as homologous to EGFR (29, 30),

contains both an analogous asparagine substitution as well as replacement of the conserved glutamate in α helix C by a histidine.

In addition to those that are proposed to interact with the tyrosine of the substrate, the catalytic loop contains amino acids that are positioned to interact with MgATP. In the ternary complex of cAPK, Lys-168 forms a hydrogen bond with the γ phosphate of ATP; this lysine is nearly invariant in the PSK family (1, 5). In the PTK family, arginine is found in the same position (Arg-Ala-Ala) or two residues away (Ala-Ala-Arg), and it is proposed that this arginine forms a similar hydrogen bond with the γ phosphate of ATP. The introduction of a guanidinium group instead of a primary amine provides a more broadly dispersed positive charge, perhaps helping to position the phenolic oxygen. There are specific differences in the chemical properties of the phenolic group as compared with a simple alcohol to which the phosphate will be transferred. In both PSK and PTK, a conserved aspartate is positioned to extract the proton. However, the pKa of serine is $\approx$13, whereas that of tyrosine is $\approx$10, making extraction of the proton much easier for tyrosine. Serine is a much better nucleophile than tyrosine. This may explain why the chemical step in the tyrosine phosphorylation reaction is relatively slow, whereas it is very rapid in the serine phosphorylation reaction (31). The difference in the properties of serine and tyrosine may explain why the two PK families have the lysine-to-arginine switch at the immediate site of phosphotransfer.

The importance of the Ala-Ala-Arg region in PTK activity is supported by mutational analysis. Moran *et al.* (28) converted the Ala-Ala-Arg of v-fps to Lys-Pro-Glu found in cAPK and observed that the mutant v-fps was defective for both PTK and PSK activity. A more conservative change of only Ala-815 to threonine in the insulin receptor results in diminished kinase activity and dominantly inherited insulin resistance (32). The essentiality of arginine in this region is demonstrated by the conversion to glutamic acid in a human mutant atk that results in X-linked agammaglobulinemia (33). Although these mutants emphasize the critical nature of the configuration of amino acids in the vicinity of the Ala-Ala-Arg region, the amino acid sequences of PTK in *Dictyostelium* indicate that conservation of Ala-Ala-Arg may not be absolutely essential for tyrosine kinase activity (34). Interestingly, the kinases that exhibit dual specificity for tyrosine and for serine and threonine residues contain sequences most characteristic of PSK (2, 35).

The conservation of particular amino acids in the active site implies that they play critical roles in kinase function. Asp-184 and Asn-171 of cAPK, which coordinate to Mg^{2+} in the crystallographic molecular model, are conserved in the PTK family. Mutation of Asp-184 in cAPK resulted in a nonviable yeast mutant (36). The chicken *klg* gene product was isolated with a naturally occurring mutation in the Asp-184 equivalent position, and it was found to lack kinase activity (37). Even relatively conservative mutations of Asp-184 equivalent to glutamate or asparagine in v-fps yielded inactive kinase (28), again emphasizing the very specific and essential function of this residue.

The EGFR model allows predictions of some structural determinants that specify PTK function. Crystallographic, chemical, and mutational studies of PTK will provide critical tests of this model and will help in the understanding of substrate recognition that characterizes this essential class of enzymes.

We thank Dr. Jack Kyte for critical discussions and advice. We gratefully acknowledge use of the San Diego SuperComputer Center. These studies were supported by National Institutes of Health Research Grants DK13149 (G.N.G.) and GM37674 (J.M.S.), by American Cancer Society Grant BE-48 (S.S.T.), and by the Markey Charitable Trust.

1. Hanks, S. K. (1991) *Curr. Opin. Struct. Biol.* **1**, 369–383.
2. Lindberg, R. A., Quinn, A. M. & Hunter, T. (1992) *Trends Biochem. Sci.* **17**, 114–119.
3. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Ashford, V. A., Xuong, N.-h., Taylor, S. S. & Sowadski, J. M. (1991) *Science* **253**, 407–414.
4. Knighton, D. R., Zheng, J., Ten Eyck, L. F., Xuong, N.-h., Taylor, S. S. & Sowadski, J. M. (1991) *Science* **253**, 414–420.
5. Hanks, S. K., Quinn, A. M. & Hunter, T. (1988) *Science* **241**, 42–52.
6. Knighton, D. R., Pearson, R. B., Sowadski, J. M., Means, A. R., Ten Eyck, L. F., Taylor, S. S. & Kemp, B. E. (1992) *Science* **258**, 130–135.
7. Zheng, J., Knighton, D. R., Ten Eyck, L. F., Karlsson, R., Xuong, N.-h., Taylor, S. S. & Sowadski, J. M. (1993) *Biochemistry* **32**, 2154–2161.
8. Taylor, S. S. (1989) *J. Biol. Chem.* **264**, 8443–8446.
9. Wedegaertner, P. B. & Gill, G. N. (1989) *J. Biol. Chem.* **264**, 11346–11353.
10. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977) *J. Mol. Biol.* **112**, 535–542.
11. Shenkin, P. S., Yarmush, D. L., Fine, R. M., Wang, H. & Levinthal, C. (1987) *Biopolymers* **26**, 2053–2085.
12. Biosym Technologies (1992) *Homology User Guide, Version 2.0* (Biosym Technologies, San Diego).
13. Skarzynski, T., Moody, P. C. E. & Wonacott, A. J. (1987) *J. Mol. Biol.* **193**, 171–187.
14. Wang, D., Bode, W. & Huber, R. (1985) *J. Mol. Biol.* **185**, 595–624.
15. Dreusicke, D., Karplus, P. A. & Schulz, G. E. (1988) *J. Mol. Biol.* **199**, 359–371.
16. Almassy, R. J., Fontecilla-Camps, J. C., Suddath, F. L. & Bugg, C. E. (1983) *J. Mol. Biol.* **170**, 497–527.
17. Navia, M. A., McKeever, B. M., Springer, J. P., Lin, T.-Y., Williams, H. R., Fluder, E. M., Dorn, C. P. & Hoogsteen, K. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7–11.
18. Chothia, C. & Lesk, A. M. (1986) *EMBO J.* **5**, 823–826.
19. Knowles, J. R. (1980) *Annu. Rev. Biochem.* **49**, 877–919.
20. Ho, M.-F., Bramson, H. N., Hansen, D. E., Knowles, J. R. & Kaiser, E. T. (1988) *J. Am. Chem. Soc.* **110**, 2680–2681.
21. Zoller, M. J., Nelson, N. C. & Taylor, S. S. (1981) *J. Biol. Chem.* **256**, 10837–10842.
22. Buhrow, S. A., Cohen, S. & Staros, J. V. (1982) *J. Biol. Chem.* **257**, 4019–4022.
23. Geahlen, R. L. & Harrison, M. L. (1990) in *Peptides and Protein Phosphorylation*, ed. Kemp, B. E. (CRC, Boca Raton, FL), pp. 239–253.
24. Pearson, R. B. & Kemp, B. E. (1991) *Methods Enzymol.* **200**, 62–81.
25. Kennelly, P. J. & Krebs, E. G. (1991) *J. Biol. Chem.* **266**, 15555–15558.
26. Downward, J., Parker, P. & Waterfield, M. D. (1984) *Nature (London)* **311**, 483–485.
27. Tan, J. C., Nocka, K., Ray, P., Traktman, P. & Besmer, P. (1990) *Science* **247**, 209–212.
28. Moran, M. F., Koch, C. A., Sadowski, I. & Pawson, T. (1988) *Oncogene* **3**, 665–672.
29. Kraus, M. H., Issing, W., Miki, T., Popescu, N. C. & Aaronson, S. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9193–9197.
30. Plowman, G. D., Whitney, G. S., Neubauer, M. G., Green, J. M., McDonald, V. L., Todaro, G. J. & Shoyab, M. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 4905–4909.
31. Yuan, C.-J., Jakes, S., Elliott, S. & Graves, D. J. (1990) *J. Biol. Chem.* **265**, 16205–16209.
32. Moller, D. E., Yokota, A., White, M. F., Pazianos, A. G. & Flier, J. S. (1990) *J. Biol. Chem.* **265**, 14979–14985.
33. Vetrie, D., Vorechovsky, I., Sideras, P., Holland, J., Davies, A., Flinter, F., Hammarstrom, L., Kinnon, C., Levinsky, R., Bobrow, M., Smith, C. I. E. & Bentley, D. R. (1993) *Nature (London)* **361**, 226–233.
34. Tan, J. L. & Spudich, J. A. (1990) *Mol. Cell. Biol.* **10**, 3578–3583.
35. Ben-David, Y., Letwink, V., Tannock, L., Bernstein, A. & Pawson, T. (1991) *EMBO J.* **10**, 317–325.
36. Gibbs, C. S. & Zoller, M. J. (1991) *J. Biol. Chem.* **266**, 8923–8931.
37. Chou, Y.-H. & Hayman, M. J. (1991) *Proc. Natl. Acad. Sci. USA* **88**, 4897–4901.