

Two Novel Protein-Tyrosine Kinases, Each with a Second Phosphotransferase-Related Catalytic Domain, Define a New Class of Protein Kinase

ANDREW F. WILKS,^{1*} AILSA G. HARPUR,¹ R. R. KURBAN,¹ STEPHEN J. RALPH,¹
GISELA ZÜRCHER,² AND ANDREW ZIEMIECKI²

*Ludwig Institute for Cancer Research, Post Office, Royal Melbourne Hospital, Victoria 3050, Australia,¹ and
Institute for Clinical and Experimental Cancer Research, University of Berne, 3004, Berne, Switzerland²*

Received 12 November 1990/Accepted 19 January 1991

The protein-tyrosine kinases (PTKs) are a burgeoning family of proteins, each of which bears a conserved domain of 250 to 300 amino acids capable of phosphorylating substrate proteins on tyrosine residues. We recently exploited the existence of two highly conserved sequence elements within the catalytic domain to generate PTK-specific degenerate oligonucleotide primers (A. F. Wilks, Proc. Natl. Acad. Sci. USA 86:1603-1607, 1989). By application of the polymerase chain reaction, portions of the catalytic domains of several novel PTKs were amplified. We describe here the primary sequence of one of these new PTKs, JAK1 (from Janus kinase), a member of a new class of PTK characterized by the presence of a second phosphotransferase-related domain immediately N terminal to the PTK domain. The second phosphotransferase domain bears all the hallmarks of a protein kinase, although its structure differs significantly from that of the PTK and threonine/serine kinase family members. A second member of this family (JAK2) has been partially characterized and exhibits a similar array of kinase-related domains. JAK1 is a large, widely expressed membrane-associated phosphoprotein of approximately 130,000 Da. The PTK activity of JAK1 has been located in the C-terminal PTK-like domain. The role of the second kinaselike domain is unknown.

Protein-tyrosine kinases (PTKs) are structurally well suited to a role in intracellular signal transduction. Many growth factor receptors, for example, transduce the extracellular stimulus they receive through interaction with their cognate ligand via an intracellular tyrosine kinase domain (5, 33, 52; reviewed in reference 60). Members of the PTK family each bear a highly related "catalytic" domain. The phylogenetic relationships established by an amino acid sequence comparison of the catalytic domains (10) are borne out in the overall structure of the PTKs. For example, families of PTKs, such as those based on the structure of the colony-stimulating factor-1 growth factor receptor (38) (including the two types of the platelet-derived growth factor receptor [4, 58]) and the protooncogene *c-kit* [59]) and those clustered around the cytoplasmic PTKs *c-src* (29) (including HCK/bmk [12], LCK [28], and *c-yes* [42], among others) and *c-fes* (37) (including *c-FER/flk* [11, 25]) each share the highest degree of identity with other members of their cluster and, in respect to their overall topology, are structurally more related to each other than to members of other classes of PTK. Hence, the recombination of the PTK catalytic domain with a wide variety of regulatory and other interactive domains suggests a strong evolutionary drive toward the rapid expansion of the use of its physiologically powerful catalytic activity. This combinatorial approach to the evolution of multidomain proteins such as the PTK family predicts the extensive utilization of the basic tyrosine kinase domain in other metabolic niches.

Application of the polymerase chain reaction (PCR) (32, 40) using degenerate PTK-specific oligonucleotides (56) to the isolation of novel PTK-related sequences has been a particularly successful strategy; to date, 12 novel protein

kinase-related molecules have been isolated (12a, 36). Two novel and highly related PTKs were isolated from the cDNA of a murine growth factor-dependent hemopoietic cell line, FDC-P1 (56). The similarity of these two sequences to each other, coupled with the presence in each of a rare sequence idiosyncrasy in a normally highly conserved motif, led to the speculation that these two molecules, although clearly members of the broader kinase family, were a distinct subfamily of PTKs. We have named these new PTKs JAK1 and JAK2 (from Janus kinase [6]). We report here the complete sequence of one of these PTKs (JAK1), confirm its structural relatedness to JAK2, and describe some of the salient features of the JAK1 protein.

MATERIALS AND METHODS

Screening of cDNA libraries. Several cDNA libraries were screened according to the protocols outlined in Maniatis et al. (27). cDNA libraries from murine NFS TPA-activated spleen (Clontech; catalog no. ML1018), murine Swiss albino 3T3 fibroblasts (Clontech; catalog no. 1023b), murine BALB/c bone marrow (Clontech; catalog no. ML1007), murine Swiss Webster whole brain (Clontech; catalog no. ML1002), murine ICR linoleic acid-activated pleural macrophage (Clontech; catalog no. ML1005b), and human first-trimester fetal liver (Clontech; catalog no. HL1005b) were all generated in λ gt11. cDNA libraries from murine BALB/c testis (Clontech; catalog no. ML1020b), murine day 10 embryonic neuroepithelium (36), and human foreskin fibroblast cell line AG1518 (4) were generated in λ gt10. Around 10^6 recombinants of each of these libraries were screened on each occasion.

Library screening was carried out as follows. The FD22 (JAK1) PCR clone was labeled by nick translation (27) and used to screen the murine libraries. A murine cDNA clone of

* Corresponding author.

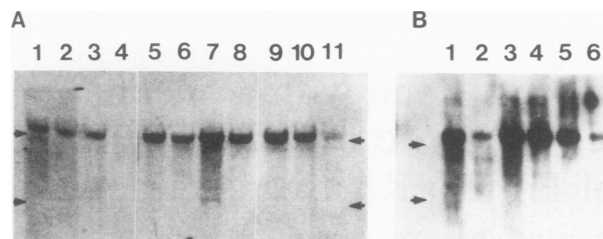


FIG. 1. Northern analysis of murine and human JAK1. (A) Aliquots (5 μ g) of poly(A)⁺ mRNA from murine tissues. Lanes: 1, lung; 2, liver; 3, kidney; 4, intestine; 5, brain; 6, skeletal muscle; 7, spleen; 8, salivary gland; 9, placenta; 10, mammary gland (lactating). Samples were fractionated on a 1.0% agarose-formaldehyde gel, and the RNA was transferred onto a nitrocellulose membrane. The transferred RNA was hybridized with a 1.8-kb ³²P-labeled murine JAK1 probe, and the filter was autoradiographed for 16 h at -70°C with two intensifying screens. The relative mobilities of 28S rRNA (upper arrow) and 18S rRNA (lower arrow) are shown. (B) Aliquots (2 μ g) of poly(A)⁺ mRNA from the human hemopoietic cell lines. Lanes: 1, HL60 (myelomonocytic); 2, U937 (monocytic); 3, LK63 (pre-B); 4, Raji (B cell); 5, CEM (T cell); 6, K562 (erythro-leukemia). Samples were fractionated on a 1.0% agarose-formaldehyde (27) gel, and the RNA was transferred onto a GeneScreen Plus (Dupont) membrane. The transferred RNA was hybridized with a full-length ³²P-labeled human JAK1 probe, and the filter was autoradiographed for 16 h at -70°C with two intensifying screens. The relative mobilities of 28S rRNA (upper arrow) and 18S rRNA (lower arrow) are shown.

1.8 kb was isolated among three other positives from the neuroepithelial and bone marrow cDNA libraries. Two full-length human JAK1 cDNA clones were isolated from the unamplified human foreskin fibroblast cell-line AG1518 by using the murine cDNA as a probe. Hybridization was at 65°C in $6\times$ SSC ($1\times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–1% sodium dodecyl sulfate (SDS)–0.5% BLOTTO–200 μ g of sonicated and denatured herring sperm DNA per ml. After hybridization, the stringency of the final wash was $0.2\times$ SSC–0.1% SDS at 65°C . Filters were autoradiographed overnight with Kodak XAR-5 X-ray film.

For JAK2, the murine macrophage was screened first with the FD17 (JAK2) PCR clone, yielding five positives, and a portion of the longest cDNA clone was isolated and used to screen the remaining cDNA libraries.

DNA sequencing. Two strategies were employed for the sequencing of JAK1 and JAK2 cDNA clones. In the case of the human JAK1 sequence, the Erase-a-Base kit (Promega) was employed to generate nested deletions of the largest *Eco*RI fragment. All of the murine JAK2 sequence data and the remainder of the human JAK1 sequence were determined by using oligonucleotide primers based on previously determined DNA sequences. In each case, the sequence information was generated by using the dideoxynucleotide chain termination method (41). All sequence information was determined on both strands.

Northern (RNA) analysis. Poly(A)⁺ mRNA samples were prepared as described elsewhere (3). Aliquots (1 μ g [Fig. 1B] or 5 μ g [Fig. 1A]) were analyzed by electrophoresis on a 1% agarose gel containing 2.2 M formaldehyde, 20 mM MOPS (morpholinepropanesulfonic acid; pH 6.8), 1 mM EDTA, and 5 mM sodium acetate and transferred to Hybond (Amersham; catalog no. RPN303N) or nitrocellulose (Schleicher & Schuell; BA85; catalog no. 401196) membranes. Filters were prehybridized for 4 h in 50% formamide containing $3\times$ SSC, $5\times$ Denhardt's solution, 10 mM HEPES (*N*-2-hydroxy-

ethylpiperazine-*N'*-2-ethanesulfonic acid; pH 7.0), 100 μ g of poly(C) per ml, 100 μ g of denatured herring sperm DNA per ml, 10 μ g of *Escherichia coli* DNA per ml, and 0.1% SDS (Fig. 1B) or in 50% formamide containing $2\times$ Denhardt's solution, 100 μ g of denatured salmon sperm DNA per ml, and 0.1% sodium pyrophosphate (Fig. 1A) and hybridized in the same solution with nick-translated or randomly primed ³²P-labeled murine or human JAK1 insert for 18 h at 42°C . Filters were washed at a final stringency of $0.1\times$ SSC–0.1% SDS at 65°C before exposure to Kodak XAR-5 X-ray film with two intensifying screens.

Antibody reagents and protein analysis. Polyclonal rabbit antisera M7 and M8 were raised against affinity-purified pGEX/JAK1/1 bacterial fusion protein (see "Protein kinase assays" below). Polyclonal antibodies M3 and M4 against the C-terminal peptide (-TSFQNLIECFEALLKC-) of JAK1 were raised in rabbits. Peptide was coupled to keyhole limpet hemocyanin with 0.05% glutaraldehyde, emulsified in Freund's complete adjuvant, and injected intradermally at several sites. The animals were boosted 4 and 7 weeks later with coupled peptide emulsified in Freund's incomplete adjuvant and bled 10 days after the last injection.

Cells were metabolically labeled with [³⁵S]methionine in methionine-free medium containing 100 μ Ci of isotope per ml. RIPA buffer (20 mM Tris [pH 7.5] containing 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1 mM EDTA, and 1 mM phenylmethylsulfonyl fluoride) extracts were incubated on ice with antiserum, and immune complexes were isolated by using protein A-bearing *Staphylococcus aureus* bacteria. Proteins were resolved by SDS-polyacrylamide gel electrophoresis (23), and radioactively labeled bands were detected by exposure to X-ray film (Kodak XAR-5). Western blot (immunoblot) analysis was performed as described by Towbin et al. (50) as modified by Ziemiecki et al. (61) with either alkaline phosphatase or ¹²⁵I-labeled protein A as a detection system.

Protein kinase assays. A variety of protocols have been tried in order to reveal the PTK activity of the JAK1 protein. First, extraction of murine mammary fibroblasts (35) has been performed in a range of buffers containing Triton X-100 or Nonidet P-40 (1.0%) alone or with added sodium deoxycholate (0.5 or 1.0%) or in RIPA buffer containing 1.0% Triton X-100, 1.0% sodium deoxycholate, and 0.1% SDS. Cells have been extracted in the presence or absence of phosphatase inhibitors such as 20 mM EDTA, 10 mM NaF, and 100 μ M Na₂VO₄.

After immunoprecipitation, kinase assays have been performed with a range of ATP concentrations (100 nM to 10 mM) or with carrier-free [γ -³²P]ATP (Amersham; catalog no. 10169) in either 20 mM Tris (pH 7.4) or 50 mM HEPES (pH 7.4), with 10 mM Mn²⁺, Mg²⁺, or Zn²⁺ as divalent cation. Incubations have been performed on ice (15 min), at 25°C (15 min), at 30°C (15 min), or at 37°C (2 min) in the presence or absence of the phosphatase inhibitor Na₂VO₄. Finally, we have employed [γ -³²P]GTP as phosphate donor in lieu of [γ -³²P]ATP, with no success.

In order to generate the JAK1-glutathione transferase fusion proteins shown in Fig. 4, domain 1 (from nucleotides 1770 to 2672 in Fig. 2) and the PTK domain (from nucleotides 2672 to the end in Fig. 2, thus including five extra amino acids beyond the ATP-binding glycine motif) were each fused into the *Bam*HI site of pGEX2 (44). The fusion protein was induced by the addition of 1 mM IPTG (isopropyl- β -D-thiogalactopyranoside) as described elsewhere (44), and Western blot analysis was performed on an induction time course with the M3 anti-JAK1 serum and the antiphospho-

tyrosine antiserum (15). Several sources of antiphosphotyrosine antisera were tried. The data in Fig. 4B were obtained by using commercially available monoclonal antibody preparation PY-20 (ICN). In control experiments, induction of the insertless pGEX or pGEX/JAK1/1 fusion protein produced no detectable tyrosine phosphorylation of bacterial substrates, and the reactivity of the antiphosphotyrosine antiserum could be completely abolished by the addition of phenyl phosphate.

Computer-aided sequence analysis. Amino acid sequence comparisons were performed by using an alignment program from the Staden-based suite of programs on a VAX VMS 5.2. Phylogenetic analysis of the two kinaselike domains of JAK1 was performed by using the tree-building concept of Fitch and Margoliash (8) as implemented by Feng and Doolittle (7). The database bearing most of the PTK catalytic domain sequences was kindly sent to us by S. Hanks of the Salk Institute and was supplemented by the addition of sequences drawn from our own local database. The SCORE program used to construct the difference matrices from which the trees were derived by the BORD and BLEN programs was the gift of R. Doolittle of the University of California-San Diego.

RESULTS

Isolation and DNA sequence of cDNA clones encoding human FD22 (JAK1). We chose to focus on PCR clone FD22 (56) for our initial studies. Northern analysis (56; Fig. 1) demonstrated that in both mouse and human tissues and cell lines, FD22 (JAK1) was encoded by a single widely expressed 5.4-kb mRNA. Human cDNA clones of FD22 (JAK1) were isolated from a human foreskin fibroblast cell line (AG1518) cDNA library (4). Two of the eight primary isolates cloned contained inserts which were candidates for being full-length cDNAs (~5.3 kb).

The nucleotide sequence of human JAK1 is shown in Fig. 2. The 5' end of the clone has stop codons in all three reading frames prior to the putative initiation ATG. Two ATG start codons in frame with the longest open reading frame were found at positions 40 and 76 in the nucleotide sequence shown in Fig. 2. The first of these is embedded in a particularly poor Kozak consensus sequence (18) (-TAA ATGCAG-), and the second matches strongly with the optimal consensus sequence defined by Kozak, namely, -GCCATGGCT-. We have chosen to consider the second ATG as the initiation codon for this protein, since the first one transgresses one of the strongest correlations found in the sequences which precede initiation codons, namely, the presence of a T residue (in lieu of an A residue) three nucleotides before the ATG sequence. At the 3' end, an in-frame stop codon at position 3502 defines the C terminus of the protein. A large (1.405 kb) 3' untranslated region containing a polyadenylation signal (data not shown) completes the mRNA sequence.

The JAK1 coding region of 3,426 bp encodes a protein of 1,142 amino acids with a calculated molecular mass of 132,000 Da. The PTK catalytic domain is located towards the C terminus of the JAK1 protein (Fig. 2). In describing the structural features of this domain, we have chosen to adopt the nomenclature of Hanks et al. (10). All of the highly conserved motifs typical of PTK domains are present within the putative PTK domain of JAK1. The presence of a tyrosine residue at position 1022 in the JAK1 protein, 11 residues C terminal to subdomain VII (a similarly placed tyrosine is a site of tyrosine autophosphorylation in *v-fps*

[54]), is a consistent feature of members of the PTK family and is considered diagnostic of membership of this class of kinases. The entire catalytic domain of 255 amino acids is approximately 28% (with *c-fes* [57]) to 37% (with TRK [19]) identical to that of other functionally defined PTKs. Finally, there is a rare variant of the highly conserved subdomain VIII motif (residues 1032 to 1039), which is believed to lie close to the active site (10). The phenylalanine and the tyrosine flanking the conserved tryptophan in this motif have been found only in the two members of this subfamily of PTKs, namely, FD22 (JAK1) and FD17 (JAK2).

A second protein kinase-related domain (here designated domain 1) is located between amino acids 578 and 824, 47 amino acids N terminal to the putative PTK domain. All of the conserved elements of protein kinases are preserved spatially in this domain. In Fig. 2 these elements are numbered with respect to their similarity to the subdomains of protein kinases described by Hanks et al. (10) (with the subscript a, e.g., III_a), and the amino acid sequences of the two kinase-related domains of JAK1 are compared with each other and with a known PTK (TRK [19]) and human CDC2 (24) in Fig. 3A. The overall structural similarity of this domain to the kinase domains of both the PTK and threonine/serine kinase families strongly suggests that this region of the protein also functions as a protein kinase. There are, however, significant differences in the sequences of key motifs within this domain which suggest that domain 1 may confer a catalytic activity other than serine/threonine or tyrosine phosphorylation. For example, subdomain VI_a is poorly conserved with respect to the equivalent motifs in the other kinase families, and the normally invariant -Asp-Phe-Gly- sequence of the PTK and threonine/serine kinase families (subdomain VII_a) is replaced by the motif -Asp-Pro-Gly- in domain 1 of JAK1. As has been noted elsewhere, the conservation of the precise sequence of subdomain VI in the PTK and threonine/serine kinase families appears to correlate with the substrate specificity of the kinase (10). Thus, it is possible that domain 1 of the JAK1 kinase has a substrate specificity other than that exhibited by the PTK and threonine/serine kinases. In support of this notion there are subtle differences in the normally consistent spacing between certain key motifs in domain 1 of JAK1. The components of the ATP-binding site (subdomains I_a and II_a) are some seven amino acids further apart in this domain than they are in both the PTK family and the threonine/serine kinase family. Moreover, the spacing between subdomains VI_a and VII_a in this region is also longer by nine amino acids. Conversely, the distance between subdomains VII_a and IX_a is seven amino acids shorter than the corresponding region in the PTK catalytic domain. The overall structure of this domain can be expected to be somewhat different from the catalytic domains of the members of the PTK and threonine/serine kinase families.

The sequences N terminal to domain 1 bear no homology to any other portion of a previously described protein kinase. Specifically, we detected no homology to the SH2 domain (34) described for the cytoplasmic PTKs such as *c-fes/fps* (39), GTPase-activating protein (51), and the phospholipase C family of proteins (46). This is a particularly interesting observation since no other nonreceptor PTK lacking this feature has been described. A hydrophilicity plot failed to demonstrate the presence of a hydrophobic domain characteristic of the growth factor receptor type of PTK (Fig. 3B), suggesting that this protein is wholly intracellular, like other members of the nonreceptor class of PTKs.

JAK1 protein is a large, widely expressed protein with a

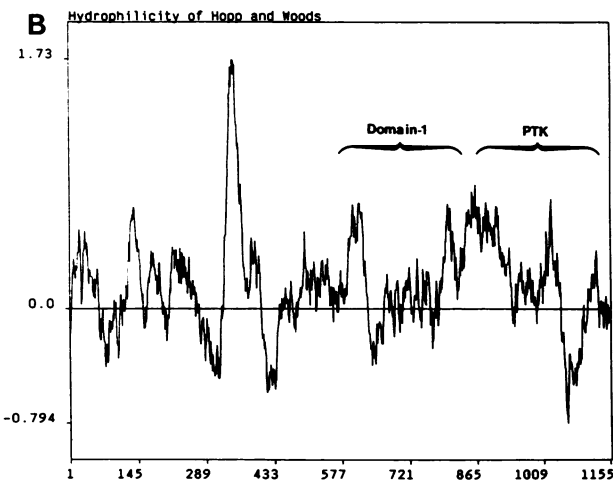
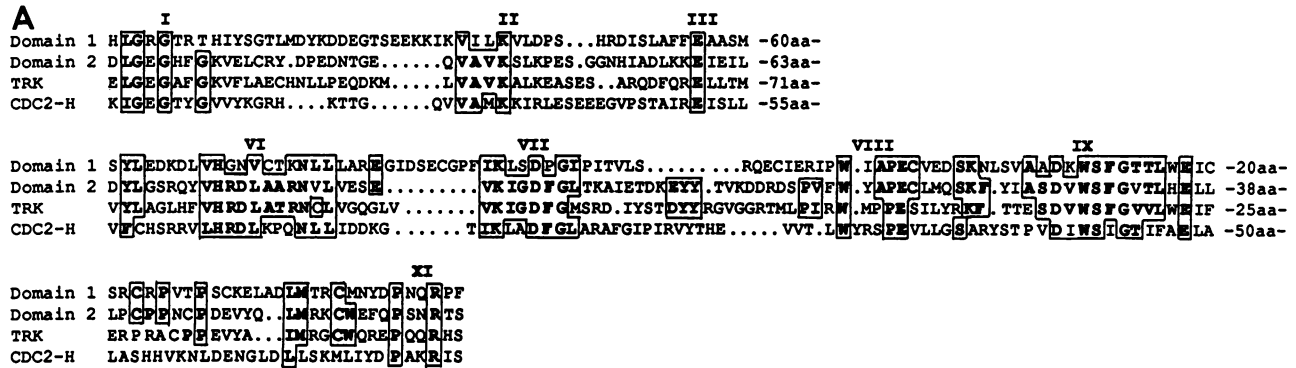


FIG. 3. (A) Amino acid sequence comparison of the two kinase-related domains of JAK1. The amino acid sequences (expressed in one-letter amino acid code) of the two kinase-related domains (domain 1, amino-acids 576 to 825; domain 2 [PTK domain], amino acids 868 to 1130) of JAK1, the PTK domain of TRK (19) (amino acids 158 to 416), and the human threonine/serine-specific kinase CDC2 (24) (amino acids 9 to 272) are aligned in order to maximize identity. The kinase-related domains have been divided into three segments, and the number of amino acid residues separating each segment appears at the end of each line. Motifs held in common between at least two of these domains are in boldface type and are boxed. Roman numerals above the alignment correspond to the conserved-domain nomenclature devised by Hanks et al. (10). (B) Hydrophathy plot of the human JAK1 protein. The protein sequence of human JAK1 (including the 10 extra amino acids which precede the most likely initiation codon) were analyzed by the hydrophilicity algorithm of Kyte and Doolittle (22), using a span length of 25 amino acids. The relative locations of the two kinase-related domains are marked domain 1 and PTK. The absence of a hydrophobic trans-membrane domain and the presence of a highly hydrophilic region between amino acids 323 and 350 can be clearly seen.

PTK domain. We have generated several antisera against the human JAK1 protein. Polyclonal antisera directed against the hexadecamer -TSFQNIECFEALLKC- (the C-terminal 15 amino acids of JAK1) were raised in rabbits and used to investigate the nature of the JAK1 protein. A second rabbit antiserum was generated by using a pGEX (44) bacterial fusion protein containing the entire domain 1 region of the human JAK1 protein (see Materials and Methods). Preliminary sequence analysis of cDNA clones of murine JAK1 demonstrated that the C termini of the human and murine versions of this protein were identical (data not shown), and the murine and human domain 1 regions exhibited a very high degree of identity. The two systems have thus been used interchangeably in the investigation of the properties of the JAK1 protein.

Both antisera have been used for Western blot analyses and immunoprecipitation studies, and the data confirm the mRNA expression studies shown in Fig. 1. For example, antisera M3 and M8 both immunoprecipitate a protein of the same apparent molecular mass (130 kDa) from [³⁵S]methionine-labeled murine breast fibroblasts (Fig. 4A). A characteristic feature of members of the PTK family is that they are able to accomplish an act of self-phosphorylation in vitro. Intriguingly, despite the high degree of sequence similarity between the PTK-related sequence of JAK1 and the PTK family in general, we have been unable to demonstrate tyrosine kinase catalytic activity in immunoprecipitates of this protein from any of the murine or human sources tested. A wide range of possibilities has been tested in search of suitable conditions for the demonstration of this activity. These are listed in the Materials and Methods section. We

are unable to draw firm conclusions from our failure to demonstrate PTK activity in vitro. The reason for the lack of activity may lie in a steric effect of the antibody in the active site of the enzyme. Alternatively, the PTK domain may be cryptic and require activation in *trans* from a cofactor or even in *cis* from the activity of domain 1.

In order to determine whether domain 1 or the PTK domain, in isolation, bore catalytic activity, we generated bacterial fusion proteins of each with the glutathione transferase protein of *Schistosoma japonicum* (44) and attempted to demonstrate with the aid of antiphosphotyrosine antibodies (15) the coordinate induction of the fusion protein and tyrosine-phosphorylated protein. In this system, as in others described in the literature (25, 26), there is no cross-reactive background for the antiphosphotyrosine antiserum, since there are no tyrosine kinases in bacteria (Fig. 4B). The phosphorylation of bacterial proteins on tyrosine is thus easily detectable with such a serum. In this series of experiments, neither pGEX without insert nor pGEX bearing domain 1 (pGEX/JAK/1/1) demonstrated any tyrosine kinase activity. We have further purified the pGEX/JAK1/1 fusion protein by affinity chromatography on a reduced-glutathione column and have failed to detect any kinase activity when histones, casein, or enolase was used as an exogenous substrate. The pattern of inducible tyrosine phosphorylation exhibited by the pGEX PTK fusion protein (pGEX/JAK1/2) (Fig. 4B) is unusually simple for an ectopically expressed PTK fusion protein (cf., for example, data in reference 26). Remarkably, autophosphorylation of the fusion protein itself does not seem to occur, a fact which may go some way

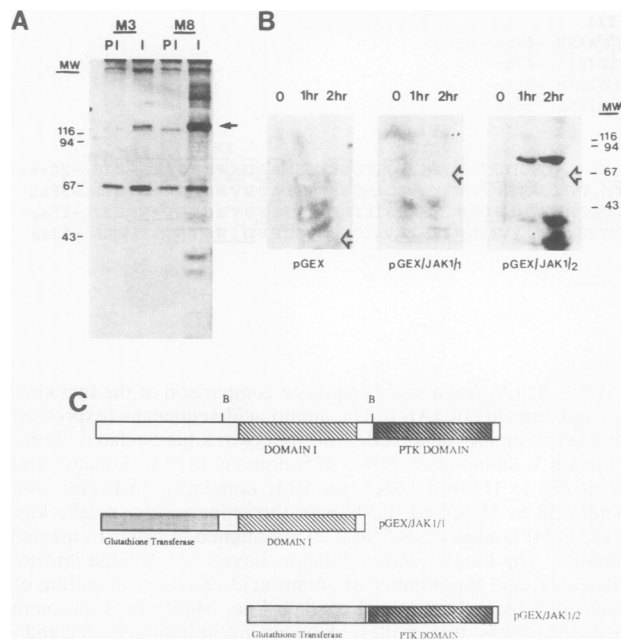


FIG. 4. Analysis of the JAK1 protein. (A) Cellular proteins of the murine mammary fibroblast cell line (35) were labeled with [³⁵S]methionine and immunoprecipitated with either preimmune (PI) or immune (I) anti-JAK rabbit antiserum (raised in rabbit M8 against the pGEX/JAK1/1 fusion protein or the C-terminal peptide [M3]) and fractionated on a 9.5% SDS-polyacrylamide gel (23). Both rabbit antisera specifically immunoprecipitated a ³⁵S-labeled protein with an apparent molecular weight (MW) of 130,000. (B) Demonstration of tyrosine kinase activity in JAK1 bacterial fusion proteins. JAK1 fusion proteins were generated by using pGEX2 (44). The entire domain 1 region was included in construct pGEX/JAK1/1. The PTK domain portion of the fusion protein extended to the *Bam*HI site 15 nucleotides 5' of the first glycine codon of the GYGXXG motif of the ATP-binding site. An empty vector control was also used. The bacteria were induced by the addition of 1 mM IPTG as described by Smith and Johnson (44), and two 1-ml aliquots of the bacteria were removed at 60 and 120 min postinduction and lysed with SDS sample buffer. Western analysis of the samples was performed by using antiphosphotyrosine antisera (PY-20 [ICN]). The arrows mark the positions of the GEX/JAK fusion proteins in each induction. (C) Construction of the pGEX/JAK fusion proteins. The locations of the two kinase-related domains of JAK1 and the structures of the fusion proteins with the glutathione S-transferase gene are shown.

toward explaining why we have had difficulty in demonstrating PTK activity in the intact protein.

A second member of the JAK family of PTKs. cDNA clones covering a significant portion of the coding region of the PCR clone FD17 (JAK2) have been isolated from a range of murine cDNA libraries. The predicted amino acid sequences of JAK2 and JAK1 show several regions of significant similarity to each other (Fig. 5). It is on the basis of this structural relatedness that these proteins can be identified as members of the same subfamily of PTKs. In recognition of this relatedness, particularly in their common possession of a second phosphotransferase-related domain, we have named these PTKs JAK1 and JAK2 (for Janus kinase [6]).

In broad structural terms, the two PTKs are closely related. The JAK2 protein has both the C-terminal PTK domain and the kinase-related domain 1 observed in JAK1. The similarity between these two proteins in these two regions is considerable (53% [64% if conservative substitutions are included] for domain 1 and 51% [63% if conserva-

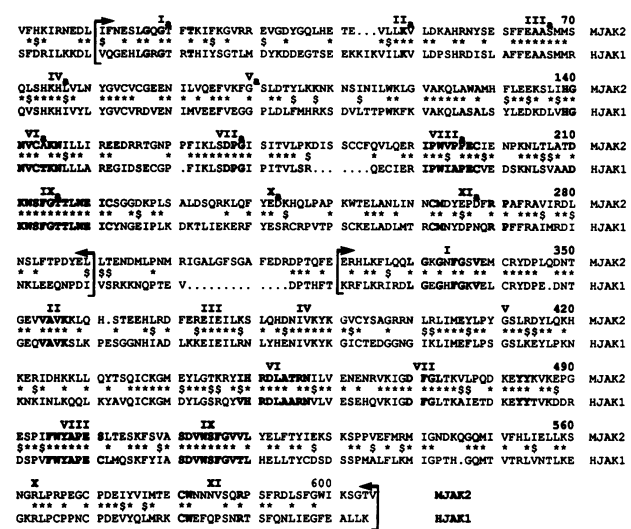


FIG. 5. Comparison of sequences in JAK1 and JAK2 kinase-related domains. The deduced amino acid sequence of murine JAK2 was compared with that of the human JAK1 by application of an alignment program of the Staden VAX-based suite of sequence analysis programs. Asterisks (*) denote identity, and dollar signs (\$) denote conservative substitutions. Sequences are numbered with respect to the JAK1 sequence. The extent of the domain 1 and PTK domains is shown by arrows above the amino acid sequence.

tive substitutions are included] for the PTK domain), clearly demonstrating their membership in the same family of PTKs. The absence of an SH2 domain in the JAK2 protein underlines the observations made in the case of the JAK1 protein and suggests that the lack of this regulatory element is a general feature of the JAK family of PTKs. An additional candidate for membership of this family is the *tyk2* PTK (21). Although the complete sequence of this PTK has not yet been described, many of the PTK domain sequence motifs common to JAK1 and JAK2 (for example, subdomain VIII, -FWYAPES-, a sequence motif characteristic of members of the JAK family) are present. Whether a second kinase-related domain is linked to the PTK domain of *tyk2* is not yet known.

PTK domain and kinase-related domain 1 are ancestrally related. The phylogenetic relationship of the catalytic domains of most of the protein kinases has been determined by using the tree-building program of Feng and Doolittle (7). In these endeavors we have been guided by the approach used by Hanks et al. (10) and have been able to replicate all of their data in addition to refining the family tree by including our own data. Figure 6 shows the phylogenetic relationship of the two kinase-related domains of the JAK1 protein to the rest of the kinase family. We conclude from this family tree that these two domains had a common ancestor which predated the development of the PTK subfamily. It is of interest to note that the kinase-related domains of the atrial natriuretic peptide (ANP)-receptor-guanylate cyclase family diverge at a point close by.

DISCUSSION

Protein phosphorylation plays a fundamental role in the regulation of a wide range of intracellular processes. It is in this respect not surprising that a large number of structurally distinct protein kinases have evolved (10, 13), each bearing a highly conserved protein kinase domain and each harnessing

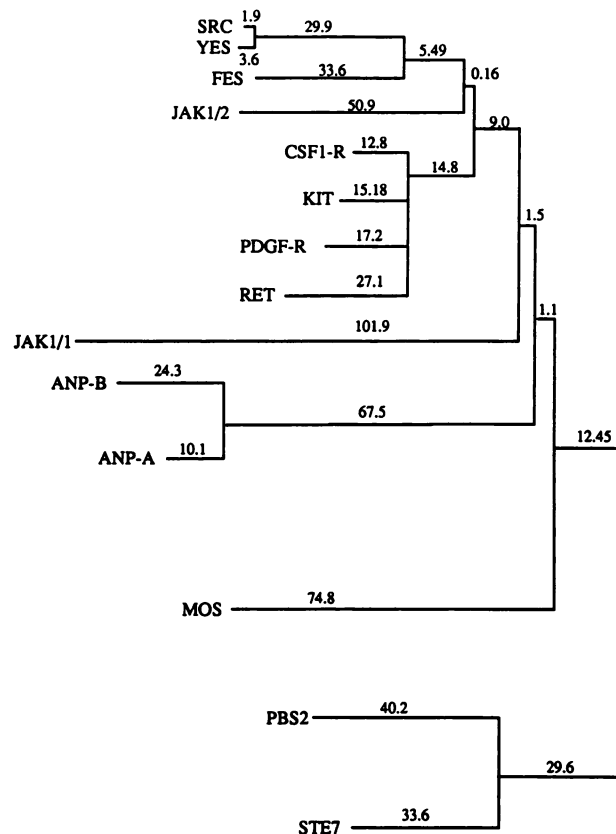


FIG. 6. Phylogenetic analysis of the two JAK1 kinaselike domains. The tree-building concept of Fitch and Margoliash (8) as implemented by Feng and Doolittle (7) and Hanks et al. (10) was used to generate a phylogenetic tree as described in Materials and Methods. In each case, the catalytic domain alone was used for comparison. The two kinase-related domains of the JAK1 protein were compared independently. Branch order is a function of structural similarity, and branch length is a function of sequence identity. Abbreviations: SRC, *c-src* (29); YES, *c-yes* (47); FES, *c-fes*, (57); CSF1-R, colony-stimulating factor-1 receptor (37); KIT, *c-kit* (59); PDGF-R, platelet-derived growth factor receptor A (58); RET, *c-RET* (48); ANP-A, ANP receptor A (43); ANP-B, ANP receptor B (42); MOS, *c-mos* (53); PBS2, polymyxin B antibiotic resistance gene product (1); STE7, sterile mutant wild-type allele gene product (49); JAK1/1, domain 1 of human JAK1; JAK1/2, PTK domain of human JAK1.

its catalytic activity to a particular appropriate metabolic signal. In the case of the PTK family, there is direct evidence implicating some members of this family (most notably the growth factor receptors) in the intracellular transmission of extracellular growth signals or differentiation signals or both (9, 16). The role(s) that other members of the PTK family of enzymes may play in related processes remains obscure. It has been presumed, however, that the paradigm exemplified by the growth factor receptor PTKs is carried through into the other members of the broader family of PTKs, namely, that the PTK catalytic domain acts as the "effector domain" of the protein, while the extracatalytic domain(s) serves to receive and process the appropriate input signal. Even though it has become clear that other important segments of these proteins can be defined (the SH2 domain [17, 39], SH3 domain [30, 45], and the intervening region in the PTK domain of the platelet-derived growth factor receptor [16]),

there has been no suspicion, to date, that the extracatalytic domains of the PTKs harbor anything other than interactive or regulatory modules. It is in this respect that the structure of the members of the JAK family of PTKs is particularly remarkable in its possession of a second kinase-related domain.

The second kinase-related domain, located 40 to 50 amino acids N terminal to the PTK catalytic domain, is a feature held in common between the two members of the JAK family of PTKs described here. There is sufficient similarity in the sequences and locations of the most highly conserved elements of this second kinase-related domain to the sequences and locations of their equally well conserved counterparts in the PTK and threonine/serine kinase families to suggest that this domain may confer a similar catalytic activity.

The existence of several other bifunctional proteins bearing protein kinase-related domains serves as a precedent for the expectation that both of these domains will be functionally active in the JAK family proteins. The S6 kinase (14) (two Ser/Thr kinase-related catalytic domains), the yeast GCN2 protein (55) (a kinase-related domain coupled to a tRNA synthetase-related domain), and the membrane ANP receptor-guanylate cyclase protein (43) (a kinase-related domain coupled to a guanylate cyclase domain) all exhibit a multidomain structure. However, the JAK proteins remain unique in the apparent tyrosine kinase activity of one of their catalytic domains. The unusual nature of domain 1 of JAK1 and JAK2 is reminiscent of the kinase-related domain of the ANP receptor, wherein, notwithstanding the overall structural similarity of the kinase-related domain to the protein kinase family at large, there are significant differences in detail between the precise amino acid sequences of certain key motifs and the protein kinase consensus sequences (10). These differences call into question the nature of the anticipated catalytic activity of this domain. For example, the presence of a conserved Asn residue in the conserved kinase motif (subdomain VI_a, -VHGNVCTKNL-) where all known protein kinases (as well as a number of bacterial phosphotransferases [2]) have an invariant Asp residue (10) is particularly unusual, although a similar variation is present in the ANP receptor-guanylate cyclase kinase-related domain (44) and in the kinase-related domain of Erb-B3 (20). The conserved Asp residue normally found in this subdomain has been shown to be essential for catalytic activity in the case of *v-fps* (31) and is likely to be a general theme in this family of proteins. However, it is conceivable that there are compensatory alterations in other subdomains that preserve the functional integrity of domain 1. Thus, while this is a highly significant deviation from the kinase consensus motifs defined by Hanks et al. (10), it is not sufficient grounds for dismissing the possibility that this domain has a kinase-related catalytic function. Rather, it suggests that other enzymatic properties, such as phosphatase activity or the capacity to phosphorylate substrates other than proteins, should be examined in the search for the function of this domain. Conversely, the significance of these sequence differences may lie in the choice of substrate for the catalytic activities encoded by these domains. Altered specificity may occur in the recognition of a different subset of substrate proteins or may encompass the phosphorylation of amino acids other than tyrosine, serine, and threonine. Alternatively, this same spectrum of amino acids may be recognized in a novel context. Whatever its function, the overall structure and conservation of this domain in the two JAK family members described here speaks to an important functional

(catalytic) role for this domain in the metabolic niche in which they serve the cell. Cogent arguments to explain the presence of two "catalytic" domains await the definition of the nature of the catalytic activity of domain 1.

Finally, in respect to the evolutionary development of a family of proteins such as the JAK family of PTKs, two scenarios, each with distinct but important ramifications, are possible. The first possibility is that the two kinase-related domains arose by duplication of a common ancestral catalytic domain followed by remodeling of one of these domains into the form described above. The second possibility is that the second kinase-related domain was "captured" by a protein already bearing a PTK-related domain, rather in the manner by which the SH2 domain of the nonreceptor PTKs has found its way into several other families of protein, such as the phospholipase C family (46) and the GTPase-activating protein of the *ras* proteins (51). The mosaic composition of most of these multidomain proteins, including the PTK family, and the observation that there are functional elements (like the SH2 domain) shared between families naturally suggest the wide application of this "modular" approach to their construction. If this mechanism is the basis by which the JAK family of PTKs has been constructed, it follows that the same kinase-related domain may also be a component of other non-PTK proteins. We are at present investigating this possibility.

ACKNOWLEDGMENTS

We are grateful to A. W. Burgess and A. R. Dunn for critical reading of the manuscript; Chris Hovens for advice on a nomenclature system; Lena Claesson-Welsh for the gift of an unamplified cDNA library; Ken Letwin, Mark Kamps, and Bart Sefton for antiphosphotyrosine antisera; Steven Hanks and Anne Marie Quinn for generously sharing their kinase database with us; R. Doolittle for kindly sending his tree-building program; Greg Thege for help with the implementation of the tree-building program; Gavin Reid and R. Simpson for synthesizing oligopeptides; Josie Discolo for synthesizing oligonucleotides; and A.-C. Andres for providing the murine mRNA.

Part of this work was supported by the Swiss National Fonds (grant 31-25726-88) and the Bernese Cancer League.

REFERENCES

- Boguslawski, G., and J. O. Polazzi. 1987. Complete nucleotide sequence of a gene conferring polymyxin B resistance on yeast: similarity of the predicted polypeptide to protein kinases. *Proc. Natl. Acad. Sci. USA* **84**:5848-5852.
- Brenner, S. 1987. Phosphotransferase sequence homology. *Nature (London)* **329**:21.
- Chowczynski, P., and N. Sacchi. 1987. Single step method for RNA isolation by acid guanidinium-thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156-159.
- Claesson-Welsh, L. A. Eriksson, B. Westermark, and C.-H. Heldin. 1989. cDNA cloning and expression of the human A-type platelet-derived growth factor (PDGF) receptor establishes structural similarity to the B-type PDGF receptor. *Proc. Natl. Acad. Sci. USA* **86**:4917-4921.
- Ebina, Y., L. Ellis, K. Jarnagin, M. Edery, L. Graf, E. Clauser, J.-H. Ou, F. Masiarz, Y. W. Kan, I. D. Godfine, R. A. Roth, and W. J. Rutter. 1985. The human insulin receptor cDNA: the structural basis for hormone activated transmembrane signaling. *Cell* **40**:747-758.
- Encyclopedia Britannica*. 1989. Janus, p. 155-156. *In Encyclopedia Britannica*, 11th ed. Vol. XV.
- Feng, D.-F., and R. F. Doolittle. 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* **25**:351-360.
- Fitch, W. M., and E. Margoliash. 1967. Construction of phylogenetic trees. *Science* **12**:279-284.
- Hafen, E., K. Basler, J. E. Edstroem, and G. M. Rubin. 1987. *sevenless*, a cell-specific homeotic gene of *Drosophila*, encodes a putative transmembrane receptor with a tyrosine kinase domain. *Science* **236**:55-63.
- Hanks, S. K., A. M. Quinn, and T. Hunter. 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**:42-52.
- Hao, Q., N. Heisterkamp, and J. Groffen. 1989. Isolation and sequence analysis of a novel human tyrosine kinase gene. *Mol. Cell. Biol.* **9**:1587-1593.
- Holtzman, D. A., W. D. Cook, and A. R. Dunn. 1987. Isolation and sequence of a cDNA corresponding to a src-related gene expressed in murine hemopoietic cells. *Proc. Natl. Acad. Sci. USA* **84**:8325-8329.
- Hovens, C., F. Clay, A. R. Dunn, A. G. Harpur, and A. F. Wilks. Unpublished data.
- Hunter, T. 1987. A thousand and one protein kinases. *Cell* **50**:823-829.
- Jones, S. W., E. Erikson, J. Blenis, J. L. Maller, and R. L. Erikson. 1988. A *Xenopus* ribosomal protein S6 kinase has two apparent kinase domains that are each similar to distinct protein kinases. *Proc. Natl. Acad. Sci. USA* **85**:3377-3381.
- Kamps, M. P., and B. M. Sefton. 1988. Identification of multiple novel polypeptide substrates of v-src, v-yes, v-fps, v-ros, and v-erbB oncogenic protein tyrosine kinases utilizing antisera against phosphotyrosine. *Oncogene* **2**:305-315.
- Kazlauskas, A., and J. A. Cooper. 1990. Autophosphorylation of the PDGF receptor in the kinase insert region regulates interactions with cell proteins. *Cell* **58**:1121-1133.
- Koch, C. A., M. Moran, I. Sadowski, and T. Pawson. 1989. The common homology region 2 domain of cytoplasmic signalling proteins is a positive effector of v-fps tyrosine kinase function. *Mol. Cell. Biol.* **9**:4131-4140.
- Kozak, M. 1984. Compilation and analysis of sequences upstream from the transcriptional start site in eukaryote mRNAs. *Nucleic Acids Res.* **12**:857-872.
- Kozma, S. C., S. M. S. Redmond, F. Xiano-Chang, S. M. Saurer, B. Groner, and N. E. Hynes. 1988. Activation of the receptor kinase domain of the trk oncogene by recombination with two different cellular sequences. *EMBO J.* **7**:147-154.
- Kraus, M. H., W. Issing, T. Miki, N. C. Popescu, and S. A. Aaronson. 1989. Isolation and characterisation of ERBB3, a third member of the ERBB/epidermal growth factor receptor family: evidence for overexpression in a subset of human mammary tumors. *Proc. Natl. Acad. Sci. USA* **86**:9193-9197.
- Krolewski, J. J., R. Lee, R. Eddy, T. B. Shows, and R. Della-Favera. 1990. Identification and chromosomal mapping of new human tyrosine kinase genes. *Oncogene* **5**:277-282.
- Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**:105-132.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
- Lee, M. G., and P. Nurse. 1987. Complementation used to clone a human homologue of the fission yeast cell cycle control gene *cdc2*. *Nature (London)* **327**:31-35.
- Letwin, K., S.-P. Yee, and T. Pawson. 1988. Novel protein tyrosine kinase cDNAs related to fps/fes and eph cloned using anti-phosphotyrosine antibody. *Oncogene* **3**:621-627.
- Lindberg, R. A., D. P. Thompson, and T. Hunter. 1988. Identification of cDNA clones that code for protein-tyrosine kinases by screening expression libraries with antibodies against phosphotyrosine. *Oncogene* **3**:629-633.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Marth, J. D., R. Peet, E. G. Krebs, and R. M. Perlmutter. 1985. A lymphocyte-specific protein-tyrosine kinase gene is rearranged and overexpressed in the murine T cell lymphoma LSTRA. *Cell* **43**:393-404.
- Martinez, R., B. Mathey-Prevot, A. Bernardis, and D. Baltimore. 1987. Neuronal pp60^{c-src} contains a six-amino acid insertion

- relative to its non-neuronal counterpart. *Science* 237:411–414.
30. Mayer, B. J., M. Hamaguchi, and H. Hanafusa. 1988. A novel viral oncogene with structural similarity to phospholipase C. *Nature (London)* 332:272–275.
 31. Moran, M. F., C. A. Koch, I. Sadowski, and T. Pawson. 1988. Mutational analysis of a phosphotransfer motif essential for v-fps tyrosine kinase activity. *Oncogene* 3:665–672.
 32. Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn, and H. A. Erlich. 1986. Specific amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symp. Quant. Biol.* 51:263–273.
 33. Park, M., M. Dean, K. Kaul, M. J. Braun, M. A. Gonda, and G. Vande Woude. 1987. Sequence of MET protooncogene cDNA has features characteristic of the tyrosine kinase family of growth-factor receptors. *Proc. Natl. Acad. Sci. USA* 84:6379–6383.
 34. Pawson, T. 1988. Non-catalytic domains of cytoplasmic protein tyrosine kinases: regulatory elements in signal transduction. *Oncogene* 3:491–495.
 35. Reichmann, E., R. Ball, B. Groner, and R. R. Friis. 1989. New mammary epithelial and fibroblastic cell clones in coculture form structures competent to differentiate functionally. *J. Cell Biol.* 108:1127–1138.
 36. Reid, H. H., A. F. Wilks, and O. Bernard. 1990. Two forms of the basic fibroblast growth factor receptor-like mRNA are expressed in the developing mouse brain. *Proc. Natl. Acad. Sci. USA* 87:1596–1600.
 37. Roebroek, A. J. M., J. A. Schalken, J. S. Verbeek, A. M. W. Van den Ouweland, C. Onnekink, H. P. J. Bloemers, and V. J. M. Van de Ven. 1985. The structure of the human c-fes/fps protooncogene. *EMBO J.* 4:2897–2903.
 38. Rothwell, V. M., and L. R. Rohrschneider. 1987. Murine c-fms cDNA: cloning sequence analysis and retroviral expression. *Oncogene Res.* 1:311–324.
 39. Sadowski, I., J. C. Stone, and T. Pawson. 1986. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus. *Mol. Cell. Biol.* 6:4396–4408.
 40. Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487–491.
 41. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* 74:5463–5467.
 42. Schultz, S., S. Singh, R. A. Bellet, G. Singh, D. J. Tubbs, H. Chin, and D. L. Garbers. 1989. The primary structure of a plasma membrane guanylate cyclase demonstrates diversity within this new receptor family. *Cell* 58:1155–1162.
 43. Singh, S., D. G. Lowe, D. S. Thorpe, H. Rodriguez, W. Kuang, L. J. Dangott, M. Chinkers, D. V. Goeddel, and D. L. Garbers. 1988. Membrane guanylate cyclase is a cell surface receptor with homology to protein kinases. *Nature (London)* 334:708–712.
 44. Smith, D. B., and K. S. Johnson. 1988. Single-step purification of polypeptides expressed in *E. coli* as fusions with glutathione S-transferase. *Gene* 67:31–40.
 45. Stahl, M. L., C. R. Ferez, K. Kelleher, R. W. Kriz, and J. Knopf. 1988. Sequence similarity of phospholipase C with the non-catalytic region of src. *Nature (London)* 332:269–272.
 46. Suh, P., S. H. Ryu, K. H. Moon, H. W. Suh, and S. G. Rhee. 1988. Cloning and sequence of multiple forms of phospholipase C. *Cell* 54:161–169.
 47. Sukegawa, J., K. Semba, Y. Yamanashi, M. Nishizawa, N. Miyajima, T. Yamamoto, and K. Toyoshima. 1987. Characterization of cDNA clones for the human c-yes gene. *Mol. Cell. Biol.* 7:41–47.
 48. Takahashi, M., and G. M. Cooper. 1987. *ret* transforming gene encodes a fusion protein homologous to tyrosine kinases. *Mol. Cell. Biol.* 7:1378–1385.
 49. Teague, M. A., D. T. Chaleef, and B. Errede. 1986. Nucleotide sequence of the yeast regulatory gene STE7 predicts a protein homologous to protein kinases. *Proc. Natl. Acad. Sci. USA* 83:7371–7375.
 50. Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA* 76:4350–4354.
 51. Trahey, M., G. Wong, R. Halenbeck, B. Rubinfeld, G. A. Martin, M. Ladner, C. M. Long, W. J. Crosier, K. Watt, K. Koths, and F. McCormick. 1988. Molecular cloning of two types of GAP complementary DNA from human placenta. *Science* 242:1697–1700.
 52. Ullrich, A., L. Coussens, J. S. Hayflick, T. J. Dull, A. Gray, A. W. Tam, J. Lee, Y. Yarden, T. A. Libermann, J. Schlessinger, J. Downward, E. L. V. Mayes, N. Whittle, M. D. Waterfield, and P. H. Seeburg. 1984. Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature (London)* 309:418–425.
 53. Watson, R., M. Oskarsson, and G. F. Vande Woude. 1982. Human DNA sequence homologous to the transforming gene (*mos*) of Moloney murine sarcoma virus. *Proc. Natl. Acad. Sci. USA* 79:4078–4082.
 54. Weinmaster, G., M. J. Zoller, M. Smith, E. Hinze, and T. Pawson. 1984. Mutagenesis of Fujinami sarcoma virus: evidence that tyrosine phosphorylation of P130^{src-fps} modulates its biological activity. *Cell* 37:559–568.
 55. Wek, R. C., B. M. Jackson, and A. G. Hinnebusch. 1989. Juxtaposition of domains homologous to protein kinases and histidyl-tRNA synthetases in GCN2 protein suggests a mechanism for coupling GCN4 expression to amino acid availability. *Proc. Natl. Acad. Sci. USA* 86:4579–4583.
 56. Wilks, A. F. 1989. Isolation of two putative protein tyrosine kinases by application of the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* 86:1603–1607.
 57. Wilks, A. F., and R. R. Kurban. 1988. Isolation and structural analysis of murine c-fes cDNA clones. *Oncogene* 3:289–294.
 58. Yarden, Y., J. A. Escobedo, W.-J. Kuang, T. L. Yang-Feng, T. O. Daniel, P. M. Tremble, E. Y. Chen, M. E. Ando, R. N. Harkins, U. Franke, V. A. Fried, A. Ullrich, and L. T. Williams. 1986. Structure of the receptor for platelet-derived growth factor helps define a family of closely related growth factor receptors. *Nature (London)* 323:226–232.
 59. Yarden, Y., W.-J. Kuang, T. Yang-Feng, L. Coussens, S. Munitz, T. J. Dull, E. Chen, J. Schlessinger, U. Franke, and A. Ullrich. 1987. Human proto-oncogene c-kit: a new cell surface receptor for an unidentified ligand. *EMBO J.* 6:3341–3351.
 60. Yarden, Y., and A. Ullrich. 1988. Growth factor receptor tyrosine kinases. *Annu. Rev. Biochem.* 57:443–478.
 61. Ziemiecki, A., R. G. Mueller, F. Xiao-Chang, N. E. Hynes, and S. Kozma. 1990. Oncogenic activation of the human trk protooncogene by recombination with the ribosomal large subunit protein L7a. *EMBO J.* 9:191–196.