S. Goodman, *J. Embryol. Morphol.* **78**, 169 (1983); F. R. Schram, in *The Biology of Crustacea*, vol. 1, *Systematics, the Fossil Record and Biogeography*, L. G. Abele, Ed. (Academic Press, New York, 1982), pp. 94–149.
30. G. E. Silvey and I. S. Wilson, *J. Exp. Biol.* **78**, 121 (1979).
31. The term "preadaptation" does not imply preordination. Gould and Vrba (*8*) have argued that this term should be replaced with "exaptation." We have retained the older term because it is more widely recognized.
32. K. Z. Lorenz, *Sci. Am.* **199**, 67 (December 1958).
33. J. E. Sulston *et al.*, *Dev. Biol.* **100**, 64 (1983).
34. W. F. Doolittle and C. Sapienza, *Nature (London)* **284**, 601 (1980).
35. K. G. Pearson and C. S. Goodman, *J. Comp. Neurol.* **184**, 141 (1979).
36. J. D. Steeves and K. G. Pearson, *J. Exp. Biol.* **103**, 47 (1983).
37. B. P. Nichols and C. Yanofsky, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5244 (1979); M. Kimura, *Nature (London)* **267**, 275 (1977).
38. E. Bizzi, W. Chapple, N. Hogan, *Trends Neurosci.* **5**, 395 (1982).
39. D. Marr, *Vision* (Freeman, San Francisco, 1983), pp. 19–27; W. E. Reichardt, in *Theoretical Approaches in Neurobiology*, W. E. Reichardt and T. Poggio, Eds. (MIT Press, Cambridge, MA, 1981), p. 1.
40. M. V. L. Bennett, in *Neural Mechanisms of Startle Behavior*, R. C. Eaton, Ed. (Plenum, New York, 1984), pp. 353–363.
41. D. H. Hubel and T. N. Wiesel, *Proc. R. Soc. London Ser. B* **198**, 1 (1977).
42. S. J. Gould, *Science* **216**, 380 (1982).
43. F. O. Albrecht, *The Anatomy of the Migratory Locust* (Athlone, London, 1953).
44. We thank R. L. Carroll, K. G. Pearson, and J. J. Wine for their criticism and discussion of this work; J. Kukalová-Peck for helpful suggestions and permission to reproduce Fig. 1B; and M. Soghikian for help in preparing the manuscript. Supported by NSF grant BNS 81-12431 and by the Natural Sciences and Engineering Research Council of Canada.

Research Articles

The Complete Primary Structure of Protein Kinase C—the Major Phorbol Ester Receptor

PETER J. PARKER, LISA COUSSENS, NICK TOTTY, LUCY RHEE, SUSAN YOUNG, ELLSON CHEN, SILVIA STABEL, MICHAEL D. WATERFIELD, AXEL ULLRICH

Protein kinase C, the major phorbol ester receptor, was purified from bovine brain and through the use of oligonucleotide probes based on partial amino acid sequence, complementary DNA clones were derived from bovine brain complementary DNA libraries. Thus, the complete amino acid sequence of bovine protein kinase C was determined, revealing a domain structure. At the amino terminal is a cysteine-rich domain with an internal duplication; a putative calcium-binding domain follows, and there is at the carboxyl terminal a domain that shows substantial homology, but not identity, to sequences of other protein kinase.

ANALYSIS OF GROWTH FACTORS AND THEIR ACTION HAS provided important insights into the mechanisms used to subvert the control of normal cell proliferation. Thus there is evidence that certain genes capable of transforming cells encode growth factors (*1*, *2*) or abnormal growth factor receptors (*3–5*); the expression of these genes allows cells to divide in a constitutive manner. In elucidating the responses of cells to growth factors, it has become evident that postreceptor events may also be in some way involved in cellular transformation. Therefore, a detailed molecular description of the intracellular pathways responsible for cell division induced by growth factors is necessary if we are to understand the normal mechanisms involved in growth factor action and in so doing to identify critical links open to subversion.

The phosphorylation of proteins plays a key role in regulating cellular functions (*6–8*). The kinases and phosphatases responsible for governing such phosphorylations are themselves targets for the action of growth factors, hormones, and other extrinsic agents participating in the control of cellular events (*6–8*). One of the major signal transduction pathways defined recently involves the enzyme protein kinase C (*9–11*), a multifunctional kinase that appears to play a central regulatory role akin to that of cyclic nucleotide-dependent and calcium-calmodulin–dependent enzymes.

Protein kinase C is a serine- and threonine-specific protein kinase that is dependent upon calcium and phospholipid for activity (*12*). However, at physiological calcium concentrations diacylglycerol is required for activity (*13*). Thus diacylglycerol has been defined as a second messenger responsible for the activation of protein kinase C in vivo (*9–11*). Agonist-induced generation of diacylglycerol has been widely described and forms part of a bifurcating signal pathway (*14*). It is thought that agonist-induced receptor-mediated activation of phospholipase C acts to generate two important second messengers from inositol phospholipids; the first, inositol 1,4,5-triphosphate, appears to be responsible for the release of calcium from intracellular stores (*15*) and the second, diacylglycerol, leads to protein kinase C activation (*13*). There is as yet only circumstantial evidence for the functioning of such a pathway in vivo (*16*).

From studies on protein kinase C in vitro, it has become apparent that those phorbol esters capable of tumor promotion can mimic the effect of diacylglycerol in enzyme activation (*17*). More recently, other structurally related and unrelated tumor promoters have also been shown to activate protein kinase C in vitro (*18–21*). The implication is that these tumor promoters elicit responses through protein kinase C and that activation of this enzyme is at least in part responsible for the activity of these hyperplasiogenic tumor-promot-

P. J. Parker, N. Totty, S. Young, S. Stabel, and M. D. Waterfield are at the Ludwig Institute for Cancer Research, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom. L. Coussens, L. Rhee, E. Chen, and A. Ullrich are in the Department of Developmental Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080.

ing agents. Indeed, there is evidence that phorbol esters do act via high-affinity binding sites to provoke cellular responses (*22*), and there is both indirect (*23–25*) and direct evidence (*26–28*) that identifies this high-affinity phorbol ester receptor as protein kinase C itself.

To probe the role of protein kinase C in growth factor action and also to define the way the function of this signal transducing protein is undermined by the action of certain tumor promoters, we have undertaken a structural analysis of the protein. We now report the complete amino acid sequence of protein kinase C obtained by sequence analysis and recombinant DNA techniques. The predicted domain structure of the enzyme has been probed with antibodies to specific regions of the polypeptide, and this provides a basis for a detailed understanding of the structure and function of this protein kinase.

Protein kinase C complementary DNA clones. Protein kinase C was purified from bovine brain as described (*27*). The yield of the enzyme was optimized by taking broad cuts at each purification step. The resulting material was approximately 70 percent pure as judged by sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis (PAGE) and specific activity (3500 units per milligram). Minor contaminating polypeptides were completely resolved from protein kinase C by gel-permeating high-performance liquid chromatography (HPLC) in SDS-phosphate (Fig. 1A). The pooled fractions from this column were dialyzed and digested with trypsin, and the peptides were purified by HPLC (Fig. 1, B to D). The high sensitivity gas phase sequence analysis of two of these tryptic peptides is shown in Fig. 1, E and F. In all, the sequence of more than 100 amino acid residues was established from analysis of eight tryptic peptides.

A complementary DNA (cDNA) library of 4×10^7 independent clones was prepared from total poly(A)$^+$-containing calf brain RNA with the use of the λgt10 vector system (*29*) and previously described procedures (*4*). Recombinant phage (2×10^6) were initially screened with radioactively labeled synthetic oligonucleotide probes that had been designed on the basis of partial amino acid sequences of peptides 3, 4, and 5 (Fig. 2A). The sequences of the double-stranded oligonucleotide probes 3, 4, and 5 were chosen on the basis of codon frequency analyses and were 54, 42, and 54 bases in length, respectively (Fig. 2A). Initial plaque hybridization with pooled probes 3, 4, and 5 resulted in three strongly hybridizing recombinant phage with cDNA inserts of similar length. Southern hybridization analysis of Eco RI–digested phage DNA with the three probes on separate blots revealed that all three hybridized with probe 5 only. Subsequent nucleotide sequence analysis, however, revealed that clone λbPKC21 contained an open reading frame that included sequences of peptide 5, as expected, and also of peptide 4. Only four codons of the corresponding probe 4 were predicted correctly (Fig. 3), resulting in 11 mismatches (26 percent) and only

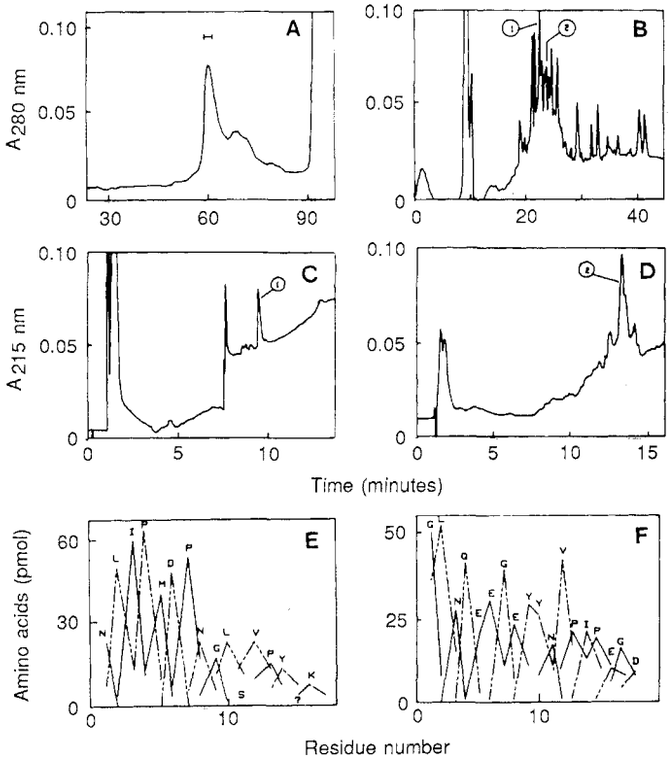

Fig. 1. Purification, trypsinolysis, and sequence analysis of protein kinase C. Protein kinase C was purified from bovine brain as described (*27*). In order to obtain optimum yields, broad cuts were taken at each step, and the final preparation was contaminated with polypeptides ≤60,000 daltons. In order to resolve these contaminants, samples (0.2 mg) of the preparation were subjected to gel permeation on a TSK 3000-GW column equilibrated in 0.1*M* sodium phosphate, *p*H 7.0, containing 0.1 percent SDS (A). In order to monitor recoveries during processing, samples of protein kinase C (1 to 2 μg) were autophosphorylated in the presence of ovalbumin as carrier (0.1 mg) and used to spike the bulk of the unlabeled material. The purified protein, which showed a single species on SDS-PAGE, was pooled from three runs as indicated (4.5 ml) and dialyzed twice against 5 liters of 20 percent methanol, 0.1*M* ammonium bicarbonate, and finally with two portions (1 liter each) of 0.05*M* ammonium bicarbonate. The protein was then digested with trypsin (1:50 by weight). The resulting tryptic peptides were resolved on a Vydak C_{18} reversed-phase column equilibrated in 0.08 percent trifluoracetic acid and developed with an acetonitrile gradient (0 to 60 percent) (B). Individual absorbing peaks from this column were further purified on a Pharmacia C_8 reversed-phase column (C and D). Absorbing material eluted from this second column was subjected to gas phase microsequencing. Examples of amino yields for two sequencer runs are shown in (E) and (F).

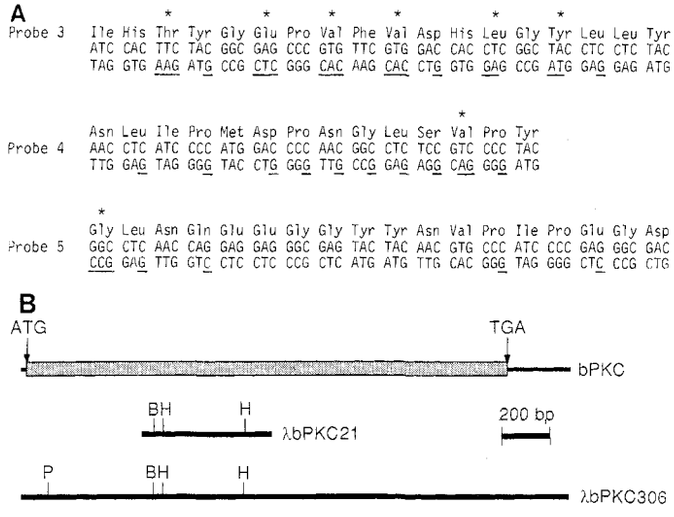

Fig. 2. Protein kinase C cDNA clones. (A) Synthetic oligonucleotide probes used for screening of cDNA libraries. Oligonucleotide sequences were predicted from codon usage frequency analyses (*37*). The bases underlined delineate differences between predicted and determined sequence (Fig. 3). The extent of mismatch in probes 3 (19 of 54) and 4 (11 of 42) resulted in neither probe hybridizing to λbPKC21 or λbPKC306 under the conditions employed. (B) Bovine brain protein kinase C cDNA clones. Recombinant λgt10 clone λbPKC21 and λbPKC306 cDNA inserts are shown as black bars. Locations of restriction endonuclease sites for Pst I (P), Bam HI (B) and Hind III (H) are indicated. The structure of bovine protein kinase C mRNA is schematically shown above with the coding region (shaded box) flanked by initiation (ATG) and termination (TGA) signals and untranslated sequences (heavy line).

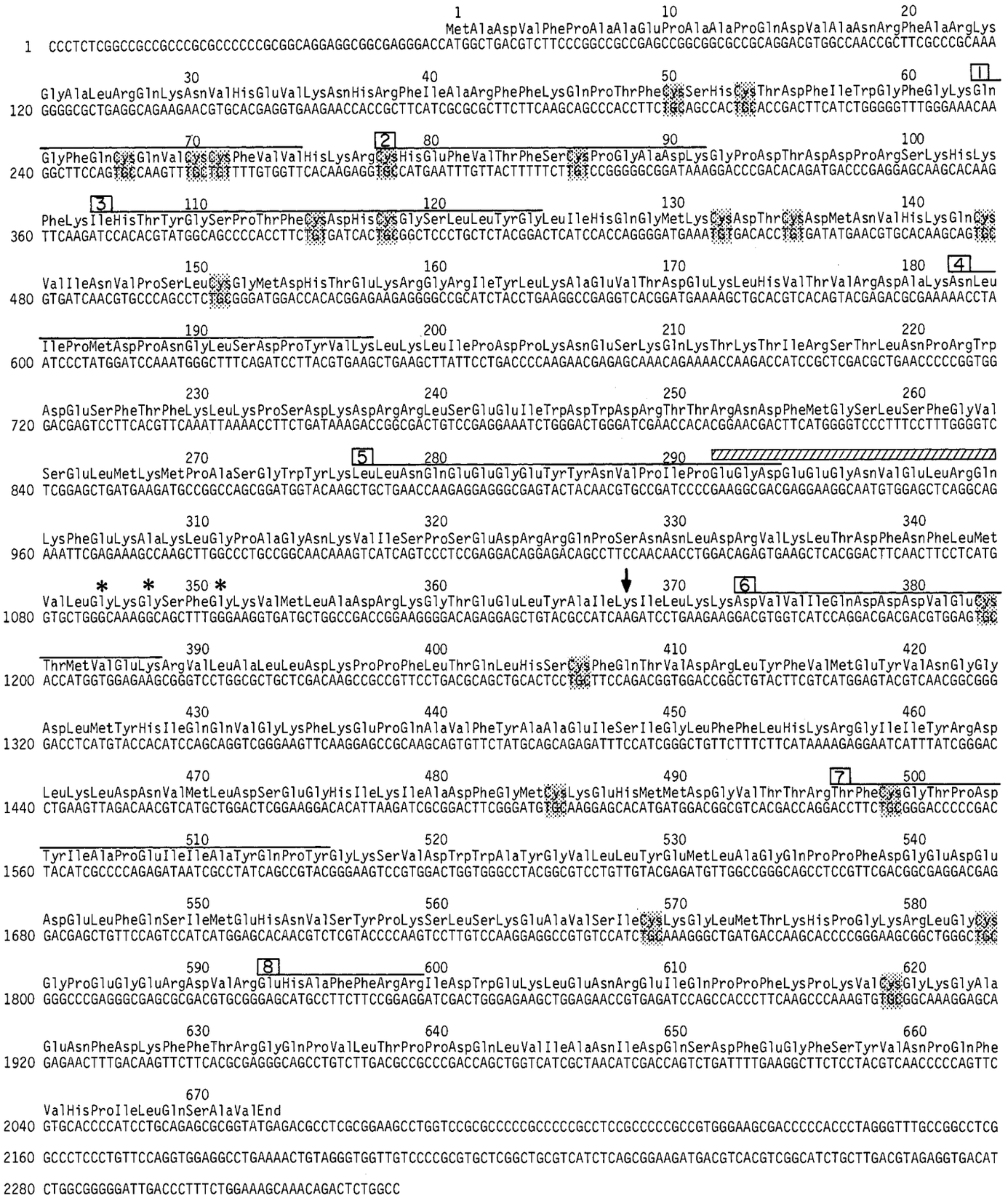

Fig. 3. Complete nucleotide sequence and deduced amino acid sequence of bovine brain protein kinase C. The nucleotide sequence was obtained from clones λbPKC21 and λbPKC306 by M13-based chain termination analysis (*38*) and is shown in 5′ to 3′ orientation. The amino acid sequence of the longest open reading frame beginning with an initiation codon is translated and contains sequences coding for eight protein kinase C peptides (overlined and sequentially numbered). Sequences constituting the proposed calcium-binding site are emphasized (cross-hatched bar), and residues predicted to be involved in formation of an ATP-binding site are marked by asterisks (Gly X Gly XX Gly) and an arrow (Lys^{368}).

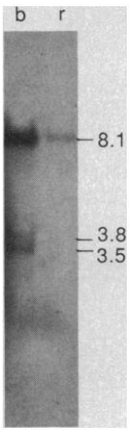

Fig. 4. Protein kinase C mRNA analysis. Polyadenylated RNA's from 4 μg of rat whole brain (r) and calf cerebellum (b) were analyzed by Northern blot hybridization (*39*); the entire λbPKC306 cDNA as a ^{32}P-labeled (10^8 cpm/μg) hybridization probe. RNA sizes are given in kilobases (kb) and were determined by comparison with RNA size standards of 9.5, 7.4, 4.4, 2.4, 1.4 and 0.3 kb (Bethesda Research Labs).

short stretches of matching sequence (Fig. 2A). Under the conditions used, this configuration was not sufficient to yield a hybridization signal.

Subsequent screening of another fraction (2×10^6 clones) of our bovine brain cDNA library with the radioactively labeled cDNA insert of λbPKC21 yielded clone λbPKC306 with a 2323-nucleotide cDNA insertion. Complete sequence analysis revealed the presence of an open reading frame (672 amino acids), which starts with a potential initiation codon. This ATG is flanked by sequences that fulfill the Kozak criteria for initiation codons (*30*) and lies within a 100-nucleotide GC-rich sequence (83 percent), which results in the absence of an in-frame stop codon upstream. Four additional cDNA clones were subsequently isolated, and all terminated within this GC-rich 5′ terminal sequence. The open reading frame initiated by this ATG codes for a 76.8-kD polypeptide, a molecular size close to that observed for the purified bovine brain enzyme (*27*). All experimentally determined protein kinase C peptide sequences are found within this predicted amino acid sequence, which provides further proof regarding the identity of clone λbPKC306. The 2016-nucleotide protein kinase C open reading frame is flanked by untranslated sequences of 50 (5′) and 258 (3′) nucleotides. Neither a poly(A) tail nor a polyadenylation signal (AATAAA) is contained in the λbPKC306 cDNA insert, suggesting that this clone was generated by nonspecific priming and that protein kinase C messenger RNA (mRNA) extends further 3′.

Protein kinase C mRNA's and their translation products. To examine protein kinase C mRNA size and complexity, we carried out Northern blot analysis with ^{32}P-labeled cDNA as a probe. There were one major and two minor mRNA hybridization signals of 8.1 kb, 3.8 kb, and 3.5 kb, respectively, in bovine brain poly(A)$^+$, and there was only one hybridization signal of ~8.6 kb in rat brain RNA (Fig. 4). Multiple mRNA's can be the product of alternative initiation, termination, or splicing events during transcription and posttranscriptional processing of a primary transcript. Alternatively, these hybridizing mRNA's may represent products of closely related but distinct genes [Coussens *et al.* (*31*) provide further information regarding this aspect]. All bovine protein kinase C RNA's are larger than our λbPKC306 cDNA and hybridize with coding as well as 3′ untranslated sequence probes, suggesting that extensions on 5′ or 3′ ends (or both) account for the observed mRNA size heterogeneity.

Since none of the protein kinase C cDNA clones described above have stop codons at the 5′ end of the open reading frame, a comparison was made between the primary protein kinase C translation products of poly(A)$^+$ RNA purified from bovine brain and in vitro transcripts from a cDNA construct covering the entire predicted amino acid sequence. The full-length cDNA in pSP65 was generated from λbPKC21 by taking the Xho I fragment covering the entire coding region and ligating into the Sal I site of the plasmid yielding pSP65-PKC1. The orientation was confirmed by restriction mapping. RNA from pSP65-PKC1 was generated in vitro with the use of SP6 polymerase, and this uncapped RNA was used as a template for translation in rabbit reticulocyte lysates.

Translation of poly(A)$^+$ RNA from bovine brain yielded a single major polypeptide that was specifically recognized by an antiserum to bovine protein kinase C. A single band at 76 kD is selectively extinguished by the addition of the protein kinase C–based peptide antigen to the reaction (Fig. 5A). Immunoprecipitation of the in

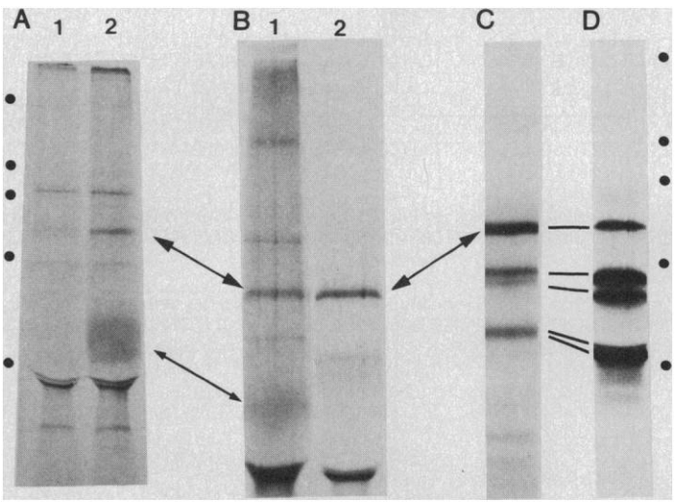

Fig. 5. In vitro translation products. An Xho I fragment covering the complete open reading frame of protein kinase C was excised from clone λbPKC306 and ligated into the Sal I site of the polylinker of plasmid pSP65 (Promega Biotec). The correct orientation of the insert was confirmed by restriction mapping. After linearization with Aha III the plasmid was used for in vitro transcription with SP6 polymerase (BioLabs) according to the manufacturer's protocols. The uncapped RNA produced was translated in vitro with a rabbit reticulocyte lysate (Amersham). In order to compare the translation product (or products) with authentic protein kinase C synthesized in vitro, total RNA was isolated from bovine brain and selected on oligo(dT)–cellulose. Approximately 0.5 μg of poly(A)$^+$ RNA from bovine brain was translated in vitro as described above. The translation products were boiled in SDS-PAGE sample buffer and diluted (1 to 10) with 1 percent deoxycholate, and 1 percent Triton X100 in phosphate-buffered saline and analyzed either directly by SDS-PAGE or immunoprecipitated. Samples were cleared with nonimmune rabbit serum and precipitated with antiserum 0442 in the presence and absence of competing peptide. Immune complexes were precipitated with protein A–Sepharose, washed with 0.5*M* NaCl in phosphate-buffered saline, and separated on 10 percent SDS-polyacrylamide gels. (A) In vitro translation of poly(A)$^+$ RNA from bovine brain yielded a single 76-kD species (large arrow) that was precipitated by protein kinase C antiserum 0442 (lane 2). Precipitation of the 76-kD protein and cross-reactive tulin (small arrow) was specifically blocked by inclusion of the peptide antigen (lane 1). Molecular size markers are indicated by dots and are in descending order: 200, 116, 93, 68, and 45 kD. (B) In vitro translation of poly(A)$^+$ RNA from bovine brain (lane 1) and manufactured putative full-length coding RNA transcribed in vitro from pSP65-PKC1 (lane 2). Antiserum 0442 was used to immunoprecipitate protein kinase C–related polypeptides. (C) In vitro translation of uncapped RNA transcribed in vitro from pSP65-PKC1 yielded a series of polypeptides ranging from the full-length translation product to several smaller species. These appear to be derived from initiation at internal methionines. Shorter translation products show apparent molecular weights of 60, 57, 54, 47 and 46 kD, and appear to be derived from initiation at each of the downstream methionine residues (residues 130 to 137, 153, 186, 256, and 266 to 268). Some smaller fragments (<40 kD) are also observed; these are likely to be generated from initiation of translation further downstream. These fragments are also precipitated by antibody 0442, which is visible on longer exposure of (B) lane 2. (D) Immunoprecipitation by polyclonal antiserum (BG 36) (*32a*) of in vitro translation products shown in (C).

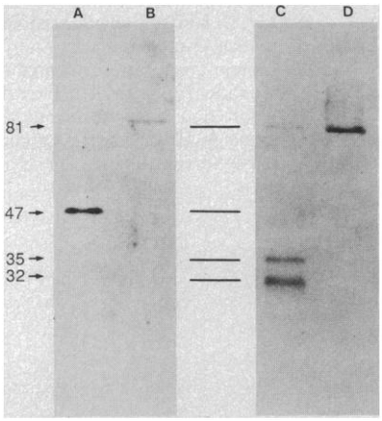

Fig. 6. Definition of the kinase domain. Antibodies to synthetic peptides based on protein kinase C residues 280 to 292 (serum 0442) and 317 to 328 (serum PP2) were obtained as described (*40*). Protein kinase C was purified to apparent homogeneity and subjected to SDS–polyacrylamide electrophoresis either without trypsin treatment (lanes B and D) or after digestion with trypsin at a final concentration of 20 μg/ml for 30 seconds at 30°C in the presence of 50 percent glycerol, 20 m*M* tris-HCl, *p*H 7.5, 2 m*M* EDTA, and 50 m*M* β-mercaptoethanol (lanes A and C). Trypsinization was terminated by addition of a tenfold molar excess of trypsin inhibitor and subsequent boiling. After SDS gel electrophoresis and transfer to nitrocellulose, protein kinase C and its fragments were subjected to immunoblot analysis with antibody 0442 (lanes C and D) or PP2 (lanes A and B). Arrows indicate molecular sizes of the immunoreactive polypeptides.

vitro translation products (Fig. 5B) from either bovine brain poly(A)$^+$ RNA (lane 1) or the SP6 polymerase transcript of the putative full-length open reading frame (Fig. 5B, lane 2, and D) yields a polypeptide of 76 kD in each case. This empirical value is close to that predicted from the sequence (76.8 kD). That these in vitro translation products are indistinguishable by size, provides some direct evidence that the initiator methionine shown in Figs. 2 and 3 is indeed the amino terminal residue of the primary translation product. Further circumstantial evidence for this is provided by the observation that the shorter in vitro translation products from these synthetic RNA's appear to be derived from each of the methionines downstream from residue 1 (Fig. 3); this pattern would not be consistent if methionine 130 was the major site of translational initiation (Fig. 5C).

The kinase domain. Protein kinase C can undergo proteolysis to generate a catalytically active fragment that is no longer dependent upon Ca^{2+} and phospholipid (*32*). This catalytic moiety can be readily generated in vitro by limited trypsinolysis (*32*); the fragment migrates as a ~50-kD species on gel filtration (*32*) and as a major band of ~47 kD on SDS-polyacrylamide gels (minor fragments of 45 and 43 kD are also present). We have used antisera to peptide fragments to define the limits of this catalytic domain. A Western blot of trypsinized protein kinase C shows that antiserum 0442 recognizes a species of 81 kD (intact protein kinase C) and a doublet at 33 to 35 kD (Fig. 6); in contrast, antiserum PP2 recognizes the intact polypeptide (81 kD) and a species at 47 kD. The polypeptide fragment at 47 kD comigrates with activity on gel filtration and therefore we can conclude that PP2 reacts with an epitope in the kinase domain, while 0442 reacts with the regulatory domain. These two antisera were made against amino acid sequences 280 to 292 (serum 0442) and 317 to 328 (serum PP2). Thus the limit on the catalytic domain is 43 kD (residues 292 to 717), this correlates with an empirical molecular size of 43 to 47 kD.

The reaction of antibody PP2 with only the uppermost of the 47- to 43-kD triplet suggests that conversion to smaller 45- and 43-kD species is associated with further cleavage at the amino terminus of the 47-kD fragment with subsequent loss of the PP2 epitope. It has not yet been possible to subfractionate this triplet in order to determine whether all these species are catalytically active.

Structural complications for protein kinase C function. Our results were designed to describe elucidation of the complete amino acid sequence of protein kinase C, the major receptor for phorbol esters. This was obtained through the isolation of cDNA clones coding for protein kinase C. The positive identification of these cDNA's has come from the presence of predicted tryptic peptides throughout the protein, whose sequences were in part predetermined from isolated bovine brain protein kinase C. In addition, translation of RNA that was derived from λbPKC-306 cDNA with the SP6 vector system (see below) yielded a 76-kD translation product that was recognized by a polyclonal antiserum to human protein kinase C (*32a*). Furthermore, antisera to a determined (antibody 0442) or predicted peptide sequence [antibody PP2, specific for alpha type protein kinase C; see legend to Fig. 5 and (*31*)] both recognize the purified bovine enzyme (see Fig. 5D).

The predicted amino acid sequence for protein kinase C shows a number of interesting features. At the amino terminus there is a tandem repeat that contains a series of six cysteine residues; the spacing of the cysteines between these two repeats is precisely conserved. Within these repeats is the sequence C-X_2-C-X_{13}-C-X_2-C; this type of pattern has been observed for various metalloproteins and also for certain DNA-binding proteins (*33*). While protein kinase C is known to bind calcium ions, this particular structure has not previously been shown to confer calcium binding and indeed other potential calcium-binding sites are present in the predicted sequence (as discussed below). The presence of this repeated pattern is of great interest; and it is now important to establish whether protein kinase C is itself a metalloprotein and whether this structure plays a role either in interaction with phospholipid (by analogy with the hydrophobic and ionic interactions of these structures with DNA) or indeed with DNA itself.

After the cysteine repeats there is a stretch of approximately 200 amino acids preceding the catalytic domain. Within this sequence lies the only potential calcium-binding site that comes close to the "E-F hand" structure characteristic of calmodulin and related calcium-binding proteins (*34*). This sequence (residues 292 to 303) is a predicted coil lying amino terminal to a predicted α-helix, with a glycine residue at the crucial turn point (Gly^{297}). This residue is surrounded by alternating amino acids capable of coordinating the calcium ion through their side chains (Glu^{292}, Asp^{294}, Glu^{296}, Asn^{298}, Glu^{300}, Gln^{303}). While there would appear to be an F helix, the presence of proline residues on the amino side of this putative binding site would not provide an equivalent E helix. There is no direct evidence that this is indeed the calcium-binding site, or that such a site in protein kinase C should fit the E-F hand model, indeed the presence of glutamic residues in such a structure would be unusual. Nevertheless, it is well documented that protein kinase C, like calmodulin, is inhibited by the phenothiazines and it might be surmised that some structural homology would exist between the calcium-binding sites of these proteins. Recent reports have suggested the possible existence of two Ca^{2+}-binding domains, one associated with catalytic activation, the other with phospholipid (*35*, *36*). The primary sequence does contain other regions that may function as Ca^{2+}-binding sites, and there is now an opportunity to directly assess the structure-function relations of these individual regions.

The catalytic domain shares substantial homology with other serine, threonine, and tyrosine protein kinases (Fig. 7). Thus by analogy with these protein kinases, one would predict that lysine-368 would lie at the ATP (adenosine triphosphate)-binding site. This residue is 17 amino acids carboxyl terminal to the last of the three conserved glycines that are also associated with the predicted nucleotide-binding site. The remainder of this domain shows further homology with these kinases; this reflects the conservation of particular sequences between all the members of this family and one would predict a functional role for these conserved residues.

A definitive proof of the location of the amino terminus has not been provided. However there are a number of considerations that

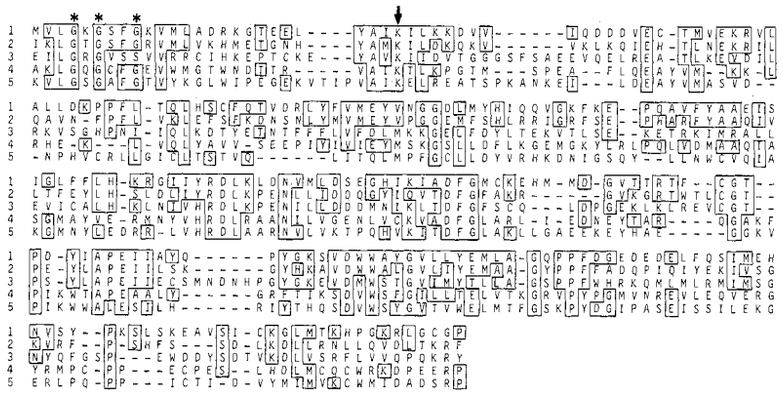

Fig. 7. Kinase domain homology to other protein kinases. The sequences of the putative catalytic (kinase) domains of a series of protein kinases are compared. Homologies to protein kinase C are shown in boxes. The asterisks indicate residues thought to be involved in the nucleotide binding site. (1) Protein kinase C; (2) cyclic AMP–dependent protein kinase (*41*); (3) γ-subunit of phosphorylase kinase (*42*); (4) v-*src* (*43*); (5) EGF-receptor (4).

are consistent with the present assignment: (i) The predicted size 77 kD is in reasonable agreement with previous determinations of the apparent molecular size of the purified protein; (ii) a related protein (*31*) is colinear at the amino terminus and retains the same amino-terminal Met-Ala-Asp triplet within a stretch that is otherwise divergent; (iii) the proteins synthesized in vitro from purified mRNA and from RNA transcribed from a putative full-length coding sequence display an identical apparent molecular size on SDS-polyacrylamide electrophoresis. Direct determination of the amino-terminal sequence has not been possible. The lack of success in obtaining this amino-terminal sequence suggests that the polypeptide might be blocked; the nature of this blocking group is not known at present.

There is evidence that protein kinase C in vivo may under certain circumstances be cleaved into a catalytically active fragment that displays an apparent molecular size of 50 kD on gel filtration. The sequence for protein kinase C would suggest that the region of this proteolytic cleavage is between the putative calcium-binding domain and the region encoding the kinase domain; within this sequence are a number of basic residues that may act as calpain or tryptic cleavage sites. Through the use of antibodies to specific regions of the polypeptide, it has been possible to identify the region that is cleaved by trypsin in vitro. Thus the catalytic domain (which retains activity) carries antigenic epitopes contained in residues 317 to 328, while not retaining those in residues 280 to 292. This puts an upper limit on the size of the fragment carrying catalytic activity and this is entirely consistent with the region of homology shared with other protein kinases (Fig. 7).

The proteolytic activation of protein kinase C generates a fragment with kinase activity that is not dependent on calcium and phospholipid (*32*). It would appear then that the regulatory domain acts to maintain the catalytic activity, and removal of this domain by proteolysis (or perhaps partial denaturation) leads to activation. Presumably the binding of calcium and phospholipid provokes a conformational change with consequent activation. It is thus possible to form a model for the activation of protein kinase C based upon these considerations (Fig. 8). It is proposed that the "hinge" region between the regulatory and catalytic domains is the exposed site of cleavage (residues 292 to 317). This type of structure has functional homology to the cyclic GMP–dependent protein kinase (GMP, guanosine monophosphate) where within a single polypeptide, a regulatory domain maintains a catalytic domain in an inactive state, until ligand is bound. This similarity suggests that perhaps as predicted for cyclic GMP–dependent protein kinase, protein kinase C was derived from the fusion of two genes and that these define the basic domain structure of the polypeptide.

When we used the isolated cDNA clones to probe the size of the mRNA for protein kinase C, we observed three major size species. We have some initial data indicating that the 8.1-kb and 3.8- or 3.5-kb messages all direct the synthesis of a 76-kD polypeptide that is recognized by antibodies to protein kinase C. This would suggest that these mRNA's are derived by alternative splicing or alternative processing outside the coding region (that is, in the 5′ or 3′ untranslated region) or that splicing variants of the coding region are indistinguishable by our criteria.

The ability of phorbol esters to bind to and activate protein kinase C is shared by a number of other tumor promoters including mezerin, aplysiatoxin, debromoaplysiatoxin, and teleocidin (*18–21*). Elucidation of the protein kinase C–binding site for these diverse structures should provide insight into the chemical features that appear to be in part responsible for tumor-promoting activity. The definition of the primary structure of protein kinase C allows us now to probe this structure, delineate the binding site and perhaps design antagonists.

Protein kinase C evidently does not show homology to any oncogenes that have been described to date. Given the important role that protein kinase C can play in the control of growth and differentiation, it might be anticipated that the expression of a mutated form of this protein, in particular cell lineages, may induce constitutive proliferation. However, it should be noted that, physiologically, protein kinase C activation by diacylglycerol is probably coupled to mobilization of Ca^{2+} and that the cellular responses are therefore a consequence of both events (*14*). Thus it may transpire that protein kinase C lies too far down a controlling hierarchy for

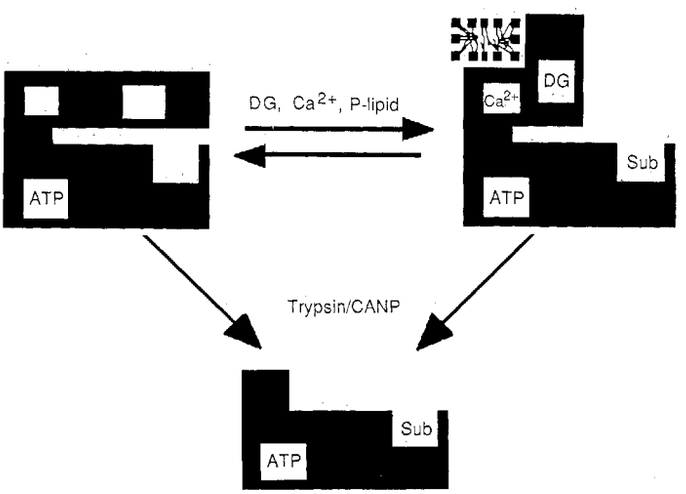

Fig. 8. Model for activation of protein kinase C. Abbreviations: DG, diacylglycerol; P-lipid, phospholipid; CANP, calpain; Sub, substrate-binding site.

aberrant activation to generate such a constitutive proliferation response (at least in the absence of other events). The generation of suitable protein kinase C cDNA constructs should make it possible to determine whether the expression of this kinase is abnormal in any human neoplasias.

REFERENCES AND NOTES

1. M. D. Waterfield *et al.*, *Nature (London)* **304**, 35 (1983).
2. R. F. Doolittle *et al.*, *Science* **221**, 275 (1983).
3. J. Downward *et al.*, *Nature (London)* **307**, 521 (1984).
4. A. Ullrich *et al.*, *ibid.* **309**, 418 (1984).
5. C. J. Sherr *et al.*, *Cell* **41**, 665 (1985).
6. P. Cohen, *Bioessays* **2**, 63 (1985).
7. ———, *Curr. Top. Cell. Reg.* **27**, 23 (1985).
8. T. Hunter and J. A. Cooper, *Annu. Rev. Biochem.* **54**, 897 (1985).
9. Y. Nishizuka, *Science* **225**, 1365 (1984).
10. ———, *Nature (London)* **308**, 693 (1984).
11. ———, *J. Natl. Cancer Inst.* **76**, 363 (1985).
12. Y. Takai *et al.*, *J. Biol. Chem.* **254**, 3692 (1979).
13. A. Kishimoto, Y. Takai, T. Mori, U. Kikkawa, Y. Nishizuka, *ibid.* **255**, 2273 (1980).
14. M. J. Berridge and R. F. Irvine, *Nature (London)* **312**, 315 (1984).
15. H. Streb, R. F. Irvine, M. J. Berridge, I. Schulz, *ibid.* **306**, 67 (1983).
16. Y. Kawahara, Y. Takai, R. Minakuchi, K. Sano, Y. Nishizuka, *Biochem. Biophys. Res. Commun.* **97**, 309 (1980).
17. M. Castagna *et al.*, *J. Biol. Chem.* **257**, 7847 (1982).
18. H. Fujiki *et al.*, *Biochem. Biophys. Res. Commun.* **120**, 339 (1984).
19. R. Miyake *et al.*, *ibid.* **121**, 649 (1984).
20. A. Couturier, S. Bazgar, M. Castagna, *ibid.*, p. 448.
21. J. P. Arcoleo and I. B. Weinstein, *Carcinogene* **6**, 213 (1985).
22. M. K. L. Collins and E. Rozengurt, *J. Cell. Physiol.* **122**, 42 (1982).
23. J. E. Niedel, L. J. Kuhn, G. R. Vandenbark, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 36 (1983).
24. C. L. Ashendel, J. M. Staller, R. K. Boutwell, *Cancer Res.* **43**, 4333 (1983).
25. K. L. Leach, M. L. James, P. M. Blumberg, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 4208 (1983).
26. U. Kikkawa, Y. Takai, Y. Tanaka, R. Miyake, Y. Nishizuka, *J. Biol. Chem.* **258**, 11442 (1983).
27. P. J. Parker, S. Stabel, M. D. Waterfield, *EMBO J.* **3**, 953 (1984).
28. T. Uchida and C. R. Filburn, *J. Biol. Chem.* **259**, 12311 (1984).
29. T. Huynh, R. Young, R. Davis, in *Practical Approaches in Biochemistry,* D. Grover, Ed. (IRL Press, Oxford, 1984).
30. M. Kozak, *Microbiol. Rev.* **47**, 1 (1983).
31. L. Coussens *et al.*, *Science* **233**, 859 (1986).
32. M. Inoue, A. Kishimoto, Y. Takai, Y. Nishizuka, *J. Biol. Chem.* **252**, 7610 (1977).
32a. M. J. Fry, A. Grebhardt, P. J. Parker, J. G. Foulkes, *EMBO J.* **4**, 3173 (1985).
33. J. M. Berg, *Science* **232**, 485 (1986).
34. L. J. van Eldik, J. G. Zendoqui, D. R. Marshak, D. M. Wattenson, *Int. Rev. Cytol.* **77**, 1 (1982).
35. M. Wolf, H. Levine, W. S. May, P. Cuatrecasas, N. Sahyoun, *Nature (London)* **317**, 546 (1985).
36. W. S. May, N. Sahyoun, M. Wolf, P. Cuatrecasas, *ibid.*, p. 549.
37. R. Grantham, C. Gautier, M. Govy, M. Jacobzone, R. Mercier, *Nucleic Acids Res.* **9**, 43 (1981).
38. J. Messing, R. Crea, P. H. Seeburg, *ibid.*, p. 309.
39. H. Lehrach, D. Diamond, J. M. Wozney, H. Boedtker, *Biochemistry* **16**, 4743 (1977).
40. P. J. Parker *et al.*, in preparation.
41. S. Shoji, L. H. Ericsson, K. A. Walsh, E. H. Fischer, K. Titani, *Biochemistry* **22**, 3702 (1983).
42. E. M. Reimann *et al.*, *ibid.* **23**, 4185 (1984).
43. A. P. Czernilofsky *et al.*, *Nature (London)* **287**, 198 (1980).

44. We thank A. Pereira for expert preparation of the manuscript.

17 July 1986; accepted 25 July 1986

Multiple, Distinct Forms of Bovine and Human Protein Kinase C Suggest Diversity in Cellular Signaling Pathways

LISA COUSSENS, PETER J. PARKER, LUCY RHEE, TERESA L. YANG-FENG, ELLSON CHEN, MICHAEL D. WATERFIELD, UTA FRANCKE, AXEL ULLRICH

A new family of protein kinase C–related genes has been identified in bovine, human, and rat genomes. The alpha-, beta-, and gamma-type protein kinase sequences are highly homologous, include a kinase domain, and potential calcium-binding sites, and they contain interspersed variable regions. The corresponding genes are located on distinct human chromosomes; the possibility of even greater genetic complexity of this gene family is suggested by Northern and Southern hybridization analyses.

ACTIVATION OF THE PHORBOL ESTER RECEPTOR PROTEIN kinase C (PKC) by calcium ions and the second messenger diacylglycerol is thought to play a central role in the induction of cellular responses by a variety of ligand-receptor systems (*1–3*) and in regulation of cellular responsiveness to external stimuli (*4, 5*). Structural characterization of bovine PKC (bPKC) by partial peptide sequencing and complementary DNA (cDNA) cloning has now provided information that facilitates the analysis of its multiple activities, including diacylglycerol and calcium binding, membrane association, serine and threonine residue phosphorylation, and substrate recognition (*6*). It remains unclear, however, how a single molecule can mediate and control a variety of diverse signaling pathways as has been suggested for PKC (*1–3*). We report here, for the first time, the existence of two bovine and human genes that encode polypeptide sequences with remarkable homology to PKC. We find homologs of bPKC-related messenger RNA's (mRNA's) in human and rat tissues and demonstrate that corresponding genes map to distinct chromosomal loci of the human genome. Southern hybridization analysis (*7*) suggests the existence of additional members of this newly identified PKC gene family.

Screening of a bovine brain cDNA library with a pool of synthetic oligonucleotide probes that had been designed on the basis of

L. Coussens, L. Rhee, E. Chen, and A. Ullrich are in the Department of Developmental Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080. P. J. Parker and M. D. Waterfield are at the Ludwig Institute for Cancer Research, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, United Kingdom. T. L. Yang-Feng and U. Francke are in the Department of Human Genetics, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06510.