Molecular Cloning and Expression of the Catalytic Subunit of Bovine Pyruvate Dehydrogenase Phosphatase and Sequence Similarity with Protein Phosphatase 2C†,‡

Janet E. Lawson,§ Xiao-Da Niu,§,∥ Karen S. Browning,§ Hai Le Trong,⊥,# Jiangong Yan,§ and Lester J. Reed*,§

Biochemical Institute and Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712, and Department of Biochemistry, University of Washington, Seattle, Washington 98195

Received June 1, 1993; Revised Manuscript Received July 16, 1993•

ABSTRACT: After many unsuccessful attempts to detect cDNA encoding the catalytic subunit of bovine pyruvate dehydrogenase phosphatase (PDPc) in bovine cDNA libraries, an approach based on the polymerase chain reaction (PCR) was undertaken. Overlapping DNA fragments were generated by PCR from bovine genomic DNA and from cDNA synthesized from total RNA with synthetic oligonucleotide primers on the basis of experimentally determined amino acid sequences. The DNA fragments were subcloned and sequenced. The complete cDNA is 1900 base pairs in length and contains an open reading frame of 1614 nucleotides encoding a putative presequence of 71 amino acid residues and a mature protein of 467 residues with a calculated M_r of 52 625. Hybridization analysis showed a single mRNA transcript of about 2.0 kilobases. Comparison of the deduced amino acid sequences of the mitochondrial PDPc and the rat cytosolic protein phosphatase 2C indicates that these protein serine/threonine phosphatases evolved from a common ancestor. The mature form of PDPc was coexpressed in *Escherichia coli* with the chaperonin proteins groEL and groES. The recombinant protein (rPDPc) was purified to near homogeneity. Its activity toward the bovine ^{32}P-labeled pyruvate dehydrogenase complex was Mg^{2+}-dependent and Ca^{2+}-stimulated and comparable to that of native bovine PDP. An active, truncated form of rPDPc, with M_r ~45 000, was produced in variable amounts during growth of cells and/or during the purification procedure.

Four major classes of protein serine/threonine phosphatases (PP1, PP2A, PP2B, and PP2C)[1] have been identified in eukaryotic cells on the basis of substrate specificities and sensitivity to activators and inhibitors (Cohen, 1989). PP1 is sensitive to the thermostable proteins inhibitor 1 and inhibitor 2, and PP1 and PP2A are sensitive to okadaic acid (Bialojan & Takai, 1988; Haystead et al., 1989). PP2B is Ca^{2+}/calmodulin-dependent, and PP2C exhibits an absolute requirement for Mg^{2+}. Unlike PP1, PP2A, and PP2B, which contain catalytic and regulatory subunits, PP2C has been isolated from the cytosol of mammalian tissues as a free catalytic subunit. Two isoforms of PP2C have been purified from rabbit skeletal muscle and rabbit liver, PP2Cα (M_r ~44 000) and PP2Cβ (M_r ~42 000) (Hiraga et al., 1981; McGowan & Cohen, 1987). PP2C apparently does not belong to the same protein phosphatase gene family as PP1, PP2A, and PP2B (Tamura et al., 1989; Shenolikar & Nairn, 1991; Wenk et al., 1992).

Pyruvate dehydrogenase phosphatase (PDP) is a Mg^{2+}-dependent and Ca^{2+}-stimulated protein serine phosphatase which catalyzes the dephosphorylation and concomitant reactivation of the mitochondrial pyruvate dehydrogenase multienzyme complex (Linn et al., 1969; Denton et al., 1972). Bovine PDP is a heterodimer with M_r ~150 000 (Teague et al., 1982; Pratt et al., 1982). It consists of a catalytic subunit (PDPc) with an apparent M_r of ~50 000 and a subunit with an apparent M_r of ~97 000. The latter subunit is a flavoprotein (FAD) of as yet unknown function. In the presence of Ca^{2+}, PDP binds to the dihydrolipoamide acetyltransferase (E_2) component of the pyruvate dehydrogenase complex, and its apparent K_m for the phosphorylated pyruvate dehydrogenase (E_1) component is decreased about 20-fold (Pettit et al., 1972). PDP possesses an "intrinsic" Ca^{2+}-binding site (K_d ~8 μM), and a second Ca^{2+}-binding site (K_d ~5 μM) is generated when both PDP and E_2 are present (Teague et al., 1982). PDP activity is not inhibited by protein phosphatase inhibitor 1, inhibitor 2, or okadaic acid (Damuni et al., 1985; Haystead et al., 1989).

Similarities in some properties of the mitochondrial PDPc and the cytosolic PP2C suggested that these two phosphatases are related. These properties include an absolute requirement for Mg^{2+}, insensitivity to inhibitor 1, inhibitor 2, and okadaic acid, and similarities in substrate specificity (P. Cohen, Z. Damuni, H. Y. L. Tung, and L. J. Reed, unpublished data).

To gain further insight into the structure, function, and regulation of PDP and its relationship to PP2C, we undertook the cloning of cDNAs coding for the two PDP subunits. This paper reports the isolation and characterization of cDNA clones encoding PDPc and expression of the coding sequence in *Escherichia coli*. Comparison of the deduced amino acid sequences of bovine mitochondrial PDPc and rat cytosolic

† This work was supported by USPHS Grant GM06590 and by a grant from the Foundation for Research.

‡ The nucleotide sequence reported in this paper has been submitted to GenBank under Accession Number L18966.

• Address correspondence to this author.

§ University of Texas at Austin.

∥ Present address: Schering-Plough Research Institute, Kenilworth, NJ 07033.

⊥ University of Washington.

Present address: ICOS Corp., Bothell, WA 98021.

• Abstract published in *Advance ACS Abstracts*, August 15, 1993.

[1] Abbreviations: PDP, pyruvate dehydrogenase phosphatase; PDPc, catalytic subunit of PDP; PP2C, protein serine/threonine phosphatase 2C; PCR, polymerase chain reaction; bp, base pair(s); kb, kilobase(s); SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; r, recombinant; IPTG, isopropyl β-thiogalactoside.

0006-2960/93/0432-8987$04.00/0

PP2C indicates that these two types of protein serine/threonine phosphatases evolved from a common ancestor.

EXPERIMENTAL PROCEDURES

Materials. Restriction endonucleases and DNA-modifying enzymes were purchased from New England Biolabs and Bethesda Research Laboratories. Random DNA primers were obtained from Boehringer Mannheim. The Gene Amp kit was obtained from Perkin-Elmer/Cetus, the TA cloning kit was from Invitrogen, the 5′ and 3′ RACE systems were from GIBCO BRL, and Zeta-Probe GT nylon membrane was from Bio-Rad. *E. coli* strain JM101 (Yanisch-Perron et al., 1983) or strain TB1 (New England Biolabs) was used for plasmid propagation, and strain JM101 was used for overexpression of proteins. Bovine kidney genomic DNA and bovine total RNA were purchased from Clontech. Plasmid pKK223-3 was purchased from Pharmacia, Bluescript was from Stratagene, and plasmid pGroESL was provided by Dr. Anthony Gatenby (Du Pont Experimental Station, Wilmington, DE). Rabbit antibodies to PDPc (Niu, 1989) were purified on protein A–agarose (Sigma) to obtain the IgG fraction.

Protein Sequence Analysis. Highly purified PDP from bovine heart or kidney (Teague et al., 1982; Niu, 1989) was subjected to SDS–PAGE (Laemmli, 1970), and the resolved protein bands were electroblotted onto an Immobilon-P membrane (Matsudaira, 1987). The band corresponding to the catalytic subunit was excised and subjected to automated sequence analysis with an Applied Biosystems Model 470A gas-phase sequencer equipped with an on-line Model 120A phenylthiohydantoin amino acid analyzer.

To obtain a sample of PDPc for internal amino acid sequence analysis, the corresponding band was eluted electrophoretically from the slab gel, and the solution was dialyzed against 5 mM NH_4HCO_3/0.05% SDS and dried in a Speed-Vac concentrator. The residue was dissolved in 50 μL of water, and the protein was precipitated with 0.45 mL of acetone/1 mM HCl for 3 h at −20 °C. Approximately 16 μg (0.32 pmol) of PDPc was subjected to digestion (24 h, 37 °C) with 0.3 μg of trypsin in 2 M urea/50 mM NH_4HCO_3, followed by reduction and S-pyridylethylation (Le Trong et al., 1990). The mixture was separated by reverse-phase HPLC on an Aquapore RP-300 C8 microbore column (2.1 × 100 mm, Pierce) with a linear gradient of 0.1% trifluoroacetic acid/acetonitrile. Peptides were sequenced as described above. In another approach, a solution of 0.9 mg of PDP (heterodimer) in 70% formic acid was incubated with CNBr in the dark at room temperature for 24 h. The digest was dried, the residue was dissolved in 0.1% trifluoroacetic acid, the peptides were separated by reverse-phase HPLC, and several were sequenced.

Preparation of Oligonucleotides. Primers for PCR and DNA sequencing were synthesized on an Applied Biosystems Model 381A DNA synthesizer.

PCR. PCR consisted of 30 cycles (30 min at 94 °C followed by 29 cycles of 1 min at 94 °C, 2 min at 55 °C, 48 °C, or 42 °C, and 3 min at 72 °C; extended to 10 min at 72 °C in the last cycle).

DNA Sequencing. Miniplasmid preparations for sequencing were obtained by a modification of the procedure of Zhou et al. (1990). The supernatant was extracted with phenol/chloroform/isoamyl alcohol (25:24:1) followed by a chloroform extraction of the aqueous layer. Double-stranded DNA was sequenced with the Sequenase V2.0 kit (U.S. Biochemicals). Both strands of each fragment were sequenced completely. Multiple clones were sequenced to minimize PCR artifacts.

Sequence Comparisons. Beckman Microgenie programs, versions 5 and 6, were used to analyze the DNA sequence data. The deduced amino acid sequence of PDPc was searched against the GenBank and PIR data bases, utilizing the BLAST program (Altschul et al., 1990). Sequence alignments were optimized using MACAW (Schuler et al., 1991) and the program described by Feng and Doolittle (1987).

Amplification of cDNA Ends. The 5′ and 3′ RACE systems were used according to the manufacturer's instructions with a few modifications. In the synthesis of fragments K1 and K2, the heat-stable reverse transcriptase *rTth* was used at 70 °C to synthesize the cDNA. A 2-μL aliquot of the first-round PCR was used as a template for the second round of PCR. The reverse transcriptase supplied with the kit was used in the synthesis of all other fragments. In the synthesis of fragment a-6, the tailing reaction time was increased to 20 min at 37 °C. Five micrograms of bovine heart total RNA was used as a template to synthesize the cDNA. Five micrograms of the tailed reaction product was used in the first round of PCR. Agarose gel slices from the first-round PCR products were used as templates for the second-round PCR. In the 3′ RACE, 5 μL of the first-round PCR was used as a template for the second-round PCR. All PCRs were done with an annealing temperature of 55 °C.

Southern Blotting. Aliquots (4 μg each) of bovine kidney genomic DNA were digested with various restriction endonucleases, and the fragments were separated by electrophoresis on a 1.0% agarose gel. The DNA was transferred to Zeta-Probe GT membrane and probed as described (Lawson et al., 1991). The probe was the K1 fragment (nucleotides 709–1378) radiolabeled with [^{32}P]dCTP using random primers (Feinberg & Vogelstein, 1983).

RNA Blot Analysis. Aliquots (40 μg each) of bovine heart, lung, or brain total RNA were dried under vacuum and resuspended in water. The electrophoresis and transfer were done as described (Miran et al., 1993). The probe was the same as described above for Southern blotting.

Immunoblotting. Proteins were separated by SDS–PAGE on 12.5% gels and transferred electrophoretically to Immobilon-P membranes. Blots were probed with a 1:4000 dilution of anti-phosphatase IgG or preimmune IgG, followed by detection with a 1:4000 dilution of goat anti-rabbit IgG conjugated to alkaline phosphatase as described by the supplier (Bio-Rad).

Expression in E. coli. Primers K19 and K22, which hybridize to the 5′ and 3′ ends, respectively, of the coding region for the mature protein, were used with p2311 as a template to generate by PCR a fragment coding for the mature protein. K19 introduced an *Eco*RI site followed by an ATG start codon, and K22 introduced a *Hin*dIII site. These sites were used to subclone the fragment into pKK223-3 to generate p1922 for expression in *E. coli*. JM101 was cotransformed with p1922 plus pGroESL or pKK223-3 plus pGroESL (pGroESL encodes *E. coli* chaperonin proteins groEL and groES). Standard methods for the transformation of competent *E. coli* cells were used (Cohen et al., 1971). Double transformants containing plasmids pGroESL and p1922 (or pKK223-3) were selected on media containing 50 μg/mL ampicillin and chloramphenicol. Fresh transformants were grown at 22, 30, or 37 °C to an OD_{600} of 0.6–1.8. Expression was induced by addition of IPTG to a final concentration of 0–1.0 mM. Incubation was continued at 22, 30, or 37 °C for 3–20 h. Maximal expression of active rPDPc was obtained by growth of fresh transformants at 30 °C in the absence of IPTG to an OD_{600} of 1.4–1.8 (18–22 h).

Table I: Peptide Sequences[a]

name	sequence	location
NT	ASTPQKFYLTPPQVN	1–15
T1	HAVGNNEFGAVDHERLSK- - - - -	414–431
T2	VIESGPDQLNDNEYTKFIPPNYytPPYLc- - - - -	297–324
T3	IVGEYLTGMHQQPIAVGGYKVTLGQMHGLLTER	361–394
T6	LRPQDKFLVLATDGLWETMHRQDVVr- - - - -	335–360
T8	LRPQDKFLVLATDGLWETMHRQDVVRIVG- -ltgmhhqq- -	
CB22	LSLPEELAR- - -	432–441

[a] Sequences were determined from tryptic (T) fragments derived from PDPc and a CNBr (CB) fragment derived from PDP. NT is the amino-terminal sequence. Location refers to the deduced sequence of the mature protein (Figure 2).

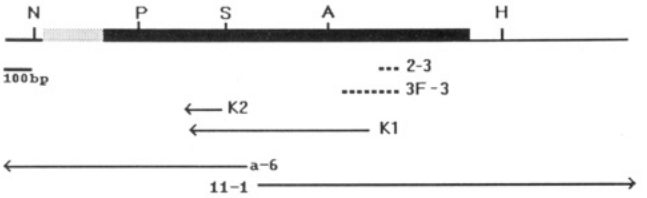

FIGURE 1: Isolation of cDNA fragments. Shown is a schematic of the full-length cDNA for PDPc. The putative presequence is indicated by the shaded bar, and the coding sequence for the mature protein is represented by the solid bar. 5′- and 3′-flanking DNA are shown as thick lines. Some unique restriction sites are *Nar*I (N), *Pvu*II (P), *Sal*I (S), *Aat*II (A), and *Hin*dIII (H). A summary of isolated DNA fragments is shown under the complete cDNA. A dashed line indicates fragments synthesized from genomic DNA. An arrow to the left indicates 5′ RACE fragments. An arrow to the right indicates a 3′ RACE fragment.

Enzyme Assay. Assay of PDPc activity was based on measurement of the initial rate of release of [^{32}P]phosphate from the bovine ^{32}P-labeled pyruvate dehydrogenase complex (Damuni et al., 1984; Niu, 1989). One unit is defined as the amount of phosphatase that releases 1 nmol of [^{32}P]phosphate/min.

Purification of Recombinant PDPc. Cells harvested from a 1.5-L culture (about 6 g wet weight) were resuspended in 30 mL of ice-cold buffer A (20 mM imidazole hydrochloride, pH 7.3, 1 mM EDTA, 0.1% 2-mercaptoethanol, 2 mM phenylmethanesulfonyl fluoride, 2 mM benzamidine, and 10% glycerol). The suspension was passed twice through a French press at 16 000 psi and then was centrifuged at 20 000 rpm for 10 min in a Beckman JA-20 rotor. The extract was applied to a SP-Sepharose column (3 × 12 cm) that had been equilibrated with buffer A. The column was washed with 15 mL of buffer A and then developed with a 300-mL linear gradient from 0.0 to 0.5 M NaCl in buffer A. Active fractions were pooled, dialyzed against buffer A, and concentrated by vacuum dialysis. The solution was applied to a DEAE-Sepharose column (3 × 8 cm) equilibrated with buffer A. The column was washed with 10 mL of buffer A and then developed with a 300-mL linear gradient from 0.0 to 0.3 M NaCl in buffer A. The active fractions were examined by SDS–PAGE to ascertain which contained rPDPc and which contained the truncated rPDPc (M_r ~45 000) and were pooled accordingly and concentrated. For further purification, 0.5-mL aliquots were subjected to fast protein liquid chromatography on a Superose 12 column equilibrated and developed with buffer A containing 0.1 M NaCl. In preliminary experiments, 0.5–2.0 mg of highly purified rPDPc and its truncated form was obtained.

RESULTS

Partial Amino Acid Sequence of PDPc. The experimentally determined peptide sequences accounted for a total of 131 residues (Table I), including the amino-terminal sequence determined by protein electroblotting and microsequencing, the sequences of five tryptic fragments derived from ~16 μg of PDPc, and the sequence of a CNBr fragment derived from PDP. The sequence of peptide T6 ended at an argininyl residue. The sequence of peptide T8 is the same as that of peptide T6, but T8 derives from incomplete digestion, and its sequence overlaps that of peptide T3. This fortunate circumstance yielded a single sequence of 60 residues.

Isolation and Characterization of Bovine Genomic DNA Fragments. Four degenerate oligonucleotide primers were designed on the basis of the combined amino acid sequences of peptides T6 and T3 (Table I). PCR was performed with different combinations of the primers, using several bovine cDNA libraries and bovine genomic DNA as templates. A 74-bp fragment, designated 2–3 (Figure 1), was amplified from bovine genomic DNA using forward primer 50ki2 and reverse primer 50ki3 (Table II). Sequencing of this fragment showed that it encoded 25 residues of the T6+T3 sequence.

Table II: Oligonucleotide Primers for PCR[a]

primer	sequence	location
50ki2	TGGG(GA)AC(AGCT)ATGCA(TC)(AC)G(AGCT)CA(GA)GA(TC)GT	1425 (F)
50ki3	GG(TC)TG(TC)TG(GA)TG(GA)TGCAT	1498 (R)
3F	HindIII TATCAAGCTTGA(TC)CA(GA)(TC)T(GACT)AA(TC)GA(TC)AA(TC)GA(GA)TA(TC)AC	1284 (F)
50ki3ulK	KpnI GCTGGTACCGGTTGCTGATGGTGCATGCCTGTT	1498 (R)
K11H	HindIII TATCAAGCTTGAATGACAATGAATACACCAAG	1292 (F)
K12H	HindIII TATCAAGCTTACACACCTCCTTATCTCACTGCTGA	1331 (F)
K13K	KpnI GCTGGTACCAAGGAGATGTCATTATCAAGCCTC	910 (R)
K14K	KpnI GCTGGTACCACTCGAAGCACTAGGTAGTTGAGG	966 (R)
K15K	KpnI GCTGGTACCTCTTGCCAGTAAGTCCTCAAGCTGTT	818 (R)
K18H	HindIII TATCAAGCTTCGAGTGGCATTTCCTGG	957 (F)
R21K	KpnI GCTGGTACCGGTGGTAAGTTACTTCTGGTTCAG	1378 (R)
c-1	GAAAGCAGATAGCAAGCTTTATTCAC	1900 (R)
K19	EcoRI AGAGGAATTCTAATGGCTTCCACACCGCAGAAG	378 (F)
K22	HindIII TAAAAGCTTTCACTGTTCCTGGTTTTGATATG	1781 (R)

[a] Sequences are listed 5′–3′. Added restriction sites are underlined. Location refers to the nucleotide of the PDPc cDNA at which hybridization to the primer begins and continues in the forward (F) or reverse (R) direction.

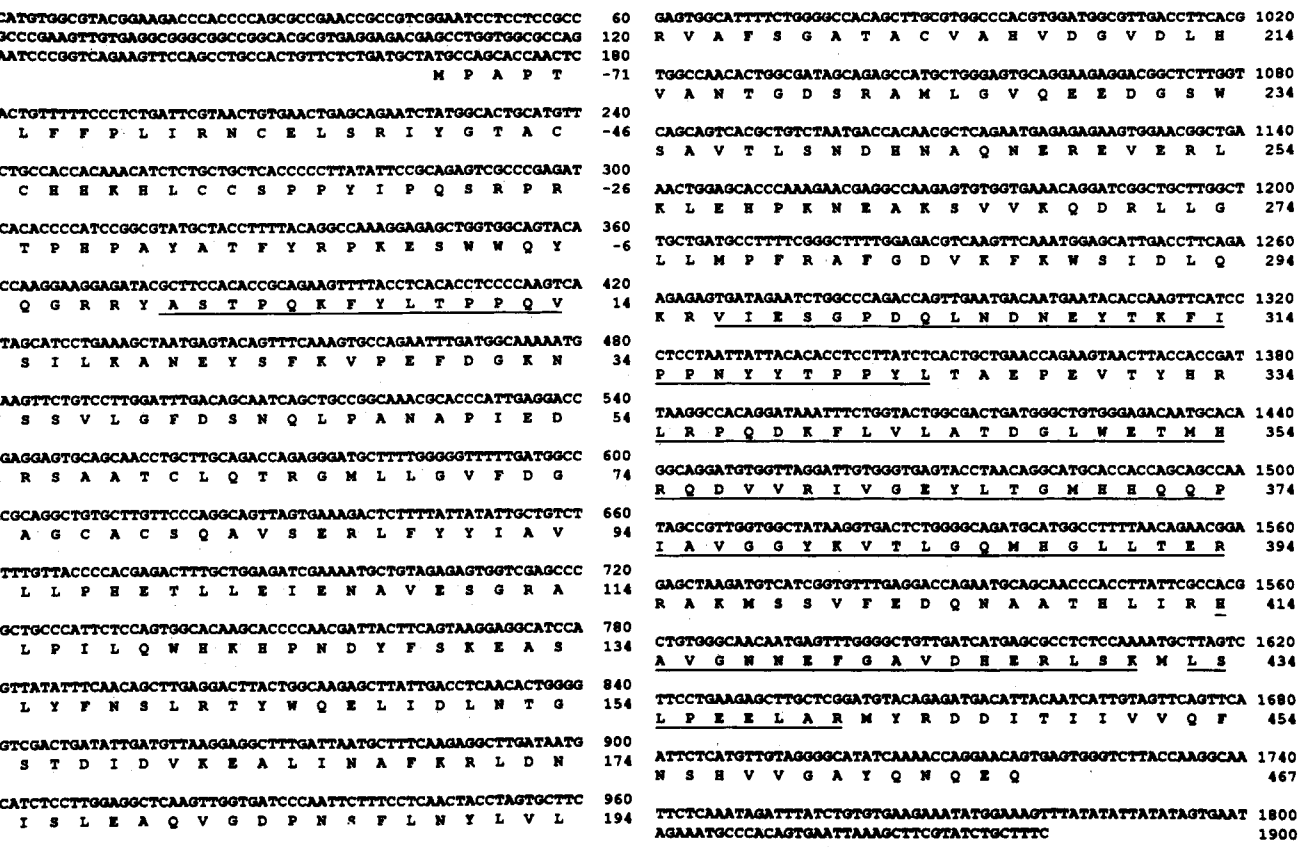

FIGURE 2: Nucleotide sequence of PDPc cDNA and the deduced amino acid sequence (one-letter amino acid symbols). The experimentally determined amino acid sequences are underlined.

On the basis of the sequence of fragment 2–3, a specific reverse primer, 50ki3ulK, was synthesized. Another DNA fragment of 215 bp was amplified from the genomic DNA using specific primer 50ki3ulK and a degenerate forward primer, 3F, based on the sequence of peptide T2 (Table I). This fragment, designated 3F-3, encoded an amino acid sequence of 72 residues, which included most of the sequences of peptides T6, T3, and T2 and an intervening segment of 25 residues. DNA fragment 3F-3 was used as a probe to screen several cDNA libraries. No positive results were obtained.

5′ RACE. In an effort to obtain a full-length cDNA, the 5′ RACE system for rapid amplification of cDNA ends was used. cDNA was synthesized from bovine total RNA with an oligo(dT) primer and the heat-stable reverse transcriptase *rTth*. Amplification primer (AP) from the 5′ RACE system kit and specific reverse primer 50ki3ulK were used to amplify the 5′ segment of cDNA upstream from fragment 3F-3. In the second round of PCR, universal amplification primer (UAP) from the 5′ RACE system and the specific primer R21K were used to amplify a 666-bp DNA fragment, designated K1. Fragment K1 was subcloned between the *Spe*I and *Kpn*I restriction sites in the plasmid Bluescript and sequenced. The nucleotide sequence of K1 contained an open reading frame encoding 222 amino acid residues, including the complete sequence of peptide T2. A similar approach using K15K, a specific reverse primer based on the sequence of fragment K1, and UAP resulted in amplification of a 137-bp fragment, designated K2. The sequence of K2 added 27 bp to the 5′ end of K1.

Specific primers based on the nucleotide sequences of K1 and K2 were used to isolate the remaining portion of the 5′ end of the cDNA using the 5′ RACE system. Primer R21K was used for the synthesis of cDNA from the total RNA. Primer K14K and the AP were used in the first round of PCR. Primer K13K and the UAP were used in the second round of PCR. The product of the second round of PCR was fragment a-6. This fragment, approximately 950 bp in length, was subcloned into pCRII (TA cloning kit) and sequenced. The sequence contained the region coding for the N-terminal segment of PDPc.

3′ RACE. The 3′ end of the cDNA was synthesized using the 3′ RACE system. Primer K18H and the UAP were used in the first round of PCR. A second round of PCR using K11H and the UAP produced fragments that contained the complete sequence of K11H followed by unidentified DNA sequences. Therefore, the second round of amplification was repeated two times using the UAP and primer K12H. The resulting fragments (300–500 bp) were segments of the phosphatase sequence previously determined, but all terminated prematurely at various points between nucleotides 1390 and 1615. Some of the clones contained a poly(A) sequence after the incomplete PDPc sequence. Several of the clones also contained a common 31-bp sequence immediately preceding the poly(A) tail. A reverse primer (c-1) based on a portion of the 31-bp common sequence was synthesized and used with primer K18H and the cDNA as a template in a PCR. The fragment produced (11–1, 1.0 kb) was subcloned into pCRII and sequenced. The sequence showed the fragment contained the intact cDNA from the K18H primer binding site through the poly(A) tail. Fragment 11–1 contained the sequences coding for peptides T6, T3, T1, T2, and CB22 (Table I).

Reconstruction of the Full-Length cDNA. Three overlapping fragments were used to reconstruct the cDNA: K1 (corresponding to nucleotides 709–1378 of the full-length cDNA), a-6 (1–912), and 11–1 (959–1900). a-6 was digested with *Sal*I at a site at the 5′ end introduced by the UAP and at nucleotide 841. K1 had been previously subcloned into

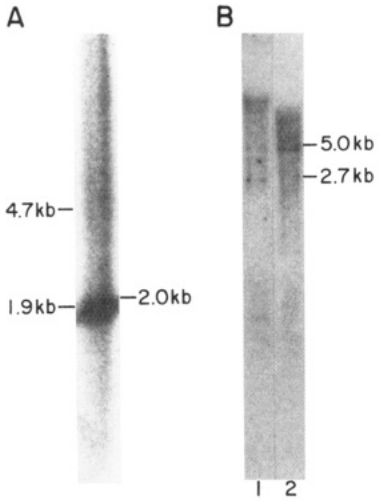

FIGURE 3: (A) RNA blot analysis. Autoradiograph of a blot of bovine lung total RNA probed with radiolabeled DNA fragment K1 (nucleotides 709–1378). The size of the hybridizing fragment is indicated on the right. The positions of the bovine ribosomal RNAs are shown on the left. The lane contained 40 μg of total RNA. (B) Southern blot analysis. Autoradiograph of radiolabeled DNA fragment K1 (nucleotides 709–1378) hybridized to a blot of bovine kidney genomic DNA. DNA was digested with *Bam*HI (lane 1) or *Hin*dIII (lane 2). Sizes of hybridizing fragments are shown on the right. Each lane contained 4 μg of genomic DNA.

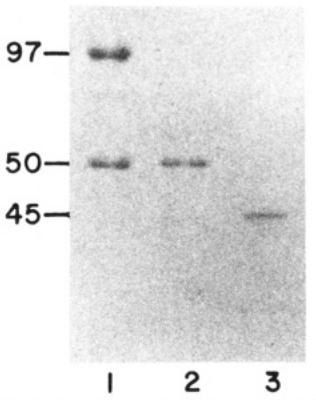

FIGURE 4: SDS–PAGE patterns of highly purified bovine PDP (lane 1), rPDPc (lane 2), and truncated rPDPc (lane 3). Approximately 2.5, 2, and 1 μg of protein was applied to lanes 1, 2, and 3, respectively. The gel was stained with Coomassie brilliant blue. Molecular weights are shown as $M_r \times 10^{-3}$.

Table III: Effects of Mg^{2+} and Ca^{2+} on Recombinant PDPc Activity

assay mixtures[a]	relative activity[b]
rPDPc	100
rPDPc – Mg^{2+}	0
rPDPc + EGTA	15
rPDPc + EGTA + Ca^{2+}	87

[a] The assay mixtures contained 0.03 unit of recombinant rPDPc, the ^{32}P-labeled pyruvate dehydrogenase complex, 10 mM $MgCl_2$, 0.1 mM $CaCl_2$, and, where indicated, 0.5 mM ethylene glycol bis(β-aminoethyl ether)-*N*,*N*,*N*′,*N*′-tetraacetic acid (EGTA) and 0.7 mM $CaCl_2$ in a total volume of 50 μL. [b] Activity is expressed as percent of the activity observed with rPDPc in the standard assay (Damuni et al., 1984).

Bluescript at the *Spe*I and *Kpn*I sites. The plasmid was digested at *Sal*I (corresponding to *Sal*I at nucleotide 842) and at *Kpn*I in the polylinker downstream of the K1 fragment. The a-6 *Sal*I and K1 *Sal*I–*Kpn*I fragments were subcloned into pCRII (digested with *Sal*I and *Kpn*I). The 11–1 fragment (already subcloned into pCRII) was digested with *Aat*II (corresponding to nucleotide 1227) and *Kpn*I (in the polylinker downstream of the insert). The fragment was inserted into the *Aat*II and *Kpn*I sites of the combined a-6/K1 plasmid described above. The final construction was termed p2311. The cDNA is 1900 bp in length and contains 164 bp of 5′-flanking DNA, 122 bp of 3′-flanking DNA, and an open reading frame of 1614 bp (Figure 2). The sequence coding for the N-terminal segment begins at nucleotide 378, leaving a 71-residue presequence. The mature protein is 467 residues in length and has a calculated M_r of 52 625. The M_r is consistent with the values reported for the isolated protein (Teague et al., 1982; Pratt et al., 1982).

Although there is an ATG codon at nucleotides 3–5 followed by an open reading frame, the ATG codon at nucleotides 165–167 corresponding to Met-71 was chosen as the probable start of the protein because initiation of protein synthesis so close to the 5′ end of the transcript would be very unusual. Furthermore, the maximum length reported for a presequence of a mitochondrial protein is about 70 amino acid residues (Hartl & Neupert, 1990). To ensure that no in-frame ATG codon downstream of ATG 165–167 was overlooked, the 5′ end of the cDNA from multiple clones was sequenced many times in both directions with several different primers.

RNA Analysis. Hybridization of DNA fragment K1 to bovine heart total RNA indicated the existence of a single transcript of approximately 2.0 kb (Figure 3A). Hybridization detected bands of similar size in total RNA from bovine brain and lung (data not shown). The size of this transcript is consistent with the length of the cloned cDNA (1.9 kb).

Analysis of Genomic DNA. Hybridization of DNA fragment K1 to restriction digests of bovine kidney genomic DNA demonstrated hybridization to a single 2.7-kb *Bam*HI DNA fragment and a single 5.0-kb *Hin*dIII fragment (Figure 3B). The small size of these fragments suggests there is a single gene which hybridizes to this probe in the bovine genome.

Expression, Purification, and Properties of Recombinant PDPc. Cultures of *E. coli* transformants which coexpressed p1922 and pGroESL were grown under conditions optimal for expression of soluble active rPDPc, as determined by immunoblot analysis and assay of phosphatase activity. A truncated form of rPDPc, with $M_r \sim 45\,000$, was detected in cell extracts by immunoblot analysis. The amount of truncated rPDPc varied with cell growth conditions and with purification procedures. Both rPDPc and the truncated rPDPc were purified to near homogeneity as described under Experimental Procedures. The specific activities of the two proteins were 292 and 345 units/mg of protein, respectively. These values are comparable to the specific activity of highly purified bovine PDP (358 units/mg of protein). When analyzed by SDS–PAGE (Figure 4) and by immunoblotting (data not shown), the purified rPDPc showed a major band ($M_r \sim 50\,000$) which comigrated with the catalytic subunit of bovine PDP. Analysis of rPDPc on a calibrated Superose 12 column indicated that it was monomeric (data not shown). rPDPc, like bovine PDP, is Mg^{2+}-dependent and Ca^{2+}-stimulated (Table III).

Comparison of Amino Acid Sequences of PDPc and PP2C. A BLAST search of the GenBank and PIR data bases revealed significant homology between PDPc and the PP2Cα (Tamura et al., 1989) and PP2Cβ (Wenk et al., 1992) isoforms but no obvious homologies between PDPc and other serine/threonine-specific protein phosphatases. Application of the progressive alignment method of Feng and Doolittle (1987) showed 20% identity between PDPc and PP2Cα and 22% identity between PDPc and PP2Cβ (Figure 5). The overall pattern indicates that PDPc and PP2C evolved from a common ancestor.

```
PDPc  ASTPQKFYLTPPQVNSILKANEYSFKVPEFDGKNVSSVLGFDSNQLPANAPIEDRRSAATCLQTRGMLLGVFDGHAGCACSQAVS   85
2Cα     MGAFLDKPKMEKHNAQGQGNGLRY      GLSSMQGWRVEMEDAHTAVIGLPSG   LETWS FFAVYDGHAGSQVAKYCC   72
2Cβ     MGAFLDKPKTEKHNAHGAGNGLRY      GLSSMQGWRVEMEDAHTAVVGIPHG   LEDWS FFAVYDGHAGSRVANYCS   72

PDPc  ERLFYYIAVSLLPHETLLEIENAVESGRALLPILQWHKHPNDYFSKEASKLYFNSLRTYWQELIDLNTGESTDIDVKEALINAFK  170
2Cα   EHLLDHITNN        QDFKGSAGA     PSVE                                        NVKNGIRTGFL  106
2Cβ   THLLEHITTN        EDFRAADKSGFALEPSVE                                        NVKTGIRTGFL  111

PDPc  RLDNDISLEAQVGDPNSFLNYLVLRVAFSGATACVAHVDGVDLHVANTGDSRAMLGVQEEDGSWSAVTLSNDHNAQNEREVERLK  255
2Cα   EIDEHMRVMSEKKHG           ADRSGSTAVGVLISPQHTYFINCGDSRGLLCRNRK     VHFFTQDHKPSNPLEKERI   175
2Cβ   KIDEYMRNFSDLRNG           MDRSGSTAVGVMISPTHIYFINCGDSRAVLCRNGQ     VCFSTQDHKPCNPMEKERI   180

PDPc  LEHPKNEAKSVVKQDRLLGLLMPFRAFGDVKFKWSIDLQKRVIESGPDQLNDNEYTKFIPPNYYTPPYLTAEPEV TYHRLRPQD  339
2Cα        QNAGGSVMIQRVNGSLAVSRALGDFDY       KCVHGKGP                   TEQLVSPEPEVHDIERSEEDD  231
2Cβ        QNAGGSVMIQRVNGSLAVSRALGDYDY       KCVDGKGP                   TEQLVSPEPEVYEILRAEEDE  236

PDPc  KFLVLATDGLWETMHRQDVVRIVGEYLTGMHHQQ      PIAVGGYKVTLGQMHGLLTERRAKMSSVFEDQNAATHLIRHAVGNN  419
2Cα   QFIILACDGIWDVMGNEELCDFVRSRLEVTDDLEKVCNEVVDTCLYKGSRDNMSVILICFPNAPKVSAEAVKKEAELDKYLENRV  316
2Cβ    FVVLACDGIWDVMSNEELCEFVNSRLEVSDDLENVCNWVVDTCLHKGSRDNMSIVLVCFANAPKVSDEAVKRDLELDKHLESRV  320

PDPc  EFGAVDHERLSKMLSLPEELARMYRDDITII      VVQFNSHVVGAYQNQEQ                                  467
2Cα   E EIIKKQG EGVPDLVHVMRTLASENIPSLPPGGELASKRNVIEAVYNRLNPYKNDDTDSASTDDMW                    382
2Cβ   E EIMQKSGEEGMPDLAHVMRILSAENIPNLPPGGGLAGKRNVIEAVYSRLNPNKDNDGGAGDLEDSLVAL                 390
```

FIGURE 5: Alignment of amino acid sequences (single-letter code) of bovine PDPc and rat PP2Cα and PP2Cβ. The deduced sequences are aligned for maximum similarity (Feng & Doolittle, 1987). Identical residues are indicated by asterisks.

DISCUSSION

All attempts to isolate clones encoding PDPc from bovine cDNA libraries were unsuccessful. These attempts included use of antisera, PCR with degenerate and specific oligonucleotide primers, and hybridization with specific probes. A PCR-based approach using bovine genomic DNA and cDNA synthesized from bovine total RNA as templates eventually proved to be successful. Initially, a 74-bp fragment (nucleotides 1425–1498) and a 215-bp fragment (nucleotides 1284–1498) were amplified from bovine genomic DNA with primers designed from experimentally determined internal amino acid sequences. The remainder of the coding region and 5′- and 3′-flanking cDNA were obtained by 5′ RACE and 3′ RACE with bovine total RNA as the template for the cDNA. Unexpected difficulty was encountered in generating the 3′ end of the cDNA. The isolation of fragments containing deletions of various sizes in the general area of nucleotides 1400–1600 followed by a poly(A) tail suggests that the reverse transcriptase had difficulty proceeding efficiently through that region during cDNA synthesis, or similarly, the DNA polymerase had difficulty during PCR. A stable stem and loop structure could cause either enzyme to "skip" a section resulting in a deletion. Such a stem and loop may also cause the poly(A) tail to fold back in such a way that it would be inaccessible for hybridization to oligo(T)–cellulose. Such a hybridization is currently used in isolation of poly(A)$^+$ RNA for the synthesis of cDNA libraries. If the mRNA for PDPc did not bind efficiently to the oligo(T)–cellulose, it would not be expected to be found in the libraries. This possibility is supported by our inability to detect the cDNA in any cDNA libraries from a variety of sources, even using sequence-specific primers.

Attempts were made to overexpress PDPc in *E. coli* as a fusion protein with glutathione *S*-transferase or with maltose-binding protein. Whether expressed alone or coexpressed with *E. coli* chaperonin proteins groES and groEL, which are known to assist the folding and assembly of polypeptides *in vivo* (Gething & Sambrook, 1992), the fusion proteins remained largely insoluble. Soluble PDPc was expressed from an inducible promoter, not fused to any other protein. Coexpression with groES and groEL increased the total amount while maintaing the solubility of the rPDPc produced. Optimal conditions for production of active soluble rPDPc were established by systematic examination of growth conditions, including temperature, time, and concentration of the inducer IPTG. Slow growth (18–22 h) at 30 °C in the absence of IPTG gave the best results. rPDPc was estimated to be 2–3% of the soluble *E. coli* protein, based on a specific activity of 292 units/mg of protein found for the purified enzyme. The properties of the rPDPc resembled those of the native enzyme, including specific activity, absolute dependence on Mg^{2+}, and stimulation by Ca^{2+}. These results confirm the previous observation that native PDPc, obtained by treatment of bovine PDP with 1 M KSCN followed by chromatography on Sephadex G-100, showed an absolute dependence on Mg^{2+} and was markedly stimulated by Ca^{2+} (Teague et al., 1982). The Ca^{2+} effect is apparently a property of the catalytic subunit rather than a property of the 97-kDa subunit or of the heterodimer. It is interesting to note in this connection that PDPc possesses a possible Ca^{2+}-binding motif, DNDISLEAQVGD (residues 173–184) (Marsden et al., 1990).

Whether or not and, if so, how the 97-kDa subunit modulates the activity of the catalytic subunit remain to be determined. An interesting possibility is that the 97-kDa subunit is involved in the insulin-induced signaling pathway leading to stimulation of the activity of the pyruvate dehydrogenase complex. This stimulation results from dephosphorylation and consequent activation of the pyruvate dehydrogenase (E_1) component of the complex and apparently involves activation of PDP (Denton et al., 1989). Cloning and expression of cDNA for PDPc and,

eventually, cDNA for the 97-kDa subunit of PDP should facilitate studies on the insulin-induced activation of pyruvate dehydrogenase.

Alignment of the deduced amino acid sequences of the cytosolic PP2C isoforms α and β with the mitochondrial PDPc (Figure 5) indicates that these protein serine/threonine phosphatases evolved from a common ancestor. Because mitochondria are believed to have evolved from bacterial endosymbionts within some ancestral type of eukaryotic cell (Gray & Doolittle, 1982), the apparent common origin of PDPc and PP2C raises an interesting question concerning the ancestry of PP2C. It is interesting to note in this connection that the mammalian mitochondrial branched-chain α-keto acid dehydrogenase kinase appears to be more closely related in sequence to prokaryotic protein histidine kinases than to eukaryotic protein serine/threonine kinases (Popov et al., 1992).

ACKNOWLEDGMENT

We thank Dr. Kenneth Walsh for assistance with amino acid analyses; Drs. Russell Doolittle and Lawrence Poulsen for assistance with amino acid sequence alignments; Drs. Dean Appling and H. Y. Lim Tung for helpful discussions; and Dr. K. Ümit Yüksel (Texas College of Osteopathic Medicine, Fort Worth) and the Protein Sequencing Center, University of Texas at Austin, for amino acid sequence analyses.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990) *J. Mol. Biol. 215*, 403–410.
Bialojan, C., & Takai, A. (1988) *Biochem. J. 256*, 283–290.
Cohen, P. (1989) *Annu. Rev. Biochem. 58*, 453–508.
Cohen, S. N., Chang, A. C. Y., & Hsu, L. (1971) *Proc. Natl. Acad. Sci. U.S.A. 69*, 2110–2114.
Damuni, Z., Humphreys, J. S., & Reed, L. J. (1984) *Biochem. Biophys. Res. Commun. 124*, 95–99.
Damuni, Z., Tung, H. Y. L., & Reed, L. J. (1985) *Biochem. Biophys. Res. Commun. 133*, 878–883.
Denton, R. M., Randle, P. J., & Martin, B. R. (1972) *Biochem. J. 128*, 161–163.
Denton, R. M., Midgley, P. J. W., Rutter, G. A., Thomas, A. P., & McCormack, J. G. (1989) *Ann. N.Y. Acad. Sci. 573*, 285–296.
Feinberg, A., & Vogelstein, B. (1983) *Anal. Biochem. 132*, 6–13.
Feng, D.-F., & Doolittle, R. F. (1987) *J. Mol. Evol. 25*, 351–360.
Gething, M. J., & Sambrook, J. (1992) *Nature 355*, 33–45.
Gray, M. W., & Doolittle, W. F. (1982) *Microbiol. Rev. 46*, 1–42.
Hartl, F.-U., & Neupert, W. (1990) *Science 247*, 930–938.
Haystead, T. A. J., Sim, A. T. R., Carling, D., Honnor, R. C., Tsukitani, Y., Cohen, P., & Hardie, D. G. (1989) *Nature 337*, 78–81.
Hiraga, A., Kikuchi, K., Tamura, S., & Tsuiki, S. (1981) *Eur. J. Biochem. 119*, 503–510.
Laemmli, U. K. (1970) *Nature 227*, 680–685.
Lawson, J. E., Behal, R. H., & Reed, L. J. (1991) *Biochemistry 30*, 2834–2839.
Le Trong, H., Beier, N., Sonnenburg, W. K., Stroop, S. D., Walsh, K. D., Beavo, J. A., & Charbonneau, H. (1990) *Biochemistry 29*, 10280–10288.
Linn, T. C., Pettit, F. H., & Reed, L. J. (1969) *Proc. Natl. Acad. Sci. U.S.A. 62*, 234–241.
Marsden, B. J., Shaw, G. S., & Sykes, B. D. (1990) *Biochem. Cell Biol. 68*, 587–601.
Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035–10038.
McGowan, C. H., & Cohen, P. (1987) *Eur. J. Biochem. 166*, 713–721.
Miran, S. G., Lawson, J. E., & Reed, L. J. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 1252–1256.
Niu, X.-D. (1989) Dissertation, University of Texas, Austin.
Pettit, F. H., Roche, T. E., & Reed, L. J. (1972) *Biochem. Biophys. Res. Commun. 49*, 563–571.
Popov, K. M., Zhao, Y., Shimomura, Y., Kuntz, M. J., & Harris, R. A. (1992) *J. Biol. Chem. 267*, 13127–13130.
Pratt, M. L., Maher, J. F., & Roche, T. E. (1982) *Eur. J. Biochem. 125*, 349–355.
Schuler, G. D., Altschul, S. F., & Lipman, D. J. (1991) *Proteins: Struct., Funct., Genet. 9*, 180–190.
Shenolikar, S., & Nairn, A. (1991) *Adv. Second Messenger Phosphoprotein Res. 23*, 1–121.
Tamura, S., Lynch, K. R., Larner, J., Fox, J., Yasui, A., Kikuchi, K., Suzuki, Y., & Tsuiki, S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 1796–1800.
Teague, W. M., Pettit, F. H., Wu, T.-L., Silberman, S. R., & Reed, L. J. (1982) *Biochemistry 21*, 5585–5592.
Wenk, J., Trompeter, H.-I., Pettrich, K.-G., Cohen, P. T. W., Campbell, D. G., & Mieskes, G. (1992) *FEBS Lett. 297*, 135–138.
Yanisch-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103–119.
Zhou, C., Yang, Y., & Jong, A. Y. (1990) *BioTechniques 8*, 172–173.