Molecular Cloning of cDNAs Encoding Two Isoforms of the Catalytic Subunit of Protein Phosphatase 2A

Stuart R. Stone,* Jan Hofsteenge, and Brian A. Hemmings
Friedrich Miescher-Institut, CH-4002 Basel, Switzerland
Received August 5, 1987; Revised Manuscript Received September 9, 1987

ABSTRACT: Clones coding for the catalytic subunit of one of the major protein phosphatases (type 2A) were isolated from a porcine cDNA library. Sequence analysis indicated that two different mRNA species coded for this enzyme. The deduced amino acid sequences of the two forms (α and β) of the enzyme were 98% identical and showed 95% identity with the partial sequence of the rabbit enzyme determined by amino acid sequencing. The use of specific oligonucleotide probes indicated that the mRNAs coding for the α and β forms were about 2 kilobases in length, present in equal amounts in a porcine cell line (LLC-PK_1), and were the products of two distinct genes. Southern analysis using the coding region of the α phosphatase cDNA as a probe suggested the existence of additional related phosphatase genes.

Protein phosphorylation plays a central role in the control of cellular functions. The kinases and phosphatases responsible for regulating this phosphorylation appear themselves to be subject to control through the action of hormones and growth factors (Cohen, 1985; Krebs, 1986). Recent work in the field of protein phosphorylation has concentrated on the characterization of protein kinases. A complete understanding of the regulation of cellular processes by protein phosphorylation will require, however, investigation of the complexity and regulation of protein phosphatases. Four different forms of protein phosphatase catalytic subunits capable of dephosphorylating phosphoserine and phosphothreonine residues have been isolated and characterized [reviewed by Ballou and Fischer (1986)]. These phosphatases have been classified into two types, depending on their sensitivity to inhibitor and modulator proteins and substrate specificity (Ingebritsen & Cohen, 1983). Type 1 protein phosphatase catalytic subunit is characterized by its sensitivity to two heat-stable inhibitor proteins. Type 2 protein phosphatase catalytic subunits are not affected by the same inhibitors and can be subdivided into three groups, termed 2A, 2B, and 2C, all of which have been purified to homogeneity and characterized enzymologically [see Ballou and Fischer (1986)]. The type 2A phosphatase can be isolated in a number of different forms, which consist of a catalytic subunit associated with different regulatory subunits (Ballou & Fischer, 1986).

To investigate the role of protein phosphatases in the regulation of cellular functions, we have initiated the molecular cloning of cDNAs coding for these proteins. This paper reports the molecular cloning of cDNAs coding for two forms of the catalytic subunit of phosphatase 2A. Sequence and Southern analysis indicated that these two cDNAs are the products of separate genes. In addition, over half of the sequence of the rabbit skeletal muscle phosphatase 2A catalytic subunit was obtained by peptide sequencing. Comparison of the sequences for the porcine and rabbit enzymes indicated that the phosphatase 2A catalytic subunit is highly conserved between species, which is consistent with a central regulatory role for this protein.

Experimental Procedures

Protein Purification. The phosphatase 2A catalytic subunit was purified as previously described (Tung et al., 1984) except that, following chromatography on Sephadex G-100, the enzyme was further purified by ion-exchange chromatography on a Mono Q column with a gradient of 0–0.5 M NaCl in 25 mM Tris-HCl[1] buffer, pH 8.0. The fractions containing protein phosphatase activity, which eluted at about 0.25 M NaCl, were analyzed by SDS–polyacrylamide gel electrophoresis and found to be homogeneous.

Protein Sequence Analysis. The purified phosphatase was reduced and carboxymethylated, and digests were made with CNBr, trypsin, and the protease from *Staphylococcus aureus* V_8 as described by Hofsteenge et al. (1983). The digests were fractionated on C_8 or C_{18} reversed-phase columns by using a linear gradient of acetonitrile (0–70%) in 0.1% (v/v) trifluoroacetic acid. Amino acid sequence analysis was performed on an Applied Biosystems gas-phase sequencer (Hewick et al., 1981).

RNA and DNA Isolations. Total RNA from LLC-PK_1 cells (Hull et al., 1976) was isolated as described by Chirgwin et al. (1979), and poly(A)+ RNA was selected by chromatography on oligo(dT)–cellulose (Aviv & Leder, 1972). Genomic DNA from LLC-PK_1 cells was isolated by the method of Blin and Stafford (1976).

cDNA Cloning. A cDNA library was constructed by using poly(A)+ RNA isolated from the porcine kidney cell line LLC-PK_1. cDNA was prepared by the method of Gubler and Hoffman (1983) as described previously (Hemmings et al., 1986). The library contained 10^7 recombinants, and about 10^6 plaques of the amplified library were screened with two oligonucleotide probes: TTCCANAC(A/G)TTNGC(A/G)TTNCC(A/G)TA, oligonucleotide 272, and AAGGT-(T/C)TC(T/C)TTAAT(A/G)TC(T/C)TGGCC, oligonucleotide 273 (see text). Oligonucleotides were end-labeled with [γ-^{32}P]ATP (>3000 Ci/mmol), and filters were hybridized with both oligonucleotides (2×10^6 cpm/mL) in 6 × SSC/2 × Denhardts/0.05% sodium pyrophosphate/0.1% SDS/100 μg mL^{-1} denaturated herring sperm DNA at 37 °C for 16 h. The filters were washed at 37 °C in 3 × SSC/0.1% SDS for 1 h, followed by a final wash at 42 °C for 15 min under the same conditions.

Southern and Northern Analysis. For genomic Southern analysis, DNA was digested with either *Eco*RI, *Bam*HI,

* Author to whom correspondence should be addressed.

[1] Abbreviations: Tris, tris(hydroxymethyl)aminomethane; SDS, sodium dodecyl sulfate; SSC, standard saline citrate; bp, base pair(s); kDa, kilodalton(s).

0006-2960/87/0426-7215$01.50/0 © 1987 American Chemical Society

*Hin*dIII, or *Bgl*II. The digested DNA was fractionated on 1% agarose gels and transferred to nitrocellulose as described by Southern (1975). The nitrocellulose filters were baked, prehybridized with 6 × SSC/5 × Denhardts/0.5% SDS, and then hybridized at 60 °C with the 1140-bp *Eco*RI fragment of λPhos 4 (10^6 cpm/mL), which had been labeled by the random priming method of Feinberg and Vogelstein (1984) to a specific activity of about 10^9 cpm/μg. The filters were washed with 2 × SSC/0.1% SDS at 60 °C for 1 h (low stringency), subjected to autoradiography for 3 days at −70 °C with intensifying screens, and then washed again with 0.5 × SSC/0.1% SDS at 65 °C for 2 h (high stringency) and autoradiographed as above. For analysis with the specific α and β probes, the oligonucleotides were end-labeled with [γ-^{32}P]ATP (>3000 Ci/mmol), and hybridization was carried out at 50 °C in 6 × SSC/2 × Denhardts/0.1% SDS/0.05% sodium pyrophosphat containing 20 μg/mL tRNA and 10^6 cpm/mL of the probe. After hybridization, the filters were washed at 55 °C with 1 × SSC/0.1% SDS before exposure.

For Northern analysis, poly(A)+ RNA was fractionated in formaldehyde/1.0% agarose gels and transferred to nitrocellulose filters by using 10 × SSC (Thomas, 1980). After transfer the filters were baked and prehybridized with 6 × SSC/2 × Denhardts/0.1% SDS/0.05% sodium pyrophosphate at 60 °C. Hybridization was carried out as described for Southern analysis. Filters were washed at 60 °C with 1 × SSC/0.1% SDS for 1– 2 h and then subjected to autoradiography.

DNA Sequence Analysis. cDNA fragments were subcloned into M13mp10 or -11 and sequenced by the dideoxynucleotide chain termination method (Sanger et al., 1977). All clones were sequenced on both strands by using specific oligonucleotide primers. For regions with a high content of G and C residues, 7-deaza-dGTP was used instead of dGTP (Mizusawa et al., 1986).

Results

Partial Amino Acid Sequence of the Rabbit Protein Phosphatase 2A Catalytic Subunit. The protein phosphatase 2A catalytic subunit isolated from rabbit skeletal muscle, both in its native form and after reduction and carboxymethylation, was not susceptible to Edman degradation. Therefore, digests of the denatured protein were made with trypsin, CNBr, and *S. aureus* V8 protease. The resulting peptides were purified by reversed-phase HPLC and a number were sequenced. Two peptides, -Y-G-N-A-N-V-W-K- and -G-Q-D-I-K-E-T-F-, were selected and used to prepare two sets of oligonucleotide probes 23 bases in length. The first (272) was a mixture of 512 sequences containing all possible codons while the second (273) contained 16 sequences and was based partially on codon preference (Lathe, 1985). All protein sequence data obtained are shown in Figure 3.

cDNA Cloning of the Porcine 2A Catalytic Subunit. Analysis of protein phosphatase activity in LLC-PK_1 cell extracts demonstrated that the type 2A was the dominant phosphatase activity synthesized by this cell line (data not shown). Therefore, it was decided to screen a cDNA library constructed from poly(A)+ RNA from the LLC-PK_1 cell line by using the oligonucleotide probes described above. This strategy was based on the assumption that the catalytic subunits of the protein phosphatases were likely to be highly conserved between species (see below). The two oligonucleotide probes (272 and 273) were used simultaneously to screen about 10^6 phage plaques, and nine positives were identified and plaque purified. Restriction mapping and DNA sequence analysis revealed, however, that five of these positives were the same cDNA clone. Thus, five unique clones were identified by the double oligonucleotide probing strategy. Subsequent analysis revealed that all the clones hybridized exclusively to oligonucleotide 272, which contained a predicted 512 sequences. Sequencing of the cDNA clones indicated that oligonucleotide 273 contained 6 mismatches in 23 bases, which explains its failure to hybridize to the cDNA clones.

Sequence Analysis of Protein Phosphatase 2A cDNAs. The nucleotide sequence of the clone λPhos 4 contained an open reading frame of 927 nucleotides starting from the assigned initiator ATG codon; the sequence of λPhos 4 corresponded to nucleotides −179 to 945 of the sequence presented in Figure 1 (the A of the putative initiation codon is numbered as 1). The deduced amino acid sequence of this reading frame contained sequences homologous to all of those obtained by amino acid sequencing of peptides from the rabbit protein as shown in Figure 3. This high homology identified λPhos 4 as a clone coding for a catalytic subunit of protein phosphatase 2A and confirmed that the proposed reading frame was correct. The proposed initiator ATG codon was preceded by a G/C-rich region, and the nucleotides flanking this codon conformed to the consensus sequence for initiator codons (Kozak, 1984). The nucleotide sequence of the clone λPhos 1 overlapped with λPhos 4 in the coding region and extended to the poly(A) tail of the mRNA that was preceded 25 bases upstream by a polyadenylation signal (AATAAA); λPhos 1 is represented by the sequence starting from nucleotide 218 in Figure 1. The protein encoded by the open reading frame starting from the proposed initiator ATG is 309 amino acids in legnth and has a molecular mass of 35.6 kDa. This molecular mass is consistent with the values observed for the protein isolated by several groups (Ballou & Fischer, 1986). A peptide corresponding to the N-terminal of the deduced amino acid sequence was isolated from a CNBr digest of the rabbit enzyme. The native protein did not, however, yield any N-terminal sequence. A likely explanation for these observations is that the N-terminal amino acid of the protein is a methionine that has been posttranslationally modified so as to render the protein resistant to Edman degradation. Treatment with CNBr would, however, efficiently remove the blocked methionine, and the remaining peptide would be susceptible to amino acid sequencing. Support for this hypothesis is provided by the occurrence of an Asp residue in the second position (Figure 1). When the N-terminal methionine is acetylated in eukaryotes, Asp is the predominant penultimate residue (Flinta et al., 1986).

The clone λPhos 3 contained an open reading frame of 880 nucleotides that was lacking an initiator codon as shown in Figure 2. The protein encoded by λPhos 3 was highly homologous to that encoded by λPhos 4 (Figure 3). We have termed the proteins encoded by λPhos 4 and λPhos 3 the α and β catalytic subunits, respectively. The deduced amino acid sequence of λPhos 3 was 293 residues long, and alignment with the sequence of the α catalytic subunit resulted in the first residue encoded by λPhos 3 being aligned with residue 17 of the α catalytic subunit. The degree of identity between the two proteins was 98% with the major part of the divergence occurring in the N-terminal region. Comparison of the nucleotide sequences revealed 82% identity in the coding region. In the 3′-noncoding region, however, no significant homology was observed which allowed oligonucleotide probes specific for the α and β catalytic subunits to be synthesized (see below).

Comparison of the partial amino acid sequence obtained for the rabbit enzyme with the predicted full amino acid sequence of the α catalytic subunit reveals about 95% identity between

```
        CGTCAGTGGCCGGTAGCCGAACACCGAGGGGCGAGTCAACCGGGCAGCGCTACTGCCGA   -120
-119    GCAGCCTTCAGCAGCCACACAAGTACCCCGGGGGCCGACGGCAGCGTGTGCGTGTGGCCC   -60
 -59    GTGTGCGGGCGGCGGCGCGGGAGTAGCGCGGAGCGGCAGCCGGTTCGGGCGGGTGGCATC   0
          .         .         .         .         .         .
   1    ATGGACGAGAAGGTGTTCACTAAGGAGCTGGACCAGTGGATCGAGCAGCTGAATGAGTGC   60
        MetAspGluLysValPheThrLysGluLeuAspGlnTrpIleGluGlnLeuAsnGluCys
          .         .         .         .         .         .
  61    AAGCAGCTGTCCGAGTCCCAGGTCAAGAGCCTCTGCGAGAAGGCTAAAGAAATCCTGACA   120
        LysGlnLeuSerGluSerGlnValLysSerLeuCysGluLysAlaLysGluIleLeuThr
          .         .         .         .         .         .
 121    AAAGAATCCAACGTGCAAGAGGTTCGATGTCCAGTCACTGTCTGTGGAGATGTGCACGGG   180
        LysGluSerAsnValGlnGluValArgCysProValThrValCysGlyAspValHisGly
          .         .         .         .         .         .
 181    CAATTTCATGATCTCATGGAACTGTTTAGAATTGGTGGCAAATCACCAGATACGAATTAC   240
        GlnPheHisAspLeuMetGluLeuPheArgIleGlyGlyLysSerProAspThrAsnTyr
          .         .         .         .         .         .
 241    CTGTTTATGGGAGATTACGTTGACAGAGGATATTATTCCGTTGAAACAGTTACTCTGCTT   300
        LeuPheMetGlyAspTyrValAspArgGlyTyrTyrSerValGluThrValThrLeuLeu
          .         .         .         .         .         .
 301    GTAGCTCTTAAGGTTCGTTACCGTGAACGCATCACCATTCTTCGAGGAAATCATGAGAGC   360
        ValAlaLeuLysValArgTyrArgGluArgIleThrIleLeuArgGlyAsnHisGluSer
          .         .         .         .         .         .
 361    AGACAGATCACACAAGTATATGGTTTCTATGATGAGTGTTTAAGGAAATATGGAAATGCA   420
        ArgGlnIleThrGlnValTyrGlyPheTyrAspGluCysLeuArgLysTyrGlyAsnAla
          .         .         .         .         .         .
 421    AATGTTTGGAAATATTTTACAGATCTTTTTGATTATCTTCCTCTCACTGCATTAGTGGAT   480
        AsnValTrpLysTyrPheThrAspLeuPheAspTyrLeuProLeuThrAlaLeuValAsp
          .         .         .         .         .         .
 481    GGGCAGATCTTCTGTCTACATGGTGGCCTCTCACCATCCATAGATACACTGGATCACATC   540
        GlyGlnIlePheCysLeuHisGlyGlyLeuSerProSerIleAspThrLeuAspHisIle
          .         .         .         .         .         .
 541    AGAGCACTTGATCGCCTGCAAGAAGTTCCCCATGAGGGTCCAATGTGTGACTTGCTGTGG   600
        ArgAlaLeuAspArgLeuGlnGluValProHisGluGlyProMetCysAspLeuLeuTrp
          .         .         .         .         .         .
 601    TCAGATCCAGATGACCGTGGAGGTTGGGGTATATCTCCTCGAGGAGCTGGTTACACTTTT   660
        SerAspProAspAspArgGlyGlyTrpGlyIleSerProArgGlyAlaGlyTyrThrPhe
          .         .         .         .         .         .
 661    GGGCAAGATATTTCCGAGACATTTAATCATGCCAATGGCCTCACGTTGGTGTCCAGAGCT   720
        GlyGlnAspIleSerGluThrPheAsnHisAlaAsnGlyLeuThrLeuValSerArgAla
          .         .         .         .         .         .
 721    CATCAGCTCGTGATGGAGGGATATAACTGGTGCCATGACCGGAATGTAGTAACGATTTTC   780
        HisGlnLeuValMetGluGlyTyrAsnTrpCysHisAspArgAsnValValThrIlePhe
          .         .         .         .         .         .
 781    AGTGCTCCAAACTATTGTTACCGTTGTGGGAACCAAGCTGCAATCATGGAACTTGATGAT   840
        SerAlaProAsnTyrCysTyrArgCysGlyAsnGlnAlaAlaIleMetGluLeuAspAsp
          .         .         .         .         .         .
 841    ACTCTAAAATACTCTTTCTTGCAGTTTGACCCAGCACCTCGCAGAGGCGAGCCACACGTT   900
        ThrLeuLysTyrSerPheLeuGlnPheAspProAlaProArgArgGlyGluProHisVal
          .         .         .         .         .         .
 901    ACTCGTCGTACCCCAGACTACTTCCTGTAATGAAATTTTAAACTTGTACAGTATTGCCAT   960
        ThrArgArgThrProAspTyrPheLeu
 961    GAACCATATATTGACCTAATGGAGATGGGAAGAGCAACAGTAACTCCACAAAGTGTCAGA   1020
1021    GAATAGTTAACATTCAAAAAAACTTGTTTTCACACGGACCAAAAAGATGTGCCATATAAA   1080
1081    ATACAAAGCCTCTTGTCATCAACAGCCGTGACCACTTTAGAATGAACCAGTCCATTGCAT   1140
1141    GCTGAAGCGACATTGTTGGTCAAGAAACCAGTTTCTGGCATAGCGCTACTTGTAGTTACT   1200
1201    TTGCTTTCTCTGAGAGACTGCAGATAATAAGATGTAAACATTAACACCTCGTGAATACAA   1260
1261    TTTAACTTCCATTTAGCTATAGCTTTACTCAGCATGACTGTAGATAAGAATAGCAGCAAA   1320
1321    CAATCATTGGAGCTTAATGAACATTTTTAAAAATAAGTACCAAGGCCTCCCCTCTACTTG   1380
1381    TGAGTTTCGAAATTGTTTTGTTTATTTTCAGGGATACCGTTTAATTTAATTGTATGATTT   1440
1441    GTCTGCACTCAGTTTATTTCCTTTCTCAAATCTCAGCCTCATGTTGTTCTTTGTTATTGT   1500
1501    CAGAACCTGGTGAGTTGTTTTGAACAGAACTGTTTTTTTTCTTCCTCCTGTAAGACGATGT   1560
1501    CAGAACCTGGTGAGTTGTTTTGAACAGAACTGTTTTTTTTCTTCCTCCTGTAAGACGATGT   1560
1561    TACTGCACAAGGAGCACTGACGTGTTTTTCATAATAAACTTGTGAACTAAGAACTGAAAA   1620
1621    AAAAAAAAAAAAAA
```

FIGURE 1: Nucleotide and deduced amino acid sequence of the protein phosphatase 2A α catalytic subunit cDNA. The nucleotides are numbered in the 5′ to 3′ direction, starting with the first nucleotide of the proposed initiator codon. Nucleotides extending 5′ of base 1 are designated with negative numbers. The deduced amino acid sequence for the phosphatase 2A α catalytic subunit is shown below the nucleotide sequence. The nucleotide sequence represents a combination of those of the two overlapping clones λPhos 4 and λPhos 1, which correspond to nucleotides −179 to 945 and 218–1634, respectively. An oligonucleotide probe complementary to the underlined sequence was used in Northern and Southern analyses to differentiate between sequences encoding the α and β forms of the catalytic subunit.

the two (Figure 3). Preliminary results obtained from cDNAs encoding the human phosphatase 2A catalytic subunit indicate that the human enzymes also display a high degree of identity with the rabbit and porcine enzymes (S. R. Stone, J. Hofsteenge, and B. A. Hemmings, unpublished results).

Expression of Phosphatase 2A Catalytic Subunit mRNA in LLC-PK$_1$ Cells. The expression of mRNA coding for the catalytic subunits of protein phosphatase 2A was examined by using the 1140-bp *Eco*RI insert of λPhos 4 (nucleotides −179 to 945 of Figure 1) and specific oligonucleotide probes for the α and β subunits. The λPhos 4 fragment hybridized to mRNA species of about 2 kb (data not shown). The specific probes were synthesized so that they would hybridize to the 3′-noncoding region of the mRNA of each form as indicated in Figures 2 and 3. Northern analysis using LLC-PK$_1$ poly(A)+ RNA revealed that the two oligonucleotide probes hybridized to mRNAs approximately 2 kb in length and that the abundance of both messages was about equal (data not shown).

Genomic Sequences Hybridizing to Phosphatase 2A Catalytic Subunit cDNA and Oligonucleotide Probes Specific for

```
   1  CCTGAACGAGTGTAAACAGCTTAATGAGAACCAAGTGCGGACCCTATGCGAAAAGGCTAA   60
       LeuAsnGluCysLysGlnLeuAsnGluAsnGlnValArgThrLeuCysGluLysAlaLy

  61  GGAAATTCTAACAAAAGAATCAAATGTGCAAGAGGTTCGTTGTCCAGTTACGGTCTGTGG   120
      sGluIleLeuThrLysGluSerAsnValGlnGluValArgCysProValThrValCysGl

 121  AGATGTGCATGGTCAATTCCATGATCTTATGGAACTCTTTAGAATCGGTGGAAAATCACC   180
      yAspValHisGlyGlnPheHisAspLeuMetGluLeuPheArgIleGlyGlyLysSerPr

 181  AGACACAAACTATCTATTCATGGGTGACTATGTAGACAGAGGTTATTATTCGGTGGAGAC   240
      oAspThrAsnTyrLeuPheMetGlyAspTyrValAspArgGlyTyrTyrSerValGluTh

 241  TGTGACTCTTCTTGTGGCATTAAAGGTGCGGTATCCAGAACGCATTACAATATTGAGAGG   300
      rValThrLeuLeuValAlaLeuLysValArgTyrProGluArgIleThrIleLeuArgGl

 301  AAACCACGAGAGCCGACAAATTACCCAAGTATATGGCTTTTATGACGAATGTCTGCGGAA   360
      yAsnHisGluSerArgGlnIleThrGlnValTyrGlyPheTyrAspGluCysLeuArgLy

 361  GTATGGAAATGCCAATGTTTGGAAATATTTTACAGATCTATTTGATTATCTTCCACTTAC   420
      sTyrGlyAsnAlaAsnValTrpLysTyrPheThrAspLeuPheAspTyrLeuProLeuTh

 421  AGCTTTAGTAGATGGACAGATATTCTGCCTCCATGGTGGCCTCTCTCCATCCATAGATAC   480
      rAlaLeuValAspGlyGlnIlePheCysLeuHisGlyGlyLeuSerProSerIleAspTh

 481  ACTGGATCATATAAGAGCCCTGGATCGTTTACAAGAAGTGCCACATGAGGGCCCAATGTG   540
      rLeuAspHisIleArgAlaLeuAspArgLeuGlnGluValProHisGluGlyProMetCy

 541  TGATCTCTTATGGTCGGATCCGGATGATCGTGGTGGGTGGGGCATCTCACCACGTGGTGC   600
      sAspLeuLeuTrpSerAspProAspAspArgGlyGlyTrpGlyIleSerProArgGlyAl

 601  TGGCTACACATTTGGACAAGATATTTCTGAAACATTTAACCATGCCAATGGCCTCACACT   660
      aGlyTyrThrPheGlyGlnAspIleSerGluThrPheAsnHisAlaAsnGlyLeuThrLe

 661  GGTTTCTCGTGCTCACCAACTGGTAATGGAGGGATACAACTGCTGTCATGACCGGAACGT   720
      uValSerArgAlaHisGlnLeuValMetGluGlyTyrAsnCysCysHisAspArgAsnVa

 721  TGTTACCATTTTCAGTGCACCCAATTACTGTTATCGTTGTGGGAACCAGGCTGCTATCAT   780
      lValThrIlePheSerAlaProAsnTyrCysTyrArgCysGlyAsnGlnAlaAlaIleMe

 781  GGAATTAGATGACACTTTGAAATATTCCTTCCTTCAATTTGACCCAGCACCTCGTCGTGG   840
      tGluLeuAspAspThrLeuLysTyrSerPheLeuGlnPheAspProAlaProArgArgGl

 841  AGAGCCTCATGTGACCCGGCGCACCCCAGACTACTTCCTGTAAATTGTTCCTGGGAAAAC   900
      yGluProHisValThrArgArgThrProAspTyrPheLeu

 901  CTACCTTTATATGTGGAAGTATACCTGGCTTTTTAAAATATATATTTCAAAACAAAAC   960
 961  AAAAAGCAACAGTAATCTATGTGTTTCTGTAGCACATTGGGATCTGTCTTGGCATTAAAC   1020
1021  CATATCATGGACCAAAATGTGCCATACTAATCTAATGATGAGCATTTAGCACAATATGAG   1080
1081  ACTAAATTTTTAGTACAACACTATGTTCTAGATTCGTCGGTCTTAACAGTTTGCCTGCTG   1140
1141  TATATGTAGTAACCATTTTCTTGGGACTGCTCAAACAAAAAAGGTAACTAATTCCATCTC   1200
1201  CTTTTGCACTTACTTGGAAATTTAGTTAGAG
```

FIGURE 2: Nucleotide and deduced amino acid sequence of the protein phosphatase 2A β catalytic subunit cDNA. The nucleotides are numbered in the 5′ to 3′ direction, and the deduced amino acid sequence for the phosphatase 2A β catalytic subunit is shown below the nucleotide sequence. An oligonucleotide probe complementary to the underlined sequence was used in Northern and Southern analyses for specific detection of sequences encoding the β form of the catalytic subunit.

the α and β Forms. Southern analysis of LLC-PK$_1$ DNA was performed to investigate the genetic complexity of the genes coding for protein phosphatases in the porcine genome. Figure 4A shows the results obtained when the λPhos 4 *Eco*RI fragment was used as the probe. Initially, the filters were washed at low stringency (60 °C, 2 × SSC) and autoradiographed. The pattern of hybridization obtained (Figure 4) suggested that there were several genes coding for protein phosphatase catalytic subunits or related proteins. Following high stringency washing (65 °C, 0.5 × SSC), the pattern of hybridization was less complex (Figure 4A), which suggested the λPhos 4 probe hybridized most strongly to at least two related genes, consistent with the cDNA sequence data.

The oligonucleotides specific for the 3′-noncoding regions of λPhos 3 and λPhos 4 yielded different hybridization patterns in Southern analysis as shown in Figure 4B, which confirmed that the α and β isoforms were derived from two distinct genes.

Discussion

The data presented in this paper indicate that there are two isoforms of the catalytic subunit of protein phosphatase 2A. Although these two forms are highly homologous, they are the products of separate genes as indicated by their nucleotide sequence and by Southern analysis. The existence of two isoforms of the phosphatase 2A catalytic subunit suggests parallels with the catalytic subunit of cAMP-dependent protein kinase (Uhler et al., 1986a,b; Showers & Maurer, 1986; Rao Adavani et al., 1987) and with protein kinase C (Coussens et al., 1986; Knopf et al., 1986; Ohno et al., 1987). The complex pattern of Southern analysis presented in Figure 4A could indicate the existence of other related genes in addition to the genes coding for the α and β catalytic subunit of phosphatase 2A. In this respect, it is interesting to note that all of the protein kinases so far sequenced contain homologous regions (Hunter & Cooper, 1986), and therefore, it seems possible that the catalytic subunits of other protein phosphatases contain regions homologous to that of phosphatase 2A. Homology with the catalytic subunit of phosphatase 1 seems a distinct possibility. This enzyme is about the same size as the phosphatase 2A catalytic subunit and has a similar amino acid composition (Silberman et al., 1984). In addition, monoclonal antibodies have been isolated that react with both proteins (Speth et al., 1984; Brautigan et al., 1986). Confirmation of this idea of a protein phosphatase multigene family must await

```
  1 MDEKVFTKELDQWIEQLNECKQLSESQVKSLCEKAKEILTKESNVQEVRCPVTVCGDVHG 1
                    LNECKQLNENQVRTLCEKAKEILTKESNVQEVRCPVTVCGDVHG 2
    mDEKVFTKELDQWIEQLNE     eSQVKSLCEKAKE          eVRCPVTVCGDVhG 3

 61 QFHDLMELFRIGGKSPDTNYLFMGDYVDRGYYSVETVTLLVALKVRYRERITILRGNHES 1
    QFHDLMELFRIGGKSPDTNYLFMGDYVDRGYYSVETVTLLVALKVRYREPITILRGNHES 2
    QFXDLmELFRIGGKSPDTNYLFMGDYV                                  3

121 RQITQVYGFYDECLRKYGNANVWKYFTDLFDYLPLTALVDGQIFCLHGGLSPSIDTLDHI 1
    RQITQVYGFYDECLRKYGNANVWKYFTDLFDYLPLTALVDGQIFCLHGGLSPSIDTLDHI 2
                    YGNANVWK                                     3

181 RALDRLQEVPHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANGLTLVSRA 1
    RALDRLQEVPHEGPMCDLLWSDPDDRGGWGISPRGAGYTFGQDISETFNHANGLTLVSRA 2
         LQEVPXEGPMcdll                GAGYTFGQDIKETFNhAKGKt     3

241 HQLVMEGYNWCHDRNVVTIFSAPNYCYRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHV 1
    HQLVMEGYNCCHDRNVVTIFSAPNYCYRCGNQAAIMELDDTLKYSFLQFDPAPRRGEPHV 2
        mEGYNVChDRNVVTIFSAPNYXYrXGNQAAIMELDDTLKYSFLQFDPAprnge    3

301 TRRTPDYFL                                                    1
    TRRTPDYFL                                                    2
```

FIGURE 3: Comparison of the predicted amino acid sequences of the α and β forms of the porcine phosphatase 2A catalytic subunit with the partial amino acid sequence determined for the rabbit skeletal muscle enzyme. The sequences are given in the single-letter amino acid code, and residues that differ between sequences are underlined. The sequences of the porcine α and β catalytic subunits are represented by sequences 1 and 2, respectively, while the rabbit sequence is numbered 3. Lower case letters represent amino acid residues inferred from the method of cleavage used or tentative assignments due to low levels obtained in the Edman degradation.

the sequencing of further phosphatases.

Protein phosphatase 2A has been shown to exist in a number of different forms, which consist of a catalytic subunit complexed with various regulatory subunits (Ballou & Fischer, 1986). Peptide mapping of the catalytic subunit of the various forms of phosphatase 2A from rabbit skeletal muscle suggested that the same catalytic subunit was found in each form (Tung et al., 1985). It could be expected, however, that the differences between the α and β forms of the catalytic subunit presented in this paper would not be detected by the method of peptide mapping used (Tung et al., 1985). Therefore, it would be of interest to determine whether the different forms of the phospatase 2A contain different catalytic subunits. Also of interest is whether the two forms show a tissue-specific distribution. Similar studies with cAMP-dependent protein kinase catalytic subunit (Uhler et al., 1986b) and protein kinase C (Ohno et al., 1987; Brandt et al., 1987) have demonstrated that the different forms of these two enzymes show a tissue-specific distribution.

Comparison of the deduced amino acid sequences of the porcine phosphatases with that of rabbit enzyme indicates that the protein is conserved between these two species. Similar observations have been made with the protein kinases (Coussens et al., 1986; Showers & Maurer, 1986; Rao Adavani et al., 1987), which is consistent with a central role in the control of metabolism for these two classes of enzymes. The catalytic subunits of phosphatase 2A seem, however, to represent the first members of a unique family of proteins. Searches that compared the nucleotide and predicted amino acid sequences of both λPhos 4 and λPhos 3 with all sequences in the NBRF and EMBL data banks (Lipman & Pearson, 1985) revealed no significant extended homologies with any other known sequences. Since protein kinases and phosphatases share the same protein substrates, some homology may be expected in the substrate binding sites of the catalytic subunits of both classes of proteins, and consequently, a more detailed search was made for homology between the phosphatases, and phosphorylase kinase (Reimann et al., 1984), protein kinase C (Parker et al., 1986) and cAMP-dependent protein kinase (Shoji et al., 1983). For this purpose the algorithm of Smith and Waterman (1981) was used, but again no significant homology was found. Similar comparisons with

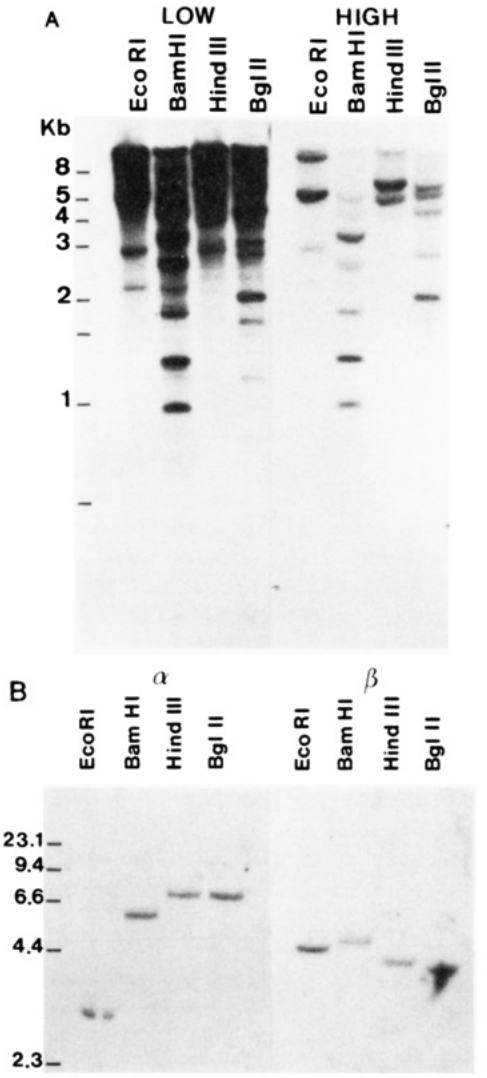

FIGURE 4: (A) Southern analysis of porcine genomic DNA using λPhos 4 cDNA. DNA from LLC-PK_1 cells were digested with the indicated restriction endonucleases, fractionated on 1% agarose gels, transferred to nitrocellulose, and hybridized with the 1140-bp *Eco*RI fragment of λPhos 4 as described under Experimental Procedures. Following hybridization, the filter was washed at "low" stringency (60 °C, 2 × SSC for 2 h) and autoradiographed. The filter was then subjected to a "high" stringency wash (65 °C, 0.5 × SSC for 2 h) and autoradiographed. In both cases, autoradiography was performed for 3 days at −70 °C with intensifying screens. The numbers mark the positions at which standards of that size in kilobases migrated. (B) Southern analysis of porcine genomic DNA using oligonucleotide probes specific for the α and β isoforms of phosphatase 2A. Southern analysis was performed essentially as described above except that, following hybridization with the specific oligonucleotides (see Figures 1 and 2), the filters were washed at 55 °C with 1 × SSC/0.1% SDS (see Experimental Procedures). Autoradiography was performed for 14 days at −70 °C with intensifying screens. The numbers mark the positions at which standards of that size in kilobases migrated.

other types of phosphatases (Berger et al., 1987; Shull et al., 1985; Brandl et al., 1986; Marcus et al., 1982) did not allow the identification with any certainty of putative active-site residues. It seems that the identification of regions of protein phosphatases involved in substrate binding and catalysis requires further protein chemical studies.

During the preparation of this paper, the cDNA sequence coding for a catalytic subunit of bovine protein phosphatase 2A was reported (Green et al., 1987). The deduced amino acid sequence of the bovine enzyme shows considerable identity with those presented in Figure 4. One major difference is that the sequence of the bovine enzyme is 325 residues in length

compared with 309 residues for the porcine α catalytic subunit (Figure 4). The difference between the two sequences begins after Pro-293 and is caused by the presence of an additional T in the porcine cDNA sequence of Figure 1 after the second nucleotide of the codon for Pro-293. We have observed a T in this position in all of six porcine and human clones that have been sequenced. Moreover, the tentative sequence of the rabbit enzyme obtained by amino acid sequencing corresponds to the deduced porcine sequence (Figure 4). An additional discrepancy occurs between Glu-33 and Leu-40 of the porcine sequence (Figure 4). Again the peptide sequence of the rabbit enzyme supports the porcine sequence. The bovine sequence of Green et al. (1987) corresponds to those of Figure 4 if a G is inserted into their nucleotide sequence at the beginning of the region of nonidentity and a G is deleted at the end.

Added in Proof

Further analysis of the bovine cDNA clone (Green et al., 1987) has indicated that the deduced amino acid sequence of this clone is identical with that of the porcine α catalytic subunit in all but one position (M. Mumby, personal communication).

Acknowledgments

We thank R. Meyer, R. Matthies, and M. Pedrocca for technical assistance, Dr. D. Jans for advice on DNA sequencing, W. Zuercher and Dr. J. Jiricny for the synthesis of oligonucleotides, and M. Altus, Dr. W. Filipowicz, Dr. S. Shibahara, and Dr. Y. Khew-Goodall for their reading of the manuscript.

Registry No. DNA (pig protein phosphatase 2A α catalytic subunit), 110717-77-6; protein phosphatase 2A (pig α catalytic subunit reduced), 110717-81-2; protein phosphatase, 9025-75-6.

References

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408–1412.

Ballou, L. M., & Fischer, E. H. (1986) *Enzymes (3rd Ed.)*, 311–361.

Berger, J., Garattini, E., Hua, J.-C., & Udenfriend, S. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 695–698.

Blin, N., & Stafford, D. W. (1976) *Nucleic Acids Res. 3*, 2303–2306.

Brandl, C. J., Green, N. M., Korczak, B., & MacLennan, D. H. (1986) *Cell (Cambridge, Mass.) 44*, 597–607.

Brandt, S. J., Niedel, J. E., Bell, R. M., & Young, W. S., III (1987) *Cell (Cambridge, Mass.) 49*, 57–63.

Brautigan, D. L., Gruppuso, P. A., & Mumby, M. (1986) *J. Biol. Chem. 261*, 14924–14928.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Cohen, P. (1985) *Eur. J. Biochem. 151*, 439–448.

Coussens, L., Parker, P. J., Rhee, L., Yang-Feng, T. L., Chen, E., Waterfield, M. D., Francke, U., & Ullrich, A. (1986) *Science (Washington, D.C.) 233*, 859–866.

Feinberg, A. P., & Vogelstein, B. (1984) *Anal. Biochem. 137*, 266–267.

Flinta, C., Persson, B., Joernvall, H., & von Heije, G. (1986) *Eur. J. Biochem. 154*, 193–196.

Green, D. D., Yang, S., & Mumby, M. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4880–4884.

Gubler, U., & Hoffman, B. J. (1983) *Gene 25*, 263–269.

Hemmings, B. A., Schwarz, M., Rao Adavani, S., & Jans, D. A. (1986) *FEBS Lett. 209*, 219–222.

Hewick, R. M., Hunkapillar, M. W., Hood, L. W., & Dreger, W. J. (1981) *J. Biol. Chem. 256*, 7990–7997.

Hofsteenge, J., Weijer, W. J., Jekel, P. A., & Beintema, J. J. (1983) *Eur. J. Biochem. 133*, 91–108.

Hull, R. N., Cherry, W. R., & Weaver, G. W. (1976) *In Vitro 12*, 670–677.

Hunter, T., & Cooper, J. A. (1986) *Enzymes (3rd Ed.)*, 191–310.

Ingebritsen, T. S., & Cohen, P. (1983) *Eur. J. Biochem. 132*, 255–261.

Knopf, J. L., Lee, M.-H., Sultzman, L. A., Kriz, R. W., Loomis, C. R., Hewick, R. M., & Bell, R. M. (1986) *Cell (Cambridge, Mass.) 46*, 491–502.

Kozak, M. (1984) *Nucleic Acids Res. 12*, 857–872.

Krebs, E. G. (1986) *Enzymes (3rd Ed.)*, 3–20.

Lathe, R. (1985) *J. Mol. Biol. 183*, 1–12.

Lipman, D. J., & Pearson, W. R. (1985) *Science (Washington, D.C.) 227*, 1435–1441.

Marcus, F., Edelstein, I., Reardon, I., & Heinrikson, R. L. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 7161–7165.

Mizusawa, S., Mishimura, S., & Seela, F. (1986) *Nucleic Acids Res. 14*, 1319–1324.

Ohno, S., Kawasaki, H., Imajoh, S., Suzuki, K., Inagaki, M., Yokokura, H., Sakoh, T., & Hidaka, H. (1987) *Nature (London) 325*, 161–166.

Parker, P. J., Coussens, L., Totty, N., Rhee, L., Young, S., Chen, E., Stabel, S., Waterfield, M. D., & Ullrich, A. (1986) *Science (Washington, D.C.) 233*, 853–859.

Rao Adavani, S., Schwarz, M., Showers, M. O., Maurer, R. A., & Hemmings, B. A. (1987) *Eur. J. Biochem. 167*, 221–226.

Reimann, E. M., Titani, K., Ericsson, L. H., Wade, R. D., Fischer, E. H., & Walsh, K. A. (1984) *Biochemistry 23*, 4185–4192.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Shoji, S., Ericsson, L. H., Walsh, K. A., Fischer, E. H., & Titani, K. (1983) *Biochemistry 22*, 3702–3709.

Showers, M. O., & Maurer, R. A. (1986) *J. Biol. Chem. 261*, 16288–16291.

Shull, G. E., Schwartz, A., & Lingrel, J. B. (1985) *Nature (London) 316*, 691–695.

Silberman, S. R., Speth, M., Nemani, R., Ganapathi, M. K., Dombradi, V., Paris, H., & Lee, E. Y. C. (1984) *J. Biol. Chem. 259*, 2913–2922.

Smith, T. F., & Waterman, M. S. (1981) *Adv. Appl. Math. 2*, 482–489.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.

Speth, M., Alejandro, R., & Lee, E. Y. C. (1984) *J. Biol. Chem. 259*, 3475–3481.

Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5201–5205.

Tung, H. Y. L., Resink, T. J., Hemmings, B. A., & Cohen, P. (1984) *Eur. J. Biochem. 138*, 635–641.

Tung, H. Y. L., Alemany, S., & Cohen, P. (1985) *Eur. J. Biochem. 148*, 253–264.

Uhler, M. D., Carmichael, D. F., Lee, D. C., Chrivia, J. C., Krebs, E. G., & McKnight, G. S. (1986a) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1300–1304.

Uhler, M. D., Chrivia, J. C., & McKnight, G. S. (1986b) *J. Biol. Chem. 261*, 15360–15363.