Zoological Studies **38**(3): 322-332 (1999)

Zoological Studies

Complementary DNA Cloning and Analysis of Gene Structure of Pyruvate Kinase from *Drosophila melanogaster*

Yi-Chih Chien*, Yu-Jing Zhu and Chun-Mei Chuen

Department of Biology, National Changhua University of Education, Changhua, Taiwan 500, R.O.C.

(Accepted May 19, 1999)

Yi-Chih Chien, Yu-Jing Zhu and Chun-Mei Chuen (1999) Complementary DNA cloning and analysis of gene structure of pyruvate kinase from *Drosophila melanogaster*. *Zoological Studies* **38**(3): 322-332. Screening a λ-gt10 cDNA library of *Drosophila melanogaster* using a human pyruvate kinase (*Pyk*) cDNA clone (pCJ11) and a rat pituitary *Pyk* cDNA clone (pCJ22) as probes, we isolated 2 cDNA clones, cDMPK15 and cDMPK06. Complete nucleotide sequencing of the two cDNA clones (GenBank AF061507) revealed that they encompassed the coding region of pyruvate kinase cDNA (1602 bp; 533 amino acids + TAA) flanked by a 5' untranslated region of 240 bp and a 3' untranslated region of 253 bp. An alignment of the deduced amino acid sequence from the *Pyk* cDNA clones with those of PK from other organisms indicated that the amino acid residues constituting the active sites have been highly conserved. In addition, the overall positional identity between the sequence of the "*Drosophila*" enzyme and those from other sources was 42%-63%.

Polymerase chain reaction was applied to amplify the genomic DNA fragments from the *Pyk* gene of *D. melanogaster*. These overlapping amplicons, which covered the complete coding region of *Pyk*, were further sequenced using cycle-sequencing with an ABI Prism 377 DNA sequencer. A total of 3447 bp of the nucleotide sequence (GenBank AF062478) was determined from these amplicons. By comparing these sequences with the sequence of *Pyk* cDNA clones isolated, 4 exons were identified of 282, 1390, 157, and 266 bp in length. The introns identified all contained the consensus 5'- and 3'-splicing sites (GT-AG). RT-PCR analysis was performed to determine the number of species of the *Pyk* transcript in adults of *D. melanogaster*. The observation that only a single amplicon appeared in each amplification suggests that a single *Pyk* transcript is expressed in adults of *D. melanogaster*, and might imply that there is only 1 *Pyk* gene in *D. melanogaster*.

Key words: Pyruvate kinase, cDNA, Cloning, RT-PCR, *Drosophila melanogaster*.

Pyruvate kinase (ATP: pyruvate 2-O-phosphotransferase, EC 2.7.1.40) is a key enzyme in the glycolytic pathway, which is an almost universal central pathway for glucose catabolism, not only in animals and plants, but also in a great number of microorganisms. It catalyzes the conversion of phosphoenolpyruvate into pyruvate by the addition of a proton and the loss of a phosphate group, which is transferred into ADP. The reaction is essentially irreversible in favor of pyruvate and ATP formation, and requires both bivalent and monovalent cations (Mg^{+2} and K^{+}).

James and Collier (1992) proposed that arginine kinase (AK) and pyruvate kinase (PK) are coordinately regulated by the 75B "early" gene in *Drosophila* imaginal discs. This hypothesis is based upon the following observations. First, detailed developmental profiles of arginine kinase and pyruvate kinase (James and Collier 1988) revealed that the specific activities of both enzymes synchronously increased and reached a peak at the prepupal stage. The increased activities of both enzymes in the prepupal stage is in part due to the increased activities of both enzymes in the imaginal discs. Second, the prepupal peaks in both AK and PK activities disappeared when temperature-sensitive ecdysteroneless mutant, *ecd-1* (Garen et al. 1977) larvae were held at a restrictive temperature (29 °C). However, when mutant larvae were shifted into a permissive temperature (20 °C) after 60 h at the restrictive

*To whom correspondence and reprint requests should be addressed. Tel: 886-4-7232105 ext. 3411. Fax: 886-4-7611156. E-mail: chien@cc.ncue.edu.tw

temperature, pupariation occurred, and the prepupal peak for both enzymes was observed (James and Collier 1990). Third, both AK (James and Collier 1992) and PK expressions were affected by altered doses of the ecdysterone-regulated 75B "early" gene.

The purpose of the coordinated regulation of PK and AK in imaginal discs is still unknown. However, it has been suggested that both PK and AK provide energy for the eversion of the imaginal discs, since they both catalyze a metabolic reaction for releasing ATP. A good understanding of the coordinated regulation of PK and AK expression requires molecular analyses of the pyruvate kinase (*Pyk*) and arginine kinase (*Ark*) genes and their associated regulatory regions. A prerequisite of any such molecular analyses is the cloning of both *Pyk* and *Ark* genes.

In this work, we isolated cDNA clones for pyruvate kinase from *D. melanogaster* and determined the primary structure of this protein from the cDNA sequence. Furthermore, using polymerase chain reaction, we were able to amplify and sequence 9 overlapping genomic DNA fragments which span the transcribed region of the pyruvate kinase gene. By comparisons of the sequences of cDNA clones and genomic DNA fragments, the structural relationship of the protein and gene of pyruvate kinase was clarified. The transcription start point (TSP) has been identified by the method of 5'RACE. Finally, it is suggested that, unlike the presence of multiple genes in mammals, there might be a single *Pyk* gene in *D. melanogaster*.

MATERIALS AND METHODS

Drosophila stocks

A wild-type stock of *Drosophila melanogaster* (Canton-S) used in this study was maintained on standard cornmeal medium (Strickberger 1962) at 24 °C. Adult flies were collected for the experiments.

General methods

Preparation of phage and plasmid DNA, electrophoresis of DNA on agarose gel, restriction enzyme digestion, DNA ligation, transformation, Southern blot analysis, rescue of single-stranded DNA from cDNA clones, and further sequencing reactions were carried out using standard procedures as described by Sambrook et al. (1989). Hybridization probes were labeled with [α-^{32}P]dCTP (800 Ci/mmol) using random oligo-primer labeling (BioRad).

cDNA cloning

We screened a λ-gt10 *Drosophila* cDNA library using a plaque hybridization procedure (Sambrook et al. 1989). This library was derived from *Drosophila* imaginal disc mRNA kindly provided by G. Rubin (pers. comm.). Approximately 5 x 10^{10} phages from the λ-gt^{10} library were screened with oligo-labeled inserts from human and rat cDNA clones, pCJ11 and pCJ22, respectively (kindly provided by Dr. Cheng, National Cancer Research, Bethesda, MD). Several independent clones were identified and characterized using polymerase chain reactions and Southern blot analysis. One of them, λPK22, was further analyzed in this study.

For sequencing purposes, the inserts of λPK22 were subcloned into bluescript phagemids (SK$^+$ or KS$^+$).

Polymerase chain reaction (PCR) for cDNA clones

The PCR primers for cDNA clones were designed by comparing the published amino acid sequences for human, chicken, yeast, and *E. coli* pyruvate kinase to identify highly conserved regions. Degenerate primers were synthesized by Operon Technologies, Inc. (Alameda, Ca, U.S.A.). The sequences of the primers were 5'-CTNGAYACNAARGGNCCNGARATN-3' and 5'-CCNAGRTCNCCNCGNGCNACCAT-3', where N represents the 4 nucleotides (A, T, C, G); R represents purines (A, G); and Y represents pyrimidines (T, C).

Amplification was performed in a GeneAmp PCR System 960 thermal cycler (Perkin-Elmer/Cetus) in 50 or 100 ml of solution containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, each dNTP at 1 mM, each primer at 1 mM, 100-1000 ng template DNA, and 2.5 units of *Taq* DNA polymerase (Perkin-Elmer/Cetus). The DNA was preheated to 95 °C for 10 min before being added to the reaction mix. Each cycle of the polymerase chain reaction consisted of denaturation for 1 min at 94 °C, annealing for 1 min at 58 °C, and extension for 1 min at 72 °C, with a final extension time of 3 min. This cycle was repeated 30 times. The degree of amplification was determined by separation of the amplicons on a 2% SeaKem agarose minigel (1x TBE) at 100 V for 1 h.

Polymerase chain reaction (PCR) of genomic DNA

Extraction of genomic DNA from adult flies was

performed according to the protocol of the DNA Isolation Kit of Puregene D-5500 (Gentra Systems). Eleven oligonucleotide primers were used to amplify the DNA fragments of the pyruvate kinase gene from the genomic DNA of *D. melanogaster*. The primers were designed according to the cDNA sequence for *D. melanogaster* pyruvate kinase. The sequences of primers are listed in table 1, and their positions relative to the *Pyk* gene are shown in figure 1.

Amplification of DNA fragments was performed in a GeneAmp PCR System 2400 or 9600 thermal cycler (Perkin-Elmer/Cetus) in 50 µl of solution containing 5 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.01 mM EDTA, 0.1 mM DTT, 5% glycerol, 0.1% Triton, 1.5 mM $MgCl_2$, 0.25 mM dNTP, 10 picomoles of primers, 25-125 ng of genomic DNA, and 1 unit of *Tag* DNA polymerase (Promega). The DNA solution was pre-heated to 95 °C for 5 min. The temperature profile was 30 cycles of 45 s at 95 °C, 45 s at 50-59 °C, and 1 min at 72 °C, with a final extension for 10 min at 72 °C. The degree of amplification was determined by separation of the amplicons on a 1. 5% agarose gel (in 0.5X TBE) at 100 V for 40-50 min. The amplicons were purified using a GENECLEAN III Kit (BIO 101) and were further used in the sequencing reaction process.

5' RACE analysis

Poly-A mRNA was isolated according to the method provided by GibcoBRL (Life Technologies, Inc.). First-strand cDNA was reversed transcribed from 50 ng of mRNA by SuperScript™ reversed transcriptase (200 units) using PK1 as primer (Table 1; Fig. 1). A homopolymeric tail was added to the 3'

Table 1. Sequences of primers used in PCR and sequencing

Primer	5' → 3'
PK1	CTG GAC ACA AAG GGG CCC GAG ATC CGT
PKp1	TCC GAG AAG GAT AAG AGC GAT CTG
PKp2	GCT GCG TTG TGT GTG TTG TTG TGC
PKp3	TCA TCG CGG TCA CCG TTT CG
PKp4	TGC TGA TTG CGT CAT GTT GTC TGG
PKp9	AAC ATG GCC GCA GGA GCT GAT ACC
PKp11	CGT TCG GTA GGC AAC AAA TGG

Primer	5' → 3'
PK2	ACC CAG ATC TCC ACG GGC CAC CAT
PKm1	CAG ATC GCT CTT ATC CTT CTC GGA
PKm2	TAG TTT GTG TGT CTG TTT GAG GCC C
PKm3	GCG ATC CGT GAG AGA AGT TCA TGC G
PKm4	GTT ACC AAC GGG AAT CGC ACT TGG G
PKm5	GAT GAA GCC GTT CTT CTT TCC GAC C
PKm12	AGA CAA CAT GAC GCA ATC AGC ACC

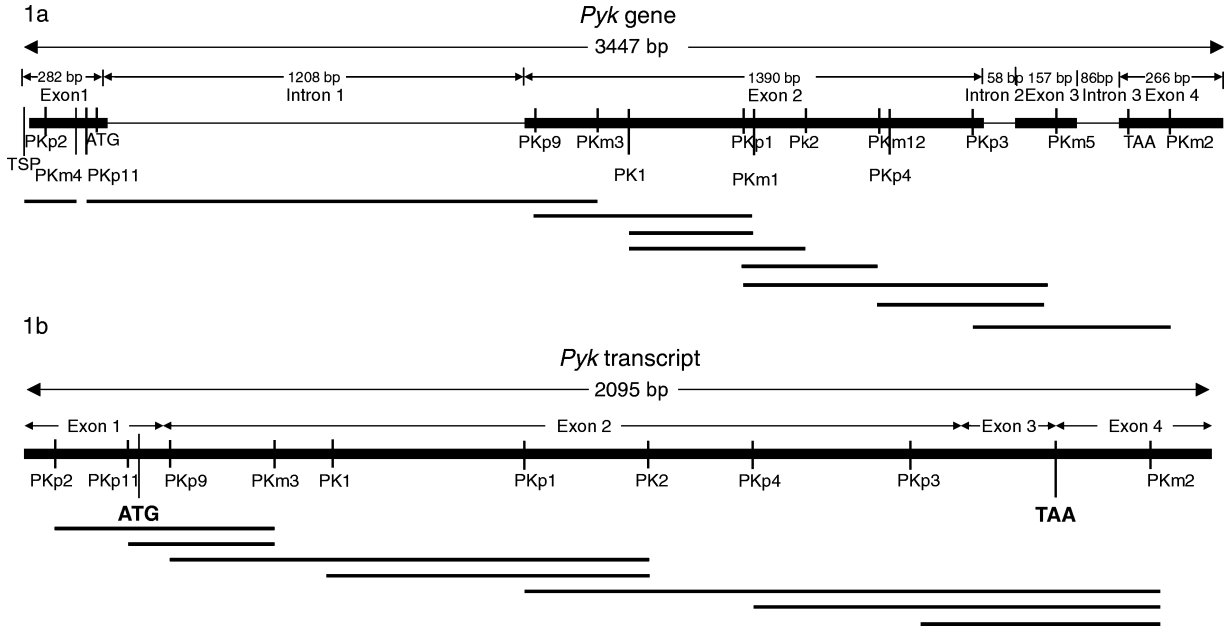

Fig. 1a. Schematic representation of the pyruvate kinase gene of *Drosophila melanogaster* that includes the primers used for PCR and the resulting amplicons. TSP represents the transcription start point as determined by 5' RACE.
Fig. 1b. Schematic representation of the pyruvate kinase transcript of *D. melanogaster* that includes the primers used for RT-PCR and the resulting amplicons.

end of the cDNA by terminal deoxynucleotidyl transferase (15 units) and dCTP. PCR amplification was then performed using PKm4 and the abridged anchor primer (5'-CUACUACUACUAGGCCACGCGTCGACTAGTACGGGIIGGGIIGGGIIG-3'), which was purchased from GibcoBRL (Life Technologies, Inc). The resulting amplicon was then sequenced to identify the transcription start point of the *Pyk* gene.

RT-PCR analysis

To identify the species of the *Pyk* transcript of *D. melanogaster*, 50 ng of mRNA from adult flies was reversed transcribed with SuperScript™ reversed transcriptase (200 units) using *Pyk*-specific PK2 or PKm2 as primers (Table 1; Fig. 1). The resulting cDNAs were amplified following the protocol mentioned above. The positions of primers relative to the *Pyk* transcript and expected amplicons are shown in figure 2. The amplicons were visualized on agarose gels and further sequenced.

DNA sequencing

Nine overlapping amplicons, which span the region of the pyruvate kinase gene from *D. melanogaster* as well as that from RT-PCR, were further sequenced (Fig. 1). All fragments were sequenced on both strands more than twice. The sequencing reactions were performed with a dye-labeled terminators kit (Perkin-Elmer) using the cycle sequencing method. Both separation of DNA fragments and sequence analysis were performed in an ABI Prism 377 DNA sequencer.

RESULTS

Screening of λ-gt10 cDNA library with human PK cDNA clone (PCJ11) and rat pituitary PK cDNA clone (PCJ22)

Five plates, each containing ca. 10^{10} plaques, were blotted onto duplicate nitrocellulose circles. One set of circles was hybridized with the 2.2-kb insert of pCJ11 and the other set of duplicate circles was hybridized with the 1.8-kb insert of pCJ22 at low stringency. From this screening, 6 agar plugs were positive on both duplicate circles.

Before secondary screening and plaque purification, phage lysates were prepared from the 6 agar plugs, from which phage DNA was isolated. The 6 pools of phage DNA, designated λC1 to λC6, were subjected to polymerase chain reactions using the 2 degenerate primers mentioned in "Materials and Methods". If *Pyk* cDNA was included in the pools of phage DNA, an amplicon of ca. 500-600 bp should be amplified. An amplicon was obtained in 4 of 6 pools of phage DNA, and the sizes of the amplicons were similar to that of those "controls" (pCJ11 and pCJ22) (data not shown). This result suggested that 4 pools of phage DNA contained *Pyk* cDNA; while the other 2 pools of phage DNA might not contain any *Pyk* cDNA or might contain only a portion of *Pyk* cDNA.

Furthermore, the 6 pools of phage DNA were digested with *Eco* R1, blotted, and hybridized with the insert of pCJ22. As a result, all 6 pools of phage DNA contained inserts that can hybridize with the probes (data not shown). λC2 contained the largest insert (ca. 2 kb) which hybridized with the probe, so it was chosen for further screening. The failure of λC5 and λC6 to amplify during PCR might be the consequence of the small inserts contained in these cDNA clones.

λC2 was further screened and plaque-purified, and 2 positive plaques, named λPK21 and λPK22, were selected for further analysis. DNA was isolated from these 2 phages, and was digested with *Eco* R1 to release the inserts. λPK21 and λPK22 appeared to be the same as they both contained 2 *Eco* R1 insert fragments, one ca. 1.5 kb and the other ca. 600 bp (data not shown).

Sequencing of the cDNA subclones: cDRPK15 and cDRPK06

For sequencing purposes, the inserts of λPK22 were subcloned into bluescript phagemids (SK^+ or KS^+). The resulting cDNA subclones were named cDMPK15 (containing a 1.5-kb insert) and cDMPK06 (containing a 0.6-kb insert). By further subcloning the available restricted fragments of the 2 cDNA fragments, the total sequences of the 2 cDNA fragments were completed.

Figure 2 shows the complete nucleotide sequence (GenBank AF061507) determined from the cDNA clones and its deduced amino acid sequence. We assumed that translation began at the 1st available d(ATG) codon and terminated at the 1st stop codon d(TAA) encountered in the reading frame. The nucleotide sequence of the 2 cDNA fragments indicated that they encompassed complete coding region of pyruvate kinase cDNA (1602 bp; 533 amino acids + TAA), flanked by a 5' untranslated region of 240 bp, and a 3' untranslated region of 253 bp (Figs. 3, 4). The polyadenylation signal, d(AATAAA), was found near the poly(A) sequence. An *Eco* R1 site,

d(GAATTC), was located at amino acids 150 (Glu) and 151 (Phe) as expected. This site was further confirmed by sequencing genomic DNA. The molecular weight of *Drosophila* PK estimated from the cDNA sequence is 58 630 Da, which is slightly smaller than the 59 000 Da estimated from purified PK protein on SDS-PAGE (data not shown).

Alignment of the deduced amino acid sequence from the *Drosophila Pyk* cDNA clones with those of PK from other organisms

Figure 3 shows the alignment of the deduced amino acid sequence from the *Drosophila Pyk* cDNA clones with those of PK from other organisms. The

```
-240                                                                                                  TCGAACTTTTTGTCTCGTCCTCGCCGCGTCCGATTT  -205
-204 TCCATATATCCCGATTTTCGTCTCAAACTTGGAAGATATATTGCCAAAATATTGCATTCGGCTAGCCGCTGCGTTGTGTGTGTTGTTGTGCGTGGAGTTGCG  -103
-102 TGTGCAGTTGCAAGAAGATATAGTTGTCTATATCGCCAGACTGATCTCGATCCCAAGTGCGATTCCCGTTGGTAACATACTGCCCGTTCGGTAGGCAACAAA    -1
   1 ATG GTG AAC GTA ACC ATT TAC GAT GAG GCA CCG CAG CTG AAG CCC AAC GAA GTA CCC CAA AAC ATG GCC GCA GGA GCT    78
   1  M   V   N   V   T   I   Y   D   E   A   P   Q   L   K   P   N   E   V   P   Q   N   M   A   A   G   A    26
  79 GAT ACC CAA CTG GAG CAC ATG TGC CGT CTG CAG TTC GAC TCG CCA GTG CCC CAT GTG CGT CTG TCC GGA ATC GTG TGC   156
  27  D   T   Q   L   E   H   M   C   R   L   Q   F   D   S   P   V   P   H   V   R   L   S   G   I   V   C    52
 157 ACC ATC GGA CCT GCC TCC AGC AGC GTG GAG ATG CTG GAG AAG ATG ATG GCC ACC GGC ATG AAC ATC GCG CGC ATG AAC   234
  53  T   I   G   P   A   S   S   S   V   E   M   L   E   K   M   M   A   T   G   M   N   I   A   R   M   N    78
 235 TTC TCT CAC GGA TCG CAC GAG TAC CAT GCC GCC ACC GTG GCC AAT GTG CGC CAG GCG GTG AAA AAC TAC TCG GCC AAG   312
  79  F   S   H   G   S   H   E   Y   H   A   A   T   V   A   N   V   R   Q   A   V   K   N   Y   S   A   K   104
 313 CTG GGC TAC GAA CAC CCC GTG GCC ATT GCC CTG GAC ACC AAG GGG CCC GAG ATC CGT ACC GGT CTG ATC GGA GGC AGC   390
 105  L   G   Y   K   H   P   V   A   I   A   L   D   T   K   G   P   E   I   R   T   G   L   I   G   G   S   130
 391 GGC ACC GCC GAG ATT GAG CTG AAG AAG GGC GAG AAG ATC AAG CTG ACC ACC AAC AAG GAA TTC CTG GAG AAG GGC TCT   468
 131  G   T   A   E   I   E   L   K   K   G   E   K   I   K   L   T   T   N   K   E   F   L   E   K   G   S   156
 469 CTG GAG ATT GTG TAC GTG GAC TAC GAG AAC ATT GTC AAT GTG GTG AAG CCC GGC AAC CGG GTG TTC GTC AAT GAC GGT   546
 157  L   E   I   V   Y   V   D   Y   E   N   I   V   N   V   V   K   P   G   N   R   V   F   V   N   D   G   182
 547 CTG ATC TCA CTG ATT GTC CGC GAG GTA GGC AAG GAT TCC CTC ACC TGC GAA GTG GAG AAC GGC GGC TCT CTG GGT TCC   624
 183  L   I   S   L   I   V   R   E   V   G   K   D   S   L   T   C   E   V   E   N   G   G   S   L   G   S   208
 625 CGC AAG GGT GTG AAC CTG CCA GGC GTG CCC GTC GAT CTG CCT GCC GTC TCC GAG AAG GAT AAG AGC GAT CTG CTG TTC   702
 209  R   K   G   V   N   L   P   G   V   P   V   D   L   P   A   V   S   E   K   D   K   S   D   L   L   F   234
 703 GGT GTG GAG CAG GAA GTG GAC ATG ATC TTT GCT TCG TTC ATC CGC AAC GCC GCT GCT TTG ACC GAG ATC CGT AAG GTT   780
 235  G   V   E   Q   E   V   D   M   I   F   A   S   F   I   R   N   A   A   A   L   T   E   I   R   K   V   260
 781 CTT GGC GAG AAG GGC AAG AAC ATC AAG ATC ATT TCC AAG ATC GAG AAC CAG CAG GGC ATG CAC AAC CTG GAC GAG ATC   858
 261  L   G   E   K   G   K   N   I   K   I   I   S   K   I   E   N   Q   Q   G   M   H   N   L   D   E   I   286
 859 ATC GAG GCC GGT GAT GGC ATT ATG GTA GCC CGT GGA GAT CTG GGT ATT GAG ATT CCC GCC GAG AAG GTG TTC CTC GCC   936
 287  I   E   A   G   D   G   I   M   V   A   R   G   D   L   G   I   E   I   P   A   E   K   V   F   L   A   312
 937 CAG AAG GCC ATG ATT GCC CGC TGC AAC AAG GCT GGC AAG CCT GTG ATC TGC GCC ACT CAG ATG TTG GAG TCA ATG GTG  1014
 313  Q   K   A   M   I   A   R   C   N   K   A   G   K   P   V   I   C   A   T   Q   M   L   E   S   M   V   338
1015 AAG AAG CCA CGT CCC ACT CGC GCT GAG ATC TCT GAT GTG GCC AAC GCT GTG CTC GAT GGT GCT GAT TGC GTC ATG TTG  1092
 339  K   K   P   R   P   T   R   A   E   I   S   D   V   A   N   A   V   L   D   G   A   D   C   V   M   L   364
1093 TCT GGT GAG ACC GCC AAG GGC GAG TAC CCG CTG GAG TGC GTC CTG ACC ATG GCC AAG ACC TGC AAG GAG GCC GAG GCT  1170
 365  S   G   E   T   A   K   G   E   Y   P   L   E   C   V   L   T   M   A   K   T   C   K   E   A   E   A   390
1171 GCC CTC TGG CAC CAG AAC TTC TTT AAC GAC TTG GTT CGC GGC GCT GGT ACC ATC GAT GCC TCT CAC GCG GCT GCC ATC  1248
 391  A   L   W   H   Q   N   F   F   N   D   L   V   R   G   A   G   T   I   D   A   S   H   A   A   A   I   416
1249 GCT GCC GTT GAG GCT GCC ACC AAG GCC AAG GCC TCC GCC ATC GTG GTG ATC ACC ACC AGC GGC AAG TCG GCC TTC CAG  1326
 417  A   A   V   E   A   A   T   K   A   K   A   S   A   I   V   V   I   T   T   S   G   K   S   A   F   Q   442
1327 GTG AGC AAG TAC CGC CCA CGC TGC CCC ATC ATC GCG GTC ACC CGT TTC GCG CAG ACC GCC CGA CAG GCC CAT CTC TAC  1404
 443  V   S   K   Y   R   P   R   C   P   I   I   A   V   T   R   F   A   Q   T   A   R   Q   A   H   L   Y   468
1405 CGT GGA CTG GTG CCA CTC ATC TAC AAG GAG CCC GGT CTT GGT GAC TGG CTG AAG GAC GTG GAC GTG CGC GTG CAG TTC  1482
 469  R   G   L   V   P   L   I   Y   K   E   P   G   L   G   D   W   L   K   D   V   D   V   R   V   Q   F   494
1483 GGT CTG CAG GTC GGA AAG AAG AAC GGC TTC ATC AAG ACC GGC GAT TCC GTC GTG GTG GTT ACC GGC TGG AAG CAG GGC  1560
 495  G   L   Q   V   G   K   K   N   G   F   I   K   T   G   D   S   V   V   V   V   T   G   W   K   Q   G   520
1561 TCC GGC TTC ACC AAC ACC ATC CGC ATT GTC ACC GTC GAA TAA                                              1602
 521  S   G   F   T   N   T   I   R   I   V   T   V   E  END                                              533
1603 ATTCTGCCGCTCAGGAAGTACCTGGATGCGTCGCAGCTGCCGCGGGCGTCGCGTTGATTACTGTTAATAACCAAGTTAATTTATTAATTACACAATATTCGG  1704
1705 AATATTGGAAACTGCAGGAGGCTGCTGGAACCGCTGTCAGAAGCGCCGAAGGCCGGCCCGGGCCTCAAACAGACACACAAATATACTATACCATTGGTGAAT  1806
1807 AATAGAGATTTCTTCGAAAATTATAATAAAGTCATCACTGTATTTATGT                                                    1855
```

Fig. 2. cDNA sequence and deduced amino acid sequence of pyruvate kinase from *Drosophila melanogaster*. The polyadenylation signal is underlined and the internal *Eco* R1 site is in bold face and underlined.

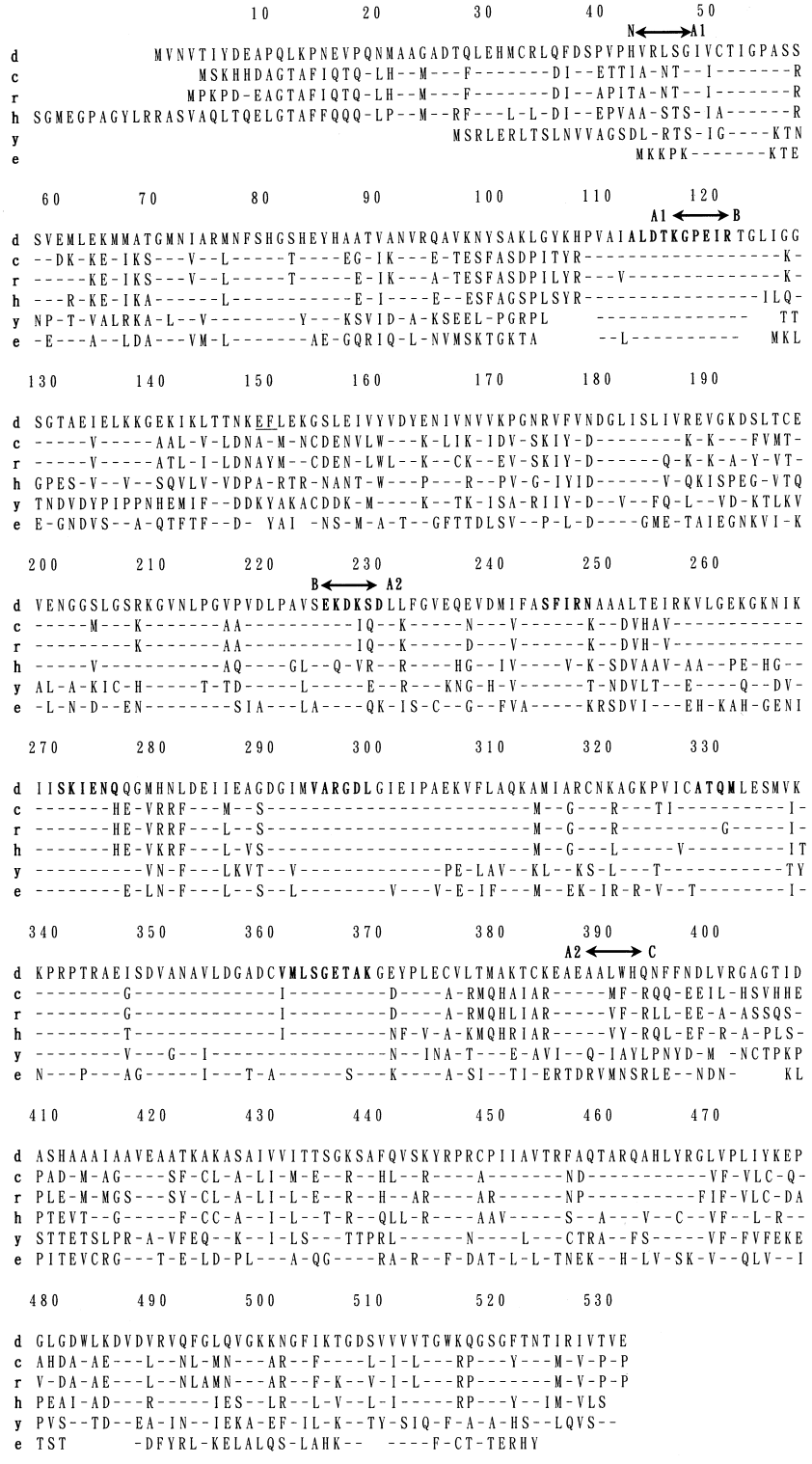

Fig. 3. Comparison of the *Drosophila melanogaster* pyruvate kinase amino acid sequence with the sequences of the protein of other organisms. Amino acid sequences have been aligned to obtain maximal positional identity. Dashed lines represent amino acids that are identical to those in corresponding positions in the *D. melanogaster* polypeptide. Blank spaces indicate the absence of amino acids at corresponding positions. The 4 domains of the pyruvate kinase subunit and the residues close to the active site, as determined for cat muscle enzyme, are marked N, A, B, C, and in bold face, respectively. The internal *Eco* RI site is underlined. The amino acid sequences presented are d: *Drosophila*; e: *E. coli* (Ohara et al. 1989); y: yeast (Burke et al. 1983); c: chicken (Lonberg and Gilbert 1983); r: rat pituitary (Parkison et al. 1989); and h: human L-PK (Tani et al. 1988).

overall positional identity between the sequence of the "*Drosophila*" enzyme and those from other sources is 42%-63% (Table 2). Using crystallographic analysis of the muscle enzyme, it has been shown that each subunit of the enzyme consists of 4 domains: an N-terminal domain, domains A (A1 and A2), B, and C (Muirhead et al. 1986). Based on crystallographic and solution studies, the enzyme active sites have been determined to lie in the pocket between domains A and B. The amino acid residues constituting the active sites (Fig. 3; Table 3) have also been determined, and it is obvious that these residues are highly conserved among species. The C domain, which presumably is involved in the regulation of enzyme activity using allosteric effectors, is much more variable than are domains A and B (Table 2), which may reflect the unique kinetic features of each individual enzyme.

A gene tree (Fig. 4) was constructed using comparisons of amino acid sequences of PK from various organisms using UPGMA analysis of PC/GENE Release 6.7 (IntelliGenetics, Inc.). As expected, the evolution of the pyruvate kinase gene is consistent with the hierarchical relationship from lower to higher organisms.

Table 2. Positional identity between the amino acid sequences of pyruvate kinase from *Drosophila melanogaster* and various other organisms

Organism	Total polypeptide positional identity (%)	Domain positional identity (%)			
		A1	B	A2	C
E. coli	42	49	42	60	31
Yeast	46	47	39	68	38
Chicken M	61	59	62	77	53
Rat pituitary	62	62	64	78	52
Human L	57	64	52	69	53

Polymerase chain reaction for genomic DNA using *Pyk*-specific primers

Twelve pyruvate kinase gene-specific primers were used to amplify genomic DNA from *D. melanogaster*. As a result, a total of 9 overlapping DNA fragments, covering the coding region of the *Pyk* gene, was amplified (Fig. 1a). Each amplicon was directly reacted using cycle-sequencing and was followed by separation of DNA fragments and sequence analyses in an ABI Prism 377 DNA sequencer. The nucleotide sequences of each amplicon were determined and reconfirmed by sequencing both strands more than twice, except that between primers PKp11 and PKm3 (Fig. 1a). Due to its length, the sequence was determined by at least 3 runs of the reaction for a single strand instead of 2 complementary strands. Fortunately, we can resolve at least 700 bases using the auto-sequencer, so we were able to determine the complete nucleotide sequence of the DNA fragment starting from both ends (GenBank AF062478).

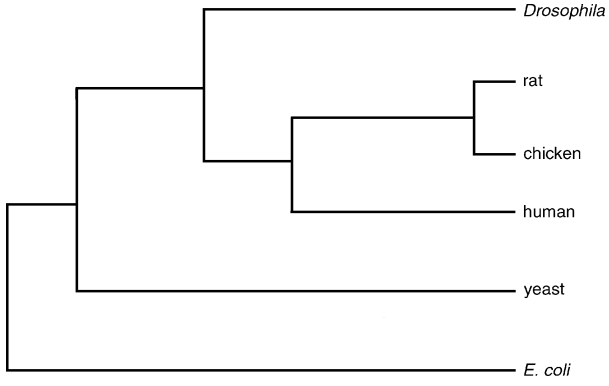

Fig. 4. Phylogenetic tree of the pyruvate kinase genes of *Drosophila*, *E. coli*, yeast, chicken, rat, and human, constructed by comparing amino acid sequences of pyruvate kinase proteins using UPGMA analysis of PC/GENE Release 6.7 (IntelliGenetics, Inc.).

Table 3. Residues constituting the active sites of pyruvate kinases. The determination of active sites of pyruvate kinase was based on the known structures of cat M1 enzyme (Muirhead et al. 1986)

Enzyme	Aα3/B	B/Aα3	Aβ4/α4	Aβ5/α5	Aβ6/α6	Aβ7/α7	ABβ/α8
Drosophila	ALDTKGPEIR	EKDKSD	SFIRN	SKIENQ	VARGDL	ATQM	VMLSGETAK
E. coli	LLDTKGPEIR	EKDKQK	SFIRK	SKIENQ	VARGDL	ATQM	VMLSGESAK
Yeast	ALDTKGPEIR	EKDKED	SFIRT	SKIENQ	VARGDL	ATQM	VMLSGETAK
Chicken M	ALDTKGPEIR	EKDIQD	SFIRK	SKIENH	VARGDL	ATQM	IMLSGETAK
Rat Pit.	ALDTKGPEIR	EKDIQD	SFIRK	SKIENH	VARGDL	ATGM	IMLSGETAK
Human L	ALDTKGPEIR	EQDVRD	SFVRK	SKIENH	VARGDL	ATQM	IMLSGETAK

```
1                                                                                                    TCGAACTTTTTGTCTCGTCCTCGCCGCGTCCGATTTTCCATAT   43
44 ATCCCGATTTTCGTCTCAAACTTGGAAGATATATTGCCAAAATATTGCATTCGGCTAGCCGCTGCGTTGTGTGTGTTGTTGTGCGTGGAGTTGCGTGTGCAG  145
146 TTGCAAGAAGATATAGTTGTCTATATCGCCAGACTGATCTCGATCCCAAGTGCGATTCCCGTTGGTAACATACTGCCCGTTCGGTAGGCAACAAAATGGTGA  247
1                                                                                                                                METValA    3
248 ACGTAACCATTTACGATGAGGCACCGCAGCTGAAGgtaggccaatgtcacggtaatagaagccctcagagttggtccataaagatattcataatcgccataa  349
3 snValThrIleTyrAspGluAlaProGlnLeuLys                                                                                                    14
350 tcagttggctggcgagttggctcgatatcagagtcattggggaggaaaatgtgggaaaagacaattatgcaactgcagctgagaaaattctgtcggtcgcgg  451
452 tcggcagaactgaccttcggcggtgcagatgtgcagcgctgccggtgtcggcttcgctgcacacttttcttccgcctgctgctttctctctgccactctctc  553
554 ttccgctctgcctttctcaggtcctctctacgggggacacccctaaccacccacgcccatccactcaagcacttcaagtacacgcactttgtttacatcagc  655
656 gaattctgcagcgacagctttgcgtttttcaaaacataacatgcacagtgcgaataaatattttgataatttaagcaatcgtaagagttaacaaatagtatt  757
758 aacgcaacagcatcatttaatattggtaatcgattataacattttatatgtcaagttttatcaatggcttatttacacatgctactaagtataatcagaatg  859
860 ataaaatactgttaatatataacgttatcattttaattaggattattagattatttaaaatcattacctttatatgtaaatagctctattttatctgtcagt  961
962 tcacaaatatttcaatcattccttttgtcatgaaaaatgattggtgcggcggcatttgggcttaaatcgccgtatttggccgtttggttctatttataagcc 1063
1064 ggtggtaaacatttaatttggtcgataaaaactgtatcgatttatattcgttgatatgaagaagggttatcatgagttttgcacgcattcaccgcgtttatg 1165
1166 ttggggttttaaatgtgtttcgcattgcatgacttcaccagggcatgattttaatatcataaaactagttggttttttttatttttgctagaacttggtagt 1267
1268 ttttaaaacgcgtagggtttcttgttgctaggaagagacttctgattattagcgtaaattggaaataaatgcttaaaatcaataaacacgactttaacctgt 1369
1370 tttattgcaaaagttggcaatcagcgtgttgatttctaaaaataaagaatcaattcgcgctgtttataaatatgaagagaaatgaaaacaaacagaattaag 1471
1472 tatagttttcattttgcagCCCAACGAAGTACCCCAAAACATGGCCGCAGGAGCTGATACCCAACTGGAGCACATGTGCCGTCTGCAGTTCGACTCGCCAGT 1573
15                                ProAsnGluValProGlnAsnMETAlaAlaGlyAlaAspThrGlnLeuGluHisMETCysArgLeuGlnPheAspSerProVa    42
1574 GCCCCATGTGCGTCTGTCCGGAATCGTGTGCACCATCGGACCTGCCTCCAGCAGCGTGGAGATGCTGGAGAAGATGATGGCCACCGGCATGAACATCGCGCG 1675
42 IProHisValArgLeuSerGlyIleValCysThrIleGlyProAlaSerSerSerValGluMETLeuGluLysMETMETAlaThrGlyMETAsnIleAlaAr    76
1676 CATGAACTTCTCTCACGGATCGCACGAGTACCATGCCGCCACCGTGGCCAATGTGCGCCAGGCGGTGAAAAACTACTCGGCCAAGCTGGGCTACAAACACCC 1777
76 gMETAsnPheSerHisGlySerHisGluTyrHisAlaAlaThrValAlaAsnValArgGlnAlaValLysAsnTyrSerAlaLysLeuGlyTyrLysHisPr   110
1778 CGTGGCCATTGCCCTGGACACCAAGGGGCCCGAGATCCGTACCGGTCTGATCGGAGGCAGCGGCACCGCCGAGATTGAGCTGAAGAAGGGCGAGAAGATCAA 1879
110 oValAlaIleAlaLeuAspThrLysGlyProGluIleArgThrGlyLeuIleGlyGlySerGlyThrAlaGluIleGluLeuLysLysGlyGluLysIleLy   144
1880 GCTGACCACCAACAAGGAATTCCTGGAGAAGGGCTCTCTGGAGATTGTGTACGTGGACTACGAGAACATTGTCAATGTGGTGAAGCCCGGCAACCGGGTGTT 1981
144 sLeuThrThrAsnLysGluPheLeuGluLysGlySerLeuGluIleValTyrValAspTyrGluAsnIleValAsnValValLysProGlyAsnArgValPh   178
1982 CGTCGATGACGGTCTGATCTCACTGATTGTCCGCGAGGTAGGCAAGGATTCCCTCACCTGCGAAGTGGAGAACGGCGGCTCTCTGGGTTCCCGCAAGGGTGT 2083
178 eValAspAspGlyLeuIleSerLeuIleValArgGluValGlyLysAspSerLeuThrCysGluValGluAsnGlyGlySerLeuGlySerArgLysGlyVa   212
2084 GAACCTGCCAGGCGTGCCCGTCGATCTGCCTGCCGTCTCCGAGAAGGATAAGAGCGATCTGCTGTTCGGTGTGGAGCAGGAAGTGGACATGATCTTTGCTTC 2185
212 IAsnLeuProGlyValProValAspLeuProAlaValSerGluLysAspLysSerAspLeuLeuPheGlyValGluGlnGluValAspMETIlePheAlaSe   246
2186 GTTCATCCGCAACGCCGCTGCTTTGACCGAGATCCGTAAGGTTCTTGGCGAGAAGGGCAAGAACATCAAGATCATTTCCAAGATCGAGAACCAGCAGGGCAT 2287
246 rPheIleArgAsnAlaAlaAlaLeuThrGluIleArgLysValLeuGlyGluLysGlyLysAsnIleLysIleIleSerLysIleGluAsnGlnGlnGlyME   280
2288 GCACAACCTGGACGAGATCATCGAGGCCGGTGATGGCATTATGGTAGCCCGTGGAGATCTGGGTATTGAGATTCCCGCCGAGAAGGTGTTCCTCGCCCAGAA 2389
280 THisAsnLeuAspGluIleIleGluAlaGlyAspGlyIleMETValAlaArgGlyAspLeuGlyIleGluIleProAlaGluLysValPheLeuAlaGlnLy   314
2390 GGCCATGATTGCCCGCTGCAACAAGGCTGGCAAGCCTGTGATCTGCGCCACTCAGATGTTGGAGTCAATGGTGAAGAAGCCACGTCCCACTCGCGCTGAGAT 2491
314 sAlaMETIleAlaArgCysAsnLysAlaGlyLysProValIleCysAlaThrGlnMETLeuGluSerMETValLysLysProArgProThrArgAlaGluIl   348
2492 CTCTGATGTGGCCAACGCTGTGCTCGATGGTGCTGATTGCGTCATGTTGTCTGGTGAGACCGCCAAGGGCGAGTACCCGCTGGAGTGCGTCCTGACCATGGC 2593
348 eSerAspValAlaAsnAlaValLeuAspGlyAlaAspCysValMETLeuSerGlyGluThrAlaLysGlyGluTyrProLeuGluCysValLeuThrMETAl   382
2594 CAAGACCTGCAAGGAGGCCGAGGCTGCCCTCTGGCACCAGAACTTCTTTAACGACTTGGTTCGCGGCGCTGGTACCATCGATGCCTCTCACGCGGCTGCCAT 2695
382 aLysThrCysLysGluAlaGluAlaAlaLeuTrpHisGlnAsnPhePheAsnAspLeuValArgGlyAlaGlyThrIleAspAlaSerHisAlaAlaAlaIl   416
2696 CGCTGCCGTTGAGGCTGCCACCAAGGCCAAGGCCTCCGCCATCGTGGTGATCACCACCAGCGGCAAGTCGGCCTTCCAGGTGAGCAAGTACCGCCCACGCTG 2797
416 eAlaAlaValGluAlaAlaThrLysAlaLysAlaSerAlaIleValValIleThrThrSerGlyLysSerAlaPheGlnValSerLysTyrArgProArgCy   450
2798 CCCCATCATCGCGGTCACCCGTTTCGCGCAGACCGCCCGACAGGCCCATCTCTACCGTGGACTGGTGCCACTCATCTACAAGGgtaggtttttgtttctcca 2899
450 sProIleIleAlaValThrArgPheAlaGlnThrAlaArgGlnAlaHisLeuTyrArgGlyLeuValProLeuIleTyrLysG                        478
2900 ggtcgttctttaaattataactctggcttgtacttgcagAGCCCGGTCTTGGTGACTGGCTGAAGGACGTGGACGTGCGCGTGCAGTTCGGTCTGCAGGTCG 3001
478                                   luProGlyLeuGlyAspTrpLeuLysAspValAspValArgValGlnPheGlyLeuGlnValG   499
3002 GAAAGAAGAACGGCTTCATCAAGACCGGCGATTCCGTCGTGGTGGTTACCGGCTGGAAGCAGGGCTCCGGCTTCACCAACACCATCCGCATTGTgtaagttt 3103
499 lyLysLysAsnGlyPheIleLysThrGlyAspSerValValValValThrGlyTrpLysGlnGlySerGlyPheThrAsnThrIleArgIleVa          530
3104 aaattctagcatgtagatatatatgtatctgtgcaaaaattatatcaacaaaactaacgcaattatctatcatttcagCACCGTCGAATAAATTCTGCCGCT 3205
531                                                                                                  IThrValGlu---    533
3206 CAGGAAGTACCTGGATGCGTCGCAGCTGCCGCGGGCGTCGCGTTGATTACTGTTAATAACCAAGTTAATTTATTAATTACACAATATTCGGAATATTGGAAA 3307
3308 CTGCAGGAGGCTGCTGGAACCGCTGTCAGAAGCGCCGAAGGCCGGCCCGGGCCTCAAACAGACACACAAATATACTATACCATTGGTGAATAATAGAGATTT 3409
3410 CTTCGAAAATTATAATAAAGTCATCACTGTATTTATGT                                                                3447
```

Fig. 5. Total sequence of pyruvate kinase genomic DNA from *Drosophila melanogaster*. The nucleotide sequence was determined from 9 overlapping amplicons. Uppercase letters indicate the sequence of exons and those in lowercase represent that of introns. The internal *Eco* R1 site is in bold face and underlined. Numbers on both ends are positions of nucleotides or amino acids.

The nucleotide sequence determined from the amplicons had a total of 3447 base pairs (GenBank AF062478). By comparson with the sequence of *Pyk* cDNA clones from *D. melanogaster*, 4 exons were identified (Figs. 1a, 5). Exon 1 contains a 240-bp 5'-untranslated region plus the 1st 14 amino acids of the protein, and is followed by a long intron (1208 bp). Exon 2 codes for amino acids 15-477, which contains the active site of the protein (Fig. 4). Exon 3 codes for 53 amino acids (478-530) and is flanked by 2 short introns (58 and 86 bp, respectively). Exon 4 codes for the last 3 amino acids of the protein plus a 3'-untranslated region. Introns were identified as those nucleotides not present in the cDNA sequence, and were further confirmed by having the consensus 5'- and 3'-splicing sites (GT-AG).

RT-PCR analysis

To ensure that the prepared cDNAs would encompass the full length of mRNA of pyruvate kinase, they were reversibly transcribed by extension from either primer PK1 or primer PKm2 (Fig. 1b). Seven amplicons (Fig. 6) were obtained from pairs of primers as shown in figure 1b. They overlapped with each other and covered the complete sequence of the pyruvate kinase transcript. The observation that there was only a single amplicon in each amplification suggests that only a single pyruvate kinase transcript is expressed in adults of *D. melanogaster* (Fig. 6). It may further imply that there is only a single *Pyk* gene in *D. melanogaster*.

5' RACE analysis

A segment of ca. 270 bp was amplified using PKm4 and the abridged anchor primer as primers from adult cDNA reverse transcribed from mRNA, and primed by PK2 (data not shown). The segment was then further sequenced, and the TSP was identified as an adenine next to poly-C in the anticoding strand, revealing thymine in the coding strand (Fig. 1a).

DISCUSSION

Comparisons of the amino acid sequences of various pyruvate kinase proteins reveal a 40%-70% homology between sequences (Ohara et al. 1989). Moreover, there are 2 domains that have amino acid homology of greater than 80%. This has allowed workers to clone *Pyk* genes using heterologous probes (Inoue et al. 1986). Quite a few pyruvate kinases from vertebrates have been cloned using heterologous *Pyk* probes (Noguchi et al. 1983 1986 1987, Cognet et al. 1987, Tani et al. 1988). Using the *Pyk* cDNA clones, pCJ11 and pCJ22, we successfully isolated 2 *Drosophila Pyk* cDNA clones (2 *Eco* R1 fragments). The 2 *Eco* R1 fragments have been subcloned and designated as cDRPK15 and cDRPK06. The 2 cDNA subclones span the complete coding sequence of pyruvate kinase gene with a 5' untranslated region of 240 bp , and a 3' untranslated region of 253 bp (Fig. 4). There is an internal *Eco* R1 site in *Drosophila Pyk* cDNA that is not found in the *Pyk* genes of other organisms. To see if the internal *Eco* R1 site was an artificial product formed during the process of construction of the cDNA library, we re-screened the λ-gt10 cDNA library using cDRPK15 as a probe. All 4 positive clones, plus 2 previously isolated, appeared to have 1.5- and 0.6-kb inserts (data not shown), indicating that the internal *Eco* R1 site was real. This was further confirmed by the nucleotide sequence of *Pyk* genomic DNA (Fig. 5).

The molecular organization of *Pyk* genes has changed during the divergence of the genes and the organisms that carry them. By matching the positions of the introns in the protein-coding sequences between rat and chicken *Pyk* genes, Noguchi et al. (1987) suggested that these genes evolved from a common ancestral gene by gene duplication. The *Pyk* gene tree constructed in this study (Fig. 4) sup-

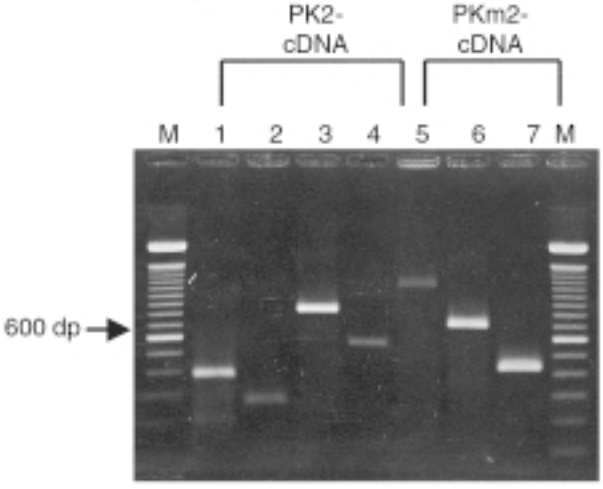

Fig. 6. RT-PCR for PK2-cDNA and PKm2-cDNA. PK2-cDNA was constructed from reverse-transcription of mRNA from *Drosophila melanogaster* using PK2 as primer, while PKm2-cDNA used PKm2 as primer. M: 100-bp DNA ladder marker. Paired primers used for PCR were lane 1: PKp2-PKm3; lane 2: PKp11-PKm3; lane 3: PKp9-PK2; lane 4: PK1-PK2; lane 5: PKp1-PKm2; lane 6: PKp4-PKm2; and lane 7: PKp3-PKm2.

ports this view, since chickens and rats are clustered together in a group. However, since there are 4 exons in the *Drosophila Pyk* gene, but 11 exons in the genes of chickens and rats (Lonberg and Gilbert 1985, Noguchi et al. 1987), we may suggest that vertebrate *Pyk* genes have diverged from that of *Drosophila* by insertion of introns, though we cannot rule out the possibility of losses of introns in *Drosophila*. Therefore, the evolving rate of *Pyk* genes might be considered rapid according to the following observations. The intron/exon structures of *Pyk* genes vary along the evolutionary lineage from lower to higher organisms, and relatively large numbers of amino acid replacements have accumulated in the regions beyond the active sites of the proteins (Fig. 3; Table 2).

Mammalian *Pyk* genes were reported to encode 4 distinct isozymes, the M_1-, M_2-, R-, and L-types (Imamura and Tanaka 1972, Imamura et al. 1972 1973). Molecular studies of gene structure of the 4 rat isozymes have confirmed that the M_1- and M_2-types are encoded by the same gene, and that the L- and R-types are produced from another gene (Noguchi et al. 1986 1987). The L- and R-type isozymes are produced by utilization of alternative promoters, whereas, the M_1- and M_2-type isozymes are produced by alternative splicings (Noguchi et al. 1986 1987). We were interested in whether there are multiple PK isozymes encoded by the *Pyk* gene in *D. melanogaster*, and whether there is more than 1 *Pyk* gene in fruit flies. Cellulose acetate electrophoresis of PK was performed on all 3 stages (larvae, pupae [imaginal disks], and adult). Since single, equal-mobility bands were present in each stage, this suggests that the PKs of the 3 stages are encoded by a single gene, and that there is no modification in production of PK in any developmental stages, either at the transcriptional or post-transcriptional levels (data not shown). Furthermore, RT-PCR is a fast and convenient method to identify the number of species of transcripts. If multiple transcripts are produced via alternative promoters as in the case of mammalian PKL- and R-isozymes, 2 or more bands should be amplified from the 5' region of the gene; however if those are produced through an alternative splicing mechanism as in the case of mammalian PKM1- and M2- isozymes, 2 or more bands would be amplified from the middle part of the gene; or there would be at least 2 bands amplified from the 3' region of the gene if those are produced by a different polyadenylation mechanism. The RT-PCR analysis (Fig. 6) shows that there is a single amplicon amplified from various portions but which complete for mRNA, implying that there is only a single species of mRNA transcribed from a *Pyk* gene, at least in the adult stage.

As mentioned in the “Introduction”, understanding the coordinated regulation of expression of the *Pyk* and *Ark* genes in *D. melanogaster* was the main objective of this study. To clarify this phenomenon, 5' *cis*-acting regulatory elements of both the *Pyk* and *Ark* genes as well as the *trans*-acting transcription factors, should be identified. To do so, genomic clones, which contain the 5' -untranscribed regulatory region of the genes, are required.

REFERENCES

Burke RL, P Tekamp-Olson, R Najarian. 1983. The isolation, characterization, and sequence of the pyruvate kinase gene of *Saccharomyces cerevisiae*. J. Biol. Chem. **258:** 2193-2201.

Cognet M, YC Lone, S Vaulont, A Kahn, J Marie. 1987. Structure of the rat L-type pyruvate kinase gene. J. Mol. Biol. **196:** 11-25.

Garen A, L Kauvar, J-A Lepesant. 1977. Roles of ecdysone in *Drosophila* development. Proc. Natl. Acad. Sci. USA **74:** 5099-5103.

Imamura K, T Tanaka. 1972. Multimolecular forms of pyruvate kinase from rat and other mammalian tissues. I. Electrophoretic studies. J. Biochem. **71:** 1043-1051.

Imamura K, T Tanaka, T Nishina, K Nakashima, S Miwa. 1973. Studies on pyruvate kinase (PK) deficiency. II. Electrophoretic, kinetic and immunological studies on pyruvate kinase of erythrocytes and other tissues. J. Biochem. **74:** 1165-1175.

Imamura K, K Taniuchi, T Tanaka. 1972. Multimolecular forms of pyruvate kinase II: purification of M_2-type pyruvate kinase from Yoshida Ascites hepatoma 130 cells and comparative studies on the enzymological and immunological properties of the three types of pyruvate kinases, L, M_1, and M_2. J. Biochem. **72:** 1001-1015.

Inoue H, T Noguchi, T Tanaka. 1986. Complete amino acid sequence of rat L-type pyruvate kinase deduced from the cDNA sequence. Eur. Biochem. **154:** 465-469.

James JM, GE Collier. 1988. Distribution and genetic basis of arginine kinase in wild type and flightless mutants of *Drosophila melanogaster*. J. Exp. Zool. **248:** 185-191.

James JM, GE Collier. 1990. Hormonally regulated expression of arginine kinase in *Drosophila melanogaster*. Roux’s Arch. Dev. Biol. **198:** 474-478.

James JM, GE Collier. 1992. Early gene interaction during prepupal expression of *Drosophila* arginine kinase. Dev. Genet. **13:** 302-305.

Lonberg N, W Gilbert. 1983. Primary structure of chicken muscle pyruvate kinase mRNA. Proc. Natl. Acad. Sci. USA **80:** 3661-3665.

Muirhead H, DA Clayden, D Barford, CG Lorimer, LA Jothergill-Gilmore, E Schiltz, W Schmitt. 1986. The structure of cat muscle pyruvate kinase. EMBO J. **5:** 475-481.

Noguchi T, H Inoue, HL Chen, KI Matsubara, T Tanaka. 1983. Molecular cloning of DNA complementary to rat L-type pyruvate kinase mRNA. J. Biol. Chem. **258:** 15220-15223.

Noguchi T, H Inoue, T Tanaka. 1986. The M_1- and M_2-type isozymes of rat pyruvate kinase are produced from the

same gene by alternative RNA splicing. J. Biol. Chem. **261:** 13807-13812.
Noguchi T, K Yamada, H Inoue, T Matsuda, T Tanaka. 1987. The L- and R-type isozymes of rat pyruvate kinase are produced from a single gene by use of different promoters. J. Biol. Chem. **262:** 14336-14371.
Ohara O, RL Dorit, W Gilbert. 1989. Direct genomic sequencing of bacterial DNA: the pyruvate kinase I gene of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **86:** 6883-6887.
Parkison C, H Kato, SY Cheng. 1989. The nucleotide sequence of a full length cDNA encoding rat pituitary pyruvate kinase. Nucleic Acid Res. **17:** 7106.
Sambrook J, EF Fritsch, T Maniatis. 1989. Molecular cloning: a laboratory manual. 2nd ed. New York: Cold Spring Harbor Laboratory Press.
Strickberger MW. 1962. Experiments in genetics with *Drosophila*. New York: J. Wiley, pp. 8-12.
Tani K, MC Yoshida, H Satoh, K Mitamura, T Noguchi, T Tanaka, H Fujii, S Miwa. 1988. Human M_2-type pyruvate kinase: cDNA cloning, chromosomal assignment and expression in hepatoma. Gene **73:** 509-516.

黃果蠅丙酮酸鹽激酶基因選殖及其基因結構之分析

簡一治[1]　朱玉菁[2]　闕淳美[3]

利用人類及大白鼠丙酮酸鹽激酶的 cDNA 當探針篩選黃果蠅 λ-gt10 cDNA 圖書館，我們得以分離出二個 cDNA 菌株，分別是 cDMPK15 及 cDMPK06。其核甘酸序列顯示此二 cDNA 菌株包含黃果蠅丙酮酸鹽激酶基因的密碼區域 (1602 鹼基對：533 個胺基酸 + TAA) 加上 240 鹼基對的 5 端未轉譯區域及 253 鹼基對的 3 端未轉譯區域。其胺基酸序列和其它生物的丙酮酸鹽激酶的胺基酸序列比較顯示構成活性中心區域的胺基酸殘基具有高度保守性。除此之外，其胺基酸序列和其它生物的丙酮酸鹽激酶的胺基酸序列的同源性則是 42-63%。聚合酶連鎖反應被應用去複製黃果蠅丙酮酸鹽激酶基因片段。這些 DNA 片段合起來包括整個丙酮酸鹽激酶的密碼區域則以 ABI 377 核酸自動定序儀進行直接定序。共 3477 個鹼基對被定序出。和 cDNA 序列比較結果 4 個表現子則被決定出，它們分別是 282，1390，157 及 266 鹼基對長。內轉子則由於它們未出現在 cDNA 序列中以及全部含有保守的 5 端及 3 端切割區域 (GT-AG)。RT-PCR 則是被用來決定黃果蠅成蟲丙酮酸鹽激酶轉錄物的種類，而從結果中每一個聚合酶連鎖反應僅複製出一條複製物建議在黃果蠅成蟲中可能只有一條丙酮酸鹽激酶轉錄物，此外更進一步暗示只有一個丙酮酸鹽激酶基因在黃果蠅中。

關鍵詞：丙酮酸鹽激酶，基因選株，反轉錄聚合酶連鎖反應，黃果蠅。

[1] 國立彰化師範大學生物學系