Proc. Natl. Acad. Sci. USA
Vol. 80, pp. 3661–3665, June 1983
Biochemistry

Primary structure of chicken muscle pyruvate kinase mRNA

(glycolytic enzymes/mRNA purification/cDNA cloning/developmental regulation)

NILS LONBERG AND WALTER GILBERT

The Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138

Contributed by Walter Gilbert, March 21, 1983

ABSTRACT **We have determined the cDNA sequence corresponding to chicken muscle pyruvate kinase mRNA; the predicted coding region spans 529 amino acids and establishes the complete amino acid sequence for the vertebrate enzyme. We demonstrate that the level of mRNA for this enzyme is under developmental control and suggest a structural model for the protein kinase-mediated regulation of the mammalian liver isozyme. We report a method for the direct analysis of, and the preparation of cDNA probes from, mRNA which has been fractionated on methylmercury/agarose gels.**

The glycolytic enzyme pyruvate kinase (PK; ATP:pyruvate 2-*O*-phosphotransferase, EC 2.7.1.40) occurs in all cells. The overall structure of the native protein is the same in bacteria, plants, and animals: a tetramer composed of identical 50- to 60-kilodalton subunits (1). Although the crystal structure of the mammalian muscle enzyme has been solved at 2.6-Å resolution (2), the primary structure is unknown. We have isolated two overlapping cDNA clones spanning the coding region for the adult chicken muscle isozyme and determined the nucleotide sequence together with 5′ nontranslated sequences not included in the clones. The predicted primary structure of the enzyme consists of 529 amino acids adding up to a molecular weight of 57,865.

MATERIALS AND METHODS

DNA and Enzymes. Enzymes were purchased from New England BioLabs, except where indicated. Plasmid DNA was purified by centrifugation in ethidium bromide/CsCl density gradients.

mRNA Isolation. Total nucleic acid was isolated from adult White Leghorn chicken breast muscle tissue by the procedure of Lomedico and Saunders (3), resuspended in 10 mM Hepes, pH 7.5/1 mM EDTA/0.5 M LiCl/0.5% $NaDodSO_4$, and bound to oligo(dT)-cellulose (type 3, Collaborative Research, Waltham, MA). Poly(A)$^+$ RNA was eluted with 10 mM Hepes, pH 7.5/1 mM EDTA and precipitated with ethanol. The pellet was washed with ethanol, resuspended in water or dilute buffer, and stored at −20°C. Concentration was measured by absorption at 260 nm.

***In Vitro* Translation.** mRNA was translated in a message-dependent rabbit reticulocyte lysate system. Final concentrations of the various components of the system were: poly(A)$^+$ RNA, 5–10 μg/ml; [^{35}S]methionine (Amersham), 1–2 mCi/ml (1 Ci = 3.7×10^{10} Bq); micrococcal nuclease-treated reticulocyte lysate (4), 20% (vol packed cells/vol); hemin, 10 μM; creatine phosphokinase, 20 μg/ml; each amino acid except methionine, 50 μM (adjusted to pH 7.0 with HCl); KOAc, 100 mM; $Mg(OAc)_2$, 500 μM; dithiothreitol, 500 μM; and creatine phosphate, 10 mM. Reaction time was 30–60 min at 37°C. ^{35}S-Labeled peptides were separated by $NaDodSO_4$/polyacrylamide gel electrophoresis (5) and detected by fluorography (EN^{3}HANCE, New England Nuclear).

Immunoprecipitation. Antiserum was raised by injecting rabbits with 0.6 mg of adult chicken muscle PK (gift of Deborah Hurt, University of North Carolina at Chapel Hill) together with Freund complete adjuvant (Colorado Serum, Denver, CO). Rabbits were given booster injections of 0.6 mg PK in incomplete adjuvant after 3 weeks and bled 2 weeks later. Immunoprecipitations were carried out according to Kessler (6).

mRNA Fractionation. Poly(A)$^+$ RNA was further fractionated on methylmercury/agarose gels (7). mRNA was denatured with 10 mM methylmercury hydroxide for 1 min at 65°C prior to electrophoresis through a 1.5% low-melting-temperature agarose (SeaPlaque, Marine Colloids, Rockland, ME) gel containing 5 mM methylmercury hydroxide. Approximately 100 μg of mRNA was used per cm^2 (cross section) of gel. After the desired separation was obtained, the gel was rinsed three times for 20 min each in 10 mM Hepes, pH 7.5/1 mM EDTA/10 mM 2-mercaptoethanol. The gel was then equilibrated for at least 30 min in a 1:3 dilution of this buffer. Plugs (5–10 μl) were removed from the gel at 2- to 3-mm intervals by using glass capillary tubes. Several plugs were removed at each interval and stored at −70°C.

The mRNA contained within the gel plugs was assayed directly without removing the agarose by *in vitro* translation in the rabbit reticulocyte lysate system described above. Each gel plug was added to an aqueous solution containing [^{35}S]methionine at 2 mCi/ml to give a final volume of 25 μl. The gel plugs were melted at 65°C for 5 min and then cooled to 37°C. The remaining components of the lysate system, equilibrated at 37°C, were then added to each sample to a final reaction volume of 50 μl.

Preparation of Hybridization Probe. Gel plugs were also used for cDNA synthesis without removing the agarose. Plugs were melted for 3 min at 65°C in 30–40 μl of reverse transcription buffer adjusted to (final concentrations): 30 μM [α-^{32}P]-dCTP (800 Ci/mmol, New England Nuclear); 1 mM dATP, dTTP, and dGTP; oligo(dT) (Collaborative Research) at 100 μg/ml; 8 mM $MgCl_2$; 50 mM Tris (pH 8.3); and 20 mM 2-mercaptoethanol. The tube was cooled to 42°C and avian myoblastosis virus reverse transcriptase (J. Beard, Life Sciences, St. Petersburg, FL) was added to a concentration of 800 units/ml. The reaction mixture was incubated for 1 hr at 42°C, phenol extracted, alkali treated to remove RNA (0.6 M NaOH, 15 min at 50°C), neutralized with HOAc, and ethanol precipitated. The resuspended cDNA was electrophoresed through a 5% polyacrylamide gel in order to remove low molecular weight products and unincorporated nucleotides. Molecules >200 nucleotides

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

Abbreviations: PK, pyruvate kinase; bp, base pair(s).

long were eluted from the gel and used as hybridization probe; $>10^7$ cpm was obtained by using a single gel plug.

Cloning of Double-Stranded cDNA. Double-stranded cDNA was synthesized from total chicken muscle poly(A)$^+$ RNA as described by Wickens *et al.* (8).

The 5′ region of the PK message was cloned by using a 106-base-pair (bp) *Hae* III/*Dde* I fragment of pPK102 (spanning bp 955 through 1,060 in Fig. 3) as a specific primer for the reverse transcriptase reaction. Single-stranded primer was isolated on a strand-separation polyacrylamide gel (9). Approximately 5 pmol of primer was hybridized to 100 μg of total poly(A)$^+$ RNA in 100 μl of 10 mM Pipes, pH 6.4/200 mM NaCl/0.1 mM EDTA for 20 min at 65°C after briefly denaturing at 90°C. Annealed primer was extended with reverse transcriptase as described above, alkali treated, and passed over a Sephadex G-150 column (Pharmacia). Short (10–20 nucleotides) poly(dC) tails were added to the 3′ ends by using terminal deoxynucleotide transferase (10). Second-strand synthesis was carried out with *Escherichia coli* polymerase I primed with oligo(dG) (Collaborative Research) as described by Land *et al.* (11).

After addition of homopolymeric dC tails, double-stranded cDNA was hybridized to dG-tailed *Pst* I-cut pBR322 (12) and used to transform *E. coli* strain HB101 (13). Tetracycline-resistant colonies were screened by colony hybridization (14).

In order to construct a library enriched for long cDNA inserts, tailed double-stranded cDNA was electrophoresed through a 1.5% low-melting-temperature agarose gel. Molecules larger than 400 bp were purified away from the gel matrix by heating at 65°C followed by phenol extraction. cDNA synthesized by primer extension was size selected on a 5% polyacrylamide gel (12) prior to annealing and transformation.

Nucleic Acid Sequence Analysis. Nucleotide sequence determinations were carried out as described by Maxam and Gilbert (9). DNA fragments isolated from cloned plasmids were labeled at their 3′ ends by using the large fragment of *E. coli* DNA polymerase I. The sequence of the extreme 5′ end of the message was obtained by analyzing the structure of a specifically primed single-stranded cDNA molecule. A 57-nucleotide primer, spanning bp 87 through 143 (see Fig. 3), was separated from its complementary strand on a 6% polyacrylamide/urea gel (9). Twenty picomoles of 5′-end-labeled (9) primer was hybridized to 140 μg of poly(A)$^+$ RNA and extended with reverse transcriptase as described above. Full-length product was purified on a 5% polyacrylamide/urea gel (9), and its sequence was determined.

RNA Blot Hybridizations. RNA was denatured in 50% formamide/6% formaldehyde and separated on a 1.5% agarose gel in the presence of 6% formaldehyde/20 mM sodium morpholinosulfonate (Mops), pH 7.0/5 mM NaOAc/1 mM Na_2EDTA (15). The gel was washed for 20 min in 50 mM NaOH, neutralized with 0.1 M Tris (pH 7.5), and equilibrated in 0.15 M NaCl/0.015 M sodium citrate before blotting (GeneScreen, New England Nuclear). Transfer was carried out in 0.15 M NaCl/0.015 M sodium citrate. The membrane-bound RNA was hybridized to gel-purified, nick-translated (16) pPK102 cDNA insert as described by Alwine *et al.* (17).

RESULTS AND DISCUSSION

cDNA Cloning Strategy. Initially, we used muscle tissue from 2- to 3-day-old chicks as a source of mRNA from which to clone PK DNA. This turned out to be a poor choice. Chicken skeletal muscle PK activity increases after hatching (18); upon observing a similar developmental increase in *in vitro*-translatable PK mRNA, we switched to adult muscle tissue. With adult mRNA, >1% of the incorporated [^{35}S]methionine is associated with PK.

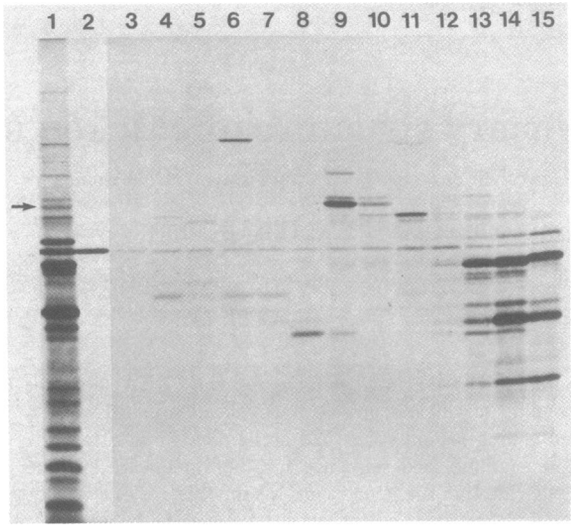

FIG. 1. *In vitro* translations of mRNA fractionated on a 1.5% agarose/methylmercury gel. Poly(A)$^+$ RNA (40 μg) was electrophoresed at 5 V/cm through a 0.5 × 4 × 11 cm gel for 4.5 hr. Adjacent 5-μl plugs were assayed by *in vitro* translation. Translation products were separated on a $NaDodSO_4$/10% polyacrylamide gel; 2.5 μl of each sample was loaded per lane. Lanes: 1, total poly(A)$^+$ RNA; 2, control (no RNA); 3–15, plugs taken at 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, and 48 mm from the origin of the methylmercury/agarose gel. Arrow, position of PK.

This relative abundance allowed us to purify PK message to 60% homogeneity on methylmercury/agarose gels, assaying the mRNA by *in vitro* translation and immunoprecipitation.

We separated poly(A)$^+$ RNA in low-melting-temperature agarose/methylmercury gels and carried out subsequent steps without removing the agarose. The melted agarose had no affect on the translation or reverse transcription of the mRNA. Direct assay avoided losses involved in separating RNA from the agarose, allowed small samples to be analyzed and transcribed, and took maximal advantage of the resolution obtained on the gel. We removed core samples at 2- to 3-mm intervals by using a glass capillary pipet. The 5- to 10-μl agarose cylinders were melted and added directly to a message-dependent translation system. Fig. 1 shows the analysis, by $NaDodSO_4$/polyacrylamide gel electrophoresis, of the translation products. By immunoprecipitation, we identified PK and located its message on the methylmercury/agarose gel. With this fractionation procedure we obtained a 50-fold purification of PK message (estimated from densitometer scans); in a typical experiment, 1.2% of the incorporated [^{35}S]methionine appeared in the PK band with total poly(A)$^+$ RNA as a template whereas 60% of the

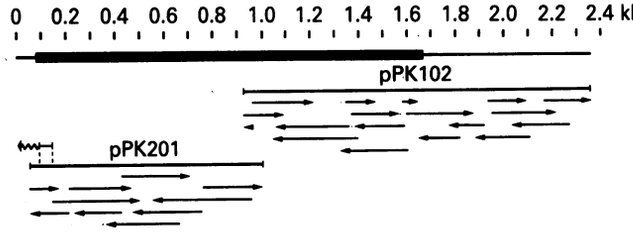

FIG. 2. Strategy for determining the sequence of the PK message. The cDNA inserts of pPK102 and pPK201 were fragmented with restriction endonucleases and analyzed to obtain the bulk of the sequence. Arrows show extent and direction (away from labeled end) of sequence determinations. The extreme 5′ sequence was determined from the primer extension product indicated by the wavy arrow. The heavy line indicates the proposed coding region. Distance from the 5′ end is given in kilobases.

label was incorporated into PK with gel-purified RNA.

A second plug, corresponding to the peak of mRNA activity (lane 9, Fig. 1), served (after melting) as template for synthesis of α-^{32}P-labeled single-stranded cDNA probe. We screened a library of 2,000 bacterial transformants with this probe, picked 5 positive clones, and assayed them by hybridization selection (19). Three of these clones contained sequences that selected PK mRNA; the two other clones selected message coding for a protein that migrated slightly slower than PK on NaDodSO$_4$/polyacrylamide gels; This is the major species which copurifies with PK mRNA; its translation product appears immediately above PK in lane 9 of Fig. 1.

Because cDNA for the first library was not size selected, the largest of the three PK clones contained an insert of only 300 bp. We labeled this insert with ^{32}P by nick-translation (16) and screened a second library constructed from cDNA that had been size selected on an agarose gel. Fifty-four positive clones appeared out of approximately 10^4 bacterial colonies. The largest clone, pPK102, contained a 1.4-kilobase insert. We determined from the nucleotide sequence that it contained the COOH-terminal half of the coding region for chicken muscle PK along with 684 bp of 3′ noncoding sequence.

To obtain the rest of the coding sequence, we primed the reverse transcription of total mRNA with a restriction fragment of pPK102. We separated the strands of the 106-bp *Hae* III–*Dde* I fragment spanning nucleotides 955 through 1,060 (see

```
                                                                                                                SerLysHis
AGCGGCGGCGGCGCTCGGGCAGGACAGGCTTTGGGCACGGCGGCGGAAGCAGCAGCAGGAGACACCGAACTCCAGTAACCATGTCGAAGC    90

  HisAspAlaGlyThrAlaPheIleGlnThrGlnGlnLeuHisAlaAlaMetAlaAspThrPheLeuGluHisMetCysArgLeuAspIle
ACCACGATGCAGGGACCGCTTTCATCCAGACCCAGCAGCTGCACGCTGCCATGGCAGACACCTTTCTGGAGCACATGTGCCGCCTGGACA   180

  AspSerGluProThrIleAlaArgAsnThrGlyIleIleCysThrIleGlyProAlaSerArgSerValAspLysLeuLysGluMetIle
TCGACTCCGAGCCAACCATTGCCAGAAACACCGGCATCATCTGCACCATCGGCCCAGCCTCCCGCTCTGTGGACAAGCTGAAGGAAATGA   270

  LysSerGlyMetAsnValAlaArgLeuAsnPheSerHisGlyThrHisGluTyrHisGluGlyThrIleLysAsnValArgGluAlaThr
TTAAATCTGGAATGAATGTTGCCCGCCTCAACTTCTCGCACGGCACCCACGAGTATCATGAGGGCACAATTAAGAACGTGCGAGAGGCCA   360

  GluSerPheAlaSerAspProIleThrTyrArgProValAlaIleAlaLeuAspThrLysGlyProGluIleArgThrGlyLeuIleLys
CAGAGAGCTTTGCCTCTGACCCGATCACCTACAGACCTGTGGCTATTGCACTGGACACCAAGGGACCTGAAATCCGAACTGGACTCATCA   450

  GlySerGlyThrAlaGluValGluLeuLysLysGlyAlaAlaLeuLysValThrLeuAspAsnAlaPheMetGluAsnCysAspGluAsn
AGGGAAGTGGCACAGCAGAGGTGGAGCTCAAGAAGGGCGCAGCTCTCAAAGTGACGCTGGACAATGCCTTCATGGAGAACTGCGATGAGA   540

  ValLeuTrpValAspTyrLysAsnLeuIleLysValIleAspValGlySerLysIleTyrValAspAspGlyLeuIleSerLeuLeuVal
ATGTGCTGTGGGTGGACTACAAGAACCTCATCAAAGTTATAGATGTGGGCAGCAAAATCTATGTGGATGACGGTCTCATTTCCTTGCTGG   630

  LysGluLysGlyLysAspPheValMetThrGluValGluAsnGlyGlyMetLeuGlySerLysLysGlyValAsnLeuProGlyAlaAla
TTAAGGAGAAAGGCAAGGACTTTGTCATGACTGAGGTTGAGAACGGTGGCATGCTTGGTAGTAAGAAGGGAGTGAACCTCCCAGGTGCTG   720

  ValAspLeuProAlaValSerGluLysAspIleGlnAspLeuLysPheGlyValGluGlnAsnValAspMetValPheAlaSerPheIle
CGGTCGACCTGCCTGCAGTCTCAGAGAAGGACATTCAGGACCTGAAATTTGGCGTGGAGCAGAATGTGGACATGGTGTTCGCTTCCTTCA   810

  ArgLysAlaAlaAspValHisAlaValArgLysValLeuGlyGluLysGlyLysHisIleLysIleIleSerLysIleGluAsnHisGlu
TCCGCAAAGCTGCTGATGTCCATGCTGTCAGGAAGGTGCTAGGGGAAAAGGGAAAGCACATCAAGATTATCAGCAAGATTGAGAATCACG   900

  GlyValArgArgPheAspGluIleMetGluAlaSerAspGlyIleMetValAlaArgGlyAspLeuGlyIleGluIleProAlaGluLys
AGGGTGTGCGCAGGTTTGATGAGATCATGGAGGCCAGCGATGGCATTATGGTGGCCCGTGGTGACCTGGGTATTGAGATCCCTGCTGAAA   990

  ValPheLeuAlaGlnLysMetMetIleGlyArgCysAsnArgAlaGlyLysProIleIleCysAlaThrGlnMetLeuGluSerMetIle
AAGTCTTCCTCGCACAGAAGATGATGATTGGGCGCTGCAACAGGGCTGGCAAACCCATCATTTGTGCCACTCAGATGTTGGAAAGCATGA  1080

  LysLysProArgProThrArgAlaGluGlySerAspValAlaAsnAlaValLeuAspGlyAlaAspCysIleMetLeuSerGlyGluThr
TCAAGAAACCTCGCCCGACCCGCGCTGAGGGCAGTGATGTTGCCAATGCAGTTCTGGATGGAGCAGACTGCATCATGCTGTCTGGGGAGA  1170

  AlaLysGlyAspTyrProLeuGluAlaValArgMetGlnHisAlaIleAlaArgGluAlaGluAlaAlaMetPheHisArgGlnGlnPhe
CCGCCAAGGGAGACTACCCACTGGAGGCTGTGCGCATGCAGCACGCTATTGCTCGTGAGGCTGAGGCCGCAATGTTCCATCGTCAGCAGT  1260

  GluGluIleLeuArgHisSerValHisHisArgGluProAlaAspAlaMetAlaAlaGlyAlaValGluAlaSerPheLysCysLeuAla
TTGAAGAAATCTTACGCCACAGTGTACACCACAGGGAGCCTGCTGATGCCATGGCAGCAGGCGCGGTGGAGGCCTCCTTTAAGTGCTTAG  1350

  AlaAlaLeuIleValMetThrGluSerGlyArgSerAlaHisLeuValSerArgTyrArgProArgAlaProIleIleAlaValThrArg
CAGCAGCTCTGATAGTTATGACCGAGTCTGGCAGGTCTGCACACCTGGTGTCCCGGTACCGCCCGCGGGCTCCCATCATCGCCGTCACCC  1440

  AsnAspGlnThrAlaArgGlnAlaHisLeuTyrArgGlyValPheProValLeuCysLysGlnProAlaHisAspAlaTrpAlaGluAsp
GCAATGACCAAACAGCACGCCAGGCACACCTGTACCGCGGCGTCTTCCCCGTGCTGTGCAAGCAGCCGGCCCACGATGCCTGGGCAGAGG  1530

  ValAspLeuArgValAsnLeuGlyMetAsnValGlyLysAlaArgGlyPhePheLysThrGlyAspLeuValIleValLeuThrGlyTrp
ATGTGGATCTCCGTGTGAACCTGGGCATGAATGTCGGCAAAGCCCGTGGATTCTTCAAGACCGGGGACCTGGTGATCGTGCTGACGGGCT  1620

  ArgProGlySerGlyTyrThrAsnThrMetArgValValProValPro***
GGCGCCCCGGCTCCGGCTACACCAACACCATGCGGGTGGTGCCCGTGCCATGAGCTGCCGGCGCCCCCGCCTCTTCCACCCCCGCGCCCC  1710

TTCCCCATGCATTAGGCCAGCAGTCGCTTGCGATGTTCTCCTCGTCCCTAGCGTGGATTTAGGTTGCTCAACACCACCGAGTGCAGCAGC  1800

GGGGGAAGAGACCTTAAATCACTCAGAATACTTGAAGTGTGTAGTTGTTTCCAGACCTCTCTCCCTGTAGTTAGCCATGTCAGAGCGGGG  1890

CCCTGTGCTCCCCACGTCCCCCTAGCTTAGCACTGCCCTCAGGGTGTGGCTGTGCCTCTCCTTGCCCCGGCTGTGGGGTGATGGGGAAGG  1980

AGTGGGCGCTTATGGTAACGGTTCTGCTGCTGCACACCGCAGCTCCCCCAGGTTGGTGCCCTCTGTCCATGAGGGGTTATAGGTTGTCCC  2070

CTCTGGGCTTACACTCCAAAGGTGTGCTCCACGGTTCCTACAAGAATCCCAAACCAGGGTGCTGCTCGTTGTCCCCCTGGGTGCAGGTGG  2160

AGTCCTGCGGGTTGTTGGGGTGCAGGATGGGTGGGGGACGCACGCTTCATCTCCATGCAGCCACCTGCCCTTCTATGGTGTCTGTCTGTG  2250

TTGTCGCCTTTCCTGAGAGCGTTCCTTTTGTCTGGGCGCCCTGTTTGCTCCGGTGGGTGTATTGGGTTGTGGAGCACCACTGTGCACGTC  2340

AATAAAGAGCTGAG
```

FIG. 3. cDNA sequence of chicken muscle PK mRNA. The predicted amino acid sequence of the mature protein is displayed above the nucleotide sequence. We assume that translation begins at the first available AUG codon and terminates at the first stop codon encountered in the reading frame. A potential polyadenylylation signal, A-A-T-A-A-A, occurs at base 2,341. The primer extension product, which was analyzed to determine the 5′ end of the RNA, was heterogeneous at its 3′ end; therefore, the sequence presented here begins with either the third or the fourth reverse-transcribable nucleotide of the message.

Fig. 3) on a polyacrylamide gel (9), hybridized the complementary strand to poly(A)$^+$ RNA, extended it with reverse transcriptase, rendered it double stranded, and cloned it into pBR322. We screened 600 colonies by using the 5′-end-labeled primer as a probe and found 1 positive colony. This clone, pPK201, contained 78 bp of sequence overlapping pPK102, along with 875 bp of upstream sequence.

Determination of the 5′ Sequence. We primed cDNA synthesis out to the 5′ end of the message by using an end-labeled fragment of pPK201 and determined a further 53 nucleotides of 5′ nontranslated sequence not included in the cloned cDNA inserts. However, 3′ heterogeneity of the primer extension product obscured the determination of the first two or three nucleotides. Fig. 2 summarizes the strategy for the sequence determination.

Primary Structure of the PK Message. Fig. 3 shows the cDNA sequence of PK mRNA and its derived amino acid sequence. An 80-nucleotide 5′ nontranslated region is followed by an open reading frame of 1,590 nucleotides, including the initiation codon. The last 684 nucleotides are 3′ nontranslated sequence; a possible polyadenylylation signal, A-A-U-A-A-A, occurs at nucleotide 2,341.

The NH_2-terminal residues of PK isolated from bovine muscle, rabbit muscle, rat liver, and human kidney have been shown to be modified; the absence of a free amino group blocks the standard dansyl-Edman reaction (20–23). The modified residue is an α-amino-acetylated serine for the rabbit and bovine muscle enzymes (20, 24). This same modification appears in tobacco mosaic virus coat protein, in which the initiator methionine is removed and the adjacent serine is acetylated (25, 26). The residue following the first AUG in our sequence is a serine; we propose that this is the first amino acid of the mature enzyme. The molecular weight of chicken muscle PK estimated from the cDNA sequence is 57,865. Sedimentation equilibrium of the purified protein has given an estimated monomer molecular weight of 53,000 (27).

Comparison with Other Muscle PK Sequences. Comparison of the predicted amino acid sequence with reported sequences of peptides isolated from rabbit, sturgeon, and bovine muscle PK reveals considerable cross-species homology. Related peptides isolated from cyanogen bromide-treated rabbit and sturgeon muscle PK (28) are identical to the corresponding chicken PK sequence (amino acids 310–329; nucleotides 1,011–1,067 in Fig. 3) at 84% and 83% of the residues; a 34-amino acid peptide isolated from trypsinized bovine muscle PK, containing a reactive lysine group (29), is completely homologous to the corresponding region of the chicken sequence (amino acid residues 341–374; nucleotides 1,104–1,205 in Fig. 3).

Comparison with the Yeast Enzyme. We aligned the chicken muscle amino acid sequence with the recently reported sequence of yeast PK (30), minimizing the number of insertions and maximizing the number of identities and conserved substitutions. The chicken enzyme is longer than the yeast enzyme by 22 amino acid residues at the NH_2 terminus and a single residue at the COOH terminus; a total of 8 additional residues must be inserted at three sites into the yeast sequence and 1 residue into the chicken sequence in order to align the two proteins (Fig. 4). If these are counted as mismatches, the overall homology between the two enzymes is approximately 45%.

The distribution of the conserved regions reflects roughly the three-dimensional structure of cat muscle PK—with regions of low homology, and positions of insertions or deletions, coinciding with sequences that connect different domains at the cat enzyme. The structure of cat PK has been determined at a resolution sufficient to trace the folding of the polypeptide chain (2). Each of the four identical subunits is divided into three domains: the central A domain, comprising approximately half the mass of the protein, flanked by the smaller B and C domains. The A domain is a symmetric, eight-stranded, parallel β barrel; adjacent β strands are connected by α helices which form a larger cylinder of parallel helices concentric with the β barrel. The course of the polypeptide chain around the barrel is interrupted between the third β sheet and the third α helix by the B domain; the chain then completes the barrel and forms the COOH-terminal C domain. The correspondence between the primary and tertiary structures of PK is then A_1–B–A_2–C; A_1 and A_2 refer to the two components of the A domain. Fig. 4 gives the approximate locations of the domains with respect to the amino acid sequence; a more precise alignment will be possible after refinement of the structure of the cat enzyme to determine the positions of the amino acid residues. The peaks of highest homology in domain A probably represent residues

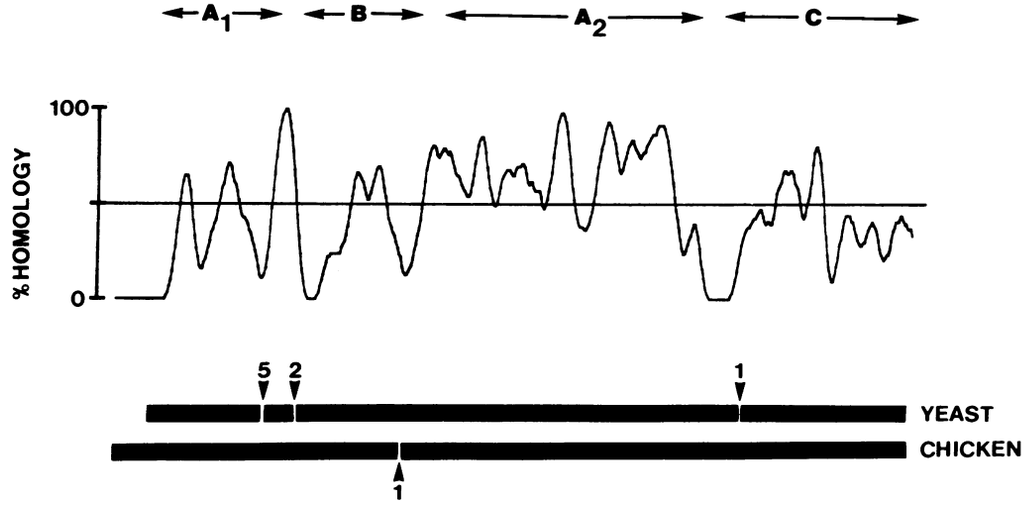

FIG. 4. Comparison of the primary structures of yeast and chicken PK. The percentage homology at each residue is a weighted average over a 15-residue window; the weighting factor is $(8 - d)/64$, in which d is the distance from the center of the window. Only exact matches are scored; insertions are counted as mismatches. Below the homology plot, arrows indicate the locations and lengths (in amino acid residues) of the minimum number of insertions required to maximize identities or conserved substitutions. The approximate locations of the structural domains are indicated above the plot (see text).

that coalesce to form the ADP and phospho*enol*pyruvate binding sites on the surface of the barrel.

Comparison with the Mammalian Liver Isozyme. At least four isozymes of PK have been reported in mammals: the skeletal muscle form (M type), the fetal or kidney form (K type), the liver form (L type), and the erythrocyte form (R type) (31). Genetic studies suggest two structural genes: one coding for the M and K types and one coding for the L and R types (32, 33). Only two isozymes have been observed in chickens: a fetal form and an adult skeletal muscle form (discussed in this paper), corresponding to the mammalian K and M types, respectively (18, 27).

To prevent a futile loop from occurring during gluconeogenesis, the mammalian liver PK isozyme is regulated by a cAMP-dependent protein kinase. It exists in two forms—phosphorylated and unphosphorylated—with the latter being more active (34). The sequence surrounding the site of phosphorylation, Glu-Gly-Pro-Ala-Gly-Tyr-Leu-Arg-Arg-Ala-Ser(P)-Leu-Ala-Gln-Leu-Thr-Gln-Glu-Leu-Gly*-Thr*-Ala*-Phe*-Phe-Gln*-Arg-Gln*-Gln*-Leu*-Pro-Ala*-Ala*-Met*, was determined by Humble (35) from a 33-amino acid cyanogen bromide fragment isolated from the pig liver enzyme. A similar sequence has been reported for rat liver PK (35). This sequence can be aligned with the NH_2 terminus of the chicken muscle enzyme (amino acids 7–20; nucleotides 102–143 in Fig. 3), matching 11 of the last 14 residues (marked with asterisks); the phosphorylated serine of the liver isozyme is three amino acids upstream from the residue corresponding to the first amino acid of the muscle enzyme. Apparently, the two forms differ in length at their NH_2 terminus, with the L type possessing a regulatory "domain" missing from the muscle form. Phosphorylation might cause this domain to interact with the rest of the enzyme so as to decrease its activity. Chymotrypsin removes the phosphorylated site of porcine liver PK without destroying its catalytic activity (35), consistent with this model.

Developmental Regulation. Fig. 5 shows that PK message increases dramatically between 3 and 126 days after hatching. This confirms the inference that we made from *in vitro* translation experiments and shows that the developmental increase in enzyme activity (18) is due to an increase in mRNA. Fig. 5 also demonstrates that, within the resolution of the gel, PK is translated from a single mRNA species in adult chicken muscle.

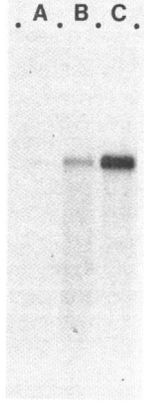

FIG. 5. RNA blot hybridization showing a developmental increase in the level of PK message. Poly(A)$^+$ RNA isolated from 3-, 15-, and 126-day-old chickens (1.6 μg each in lanes A, B, and C, respectively) was separated on a 1.5% agarose/6% formaldehyde gel, blotted, and hybridized to ^{32}P-labeled pPK102 cDNA insert.

We thank Donald Straus for multiple contributions to this project. We also thank Hilary Muirhead, Linda Fothergill, Emil Schiltz, and Tom Alber for communicating results prior to publication, David Horowitz for the plotting program used to generate Fig. 4, and Lydia Villa-Komaroff for advice on the construction and screening of cDNA libraries. We are particularly indebted to Deborah Hurt for the generous gift of purified chicken muscle PK and for discussions, both of which made this project possible. N.L. was supported, in part, by a National Science Foundation Predoctoral Fellowship. This work was supported by National Institutes of Health Grant GM09541-20.

1. Kayne, F. J. (1974) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 8, pp. 353–382.
2. Stuart, D. I., Levine, M., Muirhead, H. & Stammers, D. K. (1979) *J. Mol. Biol.* **134**, 109–142.
3. Lomedico, P. T. & Saunders, G. F. (1976) *Nucleic Acids Res.* **3**, 381–391.
4. Pelham, H. R. B. & Jackson, R. J. (1976) *Eur. J. Biochem.* **67**, 247–256.
5. Laemmli, U. K. & Favre, M. (1973) *J. Mol. Biol.* **80**, 575–599.
6. Kessler, S. W. (1975) *J. Immunol.* **115**, 1617–1624.
7. Baily, J. M. & Davidson, N. (1976) *Anal. Biochem.* **70**, 75–85.
8. Wickens, M. P., Buell, G. N. & Schimke, R. T. (1978) *J. Biol. Chem.* **253**, 2483–2495.
9. Maxam, A. M. & Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560.
10. Rouchoudhry, R., Jay, E. & Wu, R. (1976) *Nucleic Acids Res.* **3**, 863–877.
11. Land, H., Grez, M., Hauser, H., Lindenmaier, W. & Schutz, G. (1981) *Nucleic Acids Res.* **9**, 2251–2266.
12. Villa-Komaroff, L., Efstradiadis, A., Broome, S., Lomedico, P., Tizard, R., Naber, S. P., Chick, W. L. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3727–3731.
13. Mandel, M. & Higa, A. (1970) *J. Mol. Biol.* **53**, 159–162.
14. Grunstein, M. & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3961–3965.
15. Lehrach, H., Diamond, D., Wozney, J. M. & Boedtker, H. (1977) *Biochemistry* **16**, 4743–4751.
16. Rigby, W. J., Diekmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
17. Alwine, J. C., Kemp, D. J. & Stark, G. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5350–5354.
18. Cardenas, J. M., Bandman, E., Walker, C. & Strohman, R. C. (1979) *Dev. Biol.* **68**, 326–333.
19. Goldberg, M. L., Lifton, R. P., Stark, G. R. & Williams, J. G. (1979) *Methods Enzymol.* **68**, 206–220.
20. Johnson, S. C. (1980) Dissertation (Oregon State Univ., Corvallis, OR).
21. Cottam, G. L., Hollenberg, P. F. & Coon, M. J. (1969) *J. Biol. Chem.* **224**, 4363–4365.
22. Hall, E. R., Kohl, E. A. & Cottam, G. L. (1978) *Biochem. Biophys. Res. Commun.* **80**, 586–592.
23. Harkins, R. N., Black, J. A. & Rittenberg, M. B. (1977) *Biochemistry* **16**, 3831–3837.
24. Brummel, N. C., Boaz, D. P., Flashner, M. & Stegink, L. D. (1976) *Biochim. Biophys. Acta* **438**, 102–107.
25. Anderer, F. A., Uhlig, H., Weber, E. & Schramm, G. (1960) *Nature (London)* **186**, 922–925.
26. Guilley, H., Jonard, G., Kukla, B. & Richards, K. E. (1979) *Nucleic Acids Res.* **6**, 1287–1308.
27. Cardenas, J. M., Blachly, E. G., Ceccotti, P. L. & Dyson, R. D. (1975) *Biochemistry* **14**, 2247–2252.
28. Anderson, P. J. & Randall, R. F. (1975) *Biochem. J.* **145**, 575–579.
29. Johnson, S. C., Baily, T., Becker, R. R. & Cardenas, J. M. (1979) *Biochem. Biophys. Res. Commun.* **90**, 525–530.
30. Burke, R. L., Tekamp-Olson, P. & Najarian, R. (1983) *J. Biol. Chem.* **258**, 2193–2201.
31. Ibsen, K. H. (1977) *Cancer Res.* **37**, 341–353.
32. Peters, J., Nash, H. R., Eicher, E. M. & Bulfield, G. (1981) *Biochem. Genet.* **19**, 757–769.
33. Moore, K. J. & Bulfield, G. (1981) *Biochem. Genet.* **19**, 771–781.
34. Riou, J. P., Claus, T. H. & Pikas, S. J. (1978) *J. Biol. Chem.* **253**, 656–659.
35. Humble, E. (1980) *Biochim. Biophys. Acta* **626**, 179–187.