1995

The Rabbit Muscle Phosphofructokinase Gene: Six Types of mRNAs Are Tissue-Specifically Expressed by Multiple Promoters and Alternative Splicing.

Shouhua Xiao  
Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://repository.lsu.edu/gradschool_disstheses

Recommended Citation  
https://repository.lsu.edu/gradschool_disstheses/6079

This Dissertation is brought to you for free and open access by the Graduate School at LSU Scholarly Repository. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Scholarly Repository. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
THE RABBIT MUSCLE PHOSPHOFRUCTOKINASE GENE:
SIX TYPES OF mRNAs ARE TISSUE-SPECIFICALLY EXPRESSED
BY MULTIPLE PROMOTERS AND ALTERNATIVE SPLICING

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biochemistry

By

Shouhua Xiao
B.S., Jiangxi Agricultural University, China, 1983
M.S., Beijing Agricultural University, China, 1986
December 1995
Dedicated to
late Professor Shenran Ge
my father in-law
who won a First Class Scientific Accomplishment Award
from the Chinese National Science Congress
the highest honor and achievement for scientists in China
who died of lung cancer
shortly after I left China for LSU in 1989
ACKNOWLEDGEMENTS

I would like to express my many sincere thanks first and foremost to my major professor, Dr. Simon H. Chang for his academic guidance, financial support, and constant encouragement through out the course of my study. I would also like to thank the members of my Advisory Committee, Dr. Sue G. Bartlett, Dr. Patrick J. DiMario, Dr. James V. Moroney, and Dr. Ding S. Shih for their invaluable advice and kind help. I would like to extend my sincere thanks to Professor Jishu Pan of Beijing Agricultural University, China, my M.S. advisor, for her long-term encouragement and support.

I would like to acknowledge pleasant cooperation from Dr. Zhuo Chen in the early phase of this project, helpful discussion from Dr. G. A. Dunaway, and helpful assistance from Zhigong Lin and Dr. Jau-Yi Li. We would like to thank Dr. G. A. Dunaway for helpful discussion. I also sincerely thank Malcolm Byrnes, Susan DeFelice, Mary Heine, Christian Knaak, and Michele Rankin for their critical reading of this dissertation. This work is supported by National Institutes of Health Grant # DK 31676.

I am indebted to my parents and my mother-in-law for their support and sacrifice due to my absence at their high age. Last but by no means least, I would like to express my deepest appreciation to my wife, Lin, for her love, patience, and support.
This dissertation consists of four major parts: a LITERATURE REVIEW, SECTION I, SECTION II, and CONCLUSIONS AND FUTURE STUDIES. SECTION I AND SECTION II are each presented in the form of a manuscript. The LITERATURE REVIEW is intended to provide a comprehensive review of the PFK literature pertinent to this study, and therefore some overlap may exist between this part of the dissertation and the introduction in SECTION I or SECTION II. The DISCUSSION AND FUTURE STUDIES part of this work summarizes the results of this research and I have proposed in this part several experiments to further our understanding of the regulation of the rmPFK gene. Again, some overlap may be found between this part and the discussion of either SECTION I or SECTION II.
# TABLE OF CONTENTS

DEDICTION .......................................................................................................................... ii

ACKNOWLEDGEMENT .......................................................................................................... iii

PREFACE ............................................................................................................................... iv

LIST OF TABLES .................................................................................................................... vi

LIST OF FIGURES ................................................................................................................. vii

ABSTRACT ............................................................................................................................ ix

LITERATURE REVIEW .......................................................................................................... 1

SECTION I  Tissue-specific expression of six types of rabbit muscle phosphofructokinase mRNAs by multiple promoters and alternative splicing ......................... 46

SECTION II Primer extension at high temperature by *Thermus thermophilus* DNA polymerase ....... 92

CONCLUSIONS AND FUTURE STUDIES ................................................................. 105

REFERENCES ....................................................................................................................... 116

APPENDIX Restriction enzyme map of the 5' flanking region of the rabbit muscle phosphofructokinase gene .................. 124

VITA .................................................................................................................................... 136
LIST OF TABLES

1. Tissue-specific expression patterns of the mammalian PFK isozymes in human, rabbit, and rat tissues .......................................13

2. Oligonucleotides used in this study .........................50
# LIST OF FIGURES

1. The glycolytic pathway ...........................................3

2. The reaction catalyzed by PFK in the glycolytic pathway ..........4

3. Diagram depicting general view of BsPFK and its quaternary and tertiary structure changes on the allosteric transition from R to T ...............18

4. Schematic view of the 6-F loop (residue 155-162) of BsPFK showing its different interactions across the interface between dimers .................21

5. Sequence comparison of homologous PFKs and position of introns along the amino acid sequence of rabbit muscle PFK ........................................23

6. Location of substrate and effector sites in the hypothetical rmPFK protein model ..................25

7. Exon/intron organization of the rabbit muscle PFK gene and comparison of the 5’ UTRs of two rabbit and three human muscle PFK cDNAs ..........27

8. Sequence diversity in the 5’ UTRs of two rmPFK cDNAs ..................................................29

9. Alternative splicing in Drosophila somatic cell sex determination ..................33

10. Strategy of anchored PCR cloning ..................................52

11. Sequences of rmPFK cDNAs cloned by anchored PCR and RT-PCR ........................................59

12. Schematic diagram depicting the organization of the 5’ portion of the rmPFK gene and the six mRNAs identified by anchored PCR and RT-PCR ..........61

13. Sequences of the 5’ flanking region of the rmPFK gene ............................66

14. Northern blot analysis of different rmPFK mRNAs ..........68

15. Analysis of the tissue-specific expression of the six mRNAs of the rmPFK gene by RT-PCR using thermostable reverse transcriptase ..............71
16. Determination of transcription start sites of the proximal promoter of the rmPFK gene .............75
17. Determination of transcription start sites of the promoter close to the TATAAAA sequence of the rmPFK gene ..................................................78
18. Performance of Tth DNA polymerase in primer extension at different temperatures .........................99
19. Effect of extension time on the primer extension by Tth DNA polymerase ...............................101
20. Effect of actinomycin on primer extension by Tth DNA polymerase .........................................102
21. Comparison of primer extension of GAPDH RNA by Tth DNA polymerase and that by M-MLV reverse transcriptase ...........................................103
Phosphofructokinase (ATP:Fructose-6-phosphate 1-phosphotransferase, EC 2.7.1.11, PFK) catalyzes the phosphorylation of fructose 6-phosphate to generate fructose 1,6-bisphosphate, the rate-limiting step of glycolytic pathway. Li et al. (1990) have identified two cDNAs for the rabbit muscle PFK with an identical coding regions but different 5' untranslated regions (UTRs). In this dissertation research, six cDNAs for the rabbit muscle PFK mRNAs were cloned. These cDNAs are different in their 5' UTRs but have an identical open reading frame. The transcription initiation sites of two downstream promoters of this gene have been determined. A proximal promoter (Pa) is TATA-less and 86 bp upstream of the ATG start codon. The mRNA expressed from this promoter has been detected in skeletal muscle and heart tissues. The second promoter (Pb) is skeletal muscle specific and produces two mRNAs (mRNA-b1 and mRNA-b2) by alternative splicing. Part of the 5' UTR in mRNA-b1 is spliced out to yield mRNA-b2. A fourth mRNA (mRNA-c) is also skeletal muscle specific.

An additional two mRNAs (mRNA-d1 and -d2) are produced by alternative splicing. An exon of 106 bp was retained in mRNA-d1, but not in mRNA-d2. Interestingly, mRNA-d1 was detected in skeletal muscle, heart, and brain, but not in liver and kidney tissues. Its splicing variant, mRNA-d2 is detected in all five tissues examined. Therefore, the 106 bp
exon is retained in a tissue-specific manner. Since the 5' ends of mRNA-c, -d1, and -d2 are upstream of promoters a and b, these mRNAs can not be produced by these promoters. Therefore, rmPFK expresses its six mRNAs by at least three promoters and alternative splicing.

During the course of this research, I have developed a protocol for primer extension at higher temperature. In primer extension using M-MLV reverse transcriptase, a secondary structure of the RNA template was found to be problematic. I therefore optimized conditions for primer extension using thermostable DNA polymerase from Thermus thermophilus at high temperature. The primer extension at high temperature (65°C) using this enzyme reduced the artifact caused by the secondary structure in the RNA template.
LITERATURE REVIEW
PHOSPHOFRUCTOKINASE IS A RATE REGULATING ENZYME IN GLYCOLYSIS

In glycolysis, glucose is degraded to generate ATP, NADH, and pyruvate. Pyruvate is further converted to acetyl CoA, providing building blocks for other synthetic reactions, such as the formation of long-chain fatty acids (Fig. 1). In glycolysis, phosphofructokinase (ATP: D-fructose-6-phosphate 1-phosphotransferase, EC 2.7.1.11, PFK) catalyzes the phosphorylation of fructose-6-phosphate (F-6-P) by ATP, to produce fructose-1,6-bisphosphate (F-1,6-BP) and ADP (Bloxham and Lardy, 1973; Uyeda, 1979; Fig. 2). In metabolic pathways, enzymes catalyzing essentially irreversible reactions are potential sites of control. In glycolysis, the reactions catalyzed by three enzymes (hexokinase, phosphofructokinase, and pyruvate kinase) are irreversible. Therefore, these enzymes are potential sites of control in the glycolytic pathway. Hexokinase catalyzes the first irreversible reaction, the phosphorylation of glucose to produce glucose 6-phosphate. This product is not only a glycolytic intermediate, but can also be oxidized by the pentose phosphate pathway to generate NADPH or be converted into glycogen. Although PFK catalyzes the second irreversible reaction, this reaction is the first committed step in glycolysis, and therefore the most important control site in glycolysis (Bloxham and Lardy, 1973; Uyeda, 1979). One of the best known examples of PFK as the control site of glycolysis is its role in the Pasteur effect (inhibition of
Figure 1. The glycolytic pathway. The reaction catalyzed by PFK is in bold. The conversion of the glycolytic product pyruvate to acetyl CoA and further to fatty acids is shown by dashed arrows.
Figure 2. The reaction catalyzed by PFK in the glycolytic pathway. Fructose 6-phosphate is phosphorylated to produce fructose 2,6-bisphosphate.
fermentation by oxygen). In studying fermentation of yeast, Pasteur found that fermentation is inhibited by oxygen. The rate of fermentation is about seven fold lower under aerobic conditions than under anaerobic conditions. This inhibition is largely due to the inhibition of PFK activity by ATP and citrate, because the product of PFK, F-1,6-BP was readily fermented in the presence of oxygen (reviewed in Dunaway, 1983; Bloxham and Lardy, 1973).

The physiological significance of PFK is examplified in the hereditary PFK deficiency in human skeletal muscle, known as the type VII glycogenosis, or Tarui disease (Tarui et al., 1965). The symptoms of this disease include intolerance to muscular exercise, increased hemolysis, and hyperuricemia (Tarui et al., 1965; Nakajima et al., 1990). The patients with this disease have no PFK activity in their skeletal muscle (Tarui et al., 1965).

**REGULATION OF MAMMALIAN PFK ACTIVITY BY EFFECTORS**

Mammalian PFKs are allosteric enzymes. Tetramers are the smallest active form for PFK isolated from mammalian sources (Dunaway, 1983). Each tetramer has several effector binding sites in addition to the substrate binding sites. Binding of effectors, either inhibitors or activators allosterically changes the kinetic properties of this enzyme. As mentioned above, PFK is inhibited by ATP and citrate. On the other hand, PFK is activated by inorganic phosphate (Pi), AMP, cyclic AMP, ADP, GDP, and its product F-1,6-BP. The regulations of mammalian PFKs by the effectors are very
complex (Bloxham and Lardy, 1973). Here, only several major features of such regulations are summarized below:

1. The enzyme is inhibited by a high concentration of ATP.

2. The enzyme is also inhibited by citrate, an intermediate in the Krebs cycle. This inhibition is synergistic with that by ATP.

3. The inhibition by ATP can be relieved by Pi, AMP, cyclic AMP, ADP, F-1,6-BP, and F-2,6-BP.

4. Binding of substrate F-6-P is cooperative. Inhibitors decrease the affinity of PFK to F-6-P, while the activators increase the affinity.

The inhibition of rabbit PFK activity by ATP was first reported by Lardy and Parks (1956). Increasing the concentration of ATP above that necessary for the optimum catalytic activity causes a progressive decrease in PFK activity. This allosteric effect results from the decreased affinity for F-6-P upon the binding of ATP to the effector site. As the concentration of F-6-P is increased, the effect of ATP inhibition is reduced (Mansour and Ahlfors, 1968).

The inhibitory effect of ATP is counteracted by the metabolic degradation products of ATP, including Pi, AMP, and ADP (Bloxham and Lardy, 1973). Thus, PFK activity is closely regulated in relation to physiological conditions: high ATP/ADP ratio inhibits PFK activity, while high ADP/ATP ratio activates the PFK activity.
The pH of the medium also affects the inhibition of ATP. At pH 7.0 and with saturating ATP, rabbit muscle PFK is inhibited by ATP and exhibits sigmoidal kinetics with respect to F-6-P. At pH 8.0, however, ATP inhibition and cooperativity are not observed (Hanson et al., 1973). Similar results have also been reported for PFKs from brain (Tsai and Kemp, 1974; Lowry and Passonneau, 1966), heart (Mansour and Ahlfors, 1968), and liver (Kemp, 1971).

Citrate is also an important inhibitor of PFK activity. This links the regulation of glycolysis with the Krebs cycle. Using rat heart PFK, Garland et al. (1963) found that citrate is the only Krebs cycle intermediate effective in inhibiting PFK activity. However, muscle PFK has been found to be inhibited also by isocitrate, though less than by citrate (Parmeggiani and Bowman, 1963). Rat brain PFK is inhibited by most of the Krebs cycle intermediates, including cis-aconitate, isocitrate, malate, succinate, and α-ketoglutarate (Passonneau and Lowry, 1963). Tsai and Kemp (1974) observed that rabbit muscle PFK is more sensitive to citrate inhibition than the liver or brain PFK. The inhibition by citrate has been shown to be synergistic to that of ATP (Passonneau and Lowry, 1963). Mammalian PFK is also inhibited by several other phosphorylated intermediates, including phosphocreatine, 3-phosphoglycerate, 2,3-bisphosphoglycerate, and phosphoenolpyruvate (Bloxham and Lardy, 1973; Uyeda, 1979). The inhibition by these inhibitors is most pronounced in the
presence of an inhibitory concentration of ATP (Colombo et al., 1975).

The inhibition by ATP and other negative effectors can be counteracted by AMP, cAMP, ADP, Pi, F-6-P, and F-1,6-BP (Uyeda, 1979; Pilkis et al., 1981). F-1,6-BP is the strongest activator. The activation of PFK by positive effectors are synergistic. For instance, activation by F-1,6-BP is enhanced by AMP (Tornheim and Lowenstein, 1976). ADP alone is a weak activator and the presence of Pi greatly enhances the activation by ADP (Passonneau and Lowry, 1963).

In the early 1980s, a new regulator of glycolysis was discovered (Van Schaftingen et al., 1980; Claus et al. 1981; Uyeda et al., 1981). Van Schaftingen et al. (1980) found that β-D-fructose-2,6-bisphosphate (F-2,6-BP), a previously unknown metabolite, is a positive effector of PFK. Like other positive effectors, F-2,6-BP activates PFK by increasing its affinity for F-6-P and relieving the inhibitory effect of ATP. This effector is the most potent activator known. For the liver PFK, the concentration of F-2,6-BP needed to achieve a half-maximal stimulation is 1000 times lower than that of F-1,6-BP required to obtain the same effect (Van Schaftingen et al., 1980). The affinity of the rabbit muscle PFK for F-2,6-BP is approximately 10 fold higher than that for F-1,6-BP (Foe et al., 1983).

The level of F-2,6-BP has been shown to be regulated by hormones. Insulin and epinephrine, both of which stimulate lactate production, also increase the concentration of
F-2,6-BP. F-2,6-BP is synthesized and degraded by 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase (6PF-2-K/F-2,6-BPase), a bifunctional protein. The two reactions are catalyzed at two discrete active sites. 6PF-2-K/F-2,6-BPase is a dimeric protein and is regulated by reversible phosphorylation (Pilkis, 1990). The hormones regulate the F-2,6-BP level by phosphorylating or dephosphorylating the protein via a cyclic AMP mediated signal transduction pathway (Pilkis and El-Maghrabi, 1988).

**MAMMALIAN PFK EXISTS AS THREE ISOZYMES**

The possible existence of isozymes of mammalian PFK was first implicated in the aforementioned Tarui disease that is characterized by lack of muscle PFK activity (Tarui et al., 1965; Layzer et al., 1967). In patients with this disease, although the muscle PFK activity was reduced to zero, erythrocyte PFK activities were half of normal, and the PFK activities in white blood cells were unaffected (Layzer et al., 1967). Immunological studies showed that an antibody against human muscle PFK failed to inhibit erythrocyte PFK from these patients, while normal erythrocyte PFK activity was inhibited approximately 40% (Layzer et al., 1969; Tarui et al., 1969). These observations suggest that the enzymes from muscle and erythrocyte are products of different genes.

In rabbit tissues, there are three isozymes (Foe and Kemp, 1985; Gonzalez and Kemp, 1978; Tsai and Kemp, 1973; 1974; Kemp, 1971). The three isozymes purified from
different tissues were named as M from muscle, L from liver, and C from brain (Kemp, 1971; Tsai and Kemp, 1974; Foe and Kemp, 1985). The apparent monomer molecular weight of the L, M, and C subunits are 80,000, 85,000, and 87,500, respectively (Dunaway, 1983). The three isozymes are immunologically distinct. Antiserum against purified muscle PFK removed 98% activity from solution containing purified muscle PFK, but removed only partial activity (32%) of purified liver enzyme. Antiserum against purified liver PFK removed 96% activity of the purified liver PFK, but removed a mere 3% activity of the purified muscle enzyme. The combination of both antisera did not remove all the brain PFK activity, because the C type isozyme also is found in this organ (Tsai and Kemp, 1973). Antisera raised against muscle or liver PFK removed little purified PFK C activity (Foe and Kemp, 1985). When the protomers of acid dissociated muscle and liver PFK recombine at neutral pH, they form three electrophoretically distinct isozymes in addition to the two parent enzymes. This result agrees with the concept that muscle and liver PFKs are different proteins. The three hybrids are be M\(_3\)L, L\(_3\)M, and M\(_2\)L\(_2\) (Gozalez and Kemp, 1978).

In human and rat tissues, there are also three isozymes (Dunaway, et al., 1988; Dunaway and Kasten, 1987). In general, mammalian PFK exists as a homo- or hetero-tetramer of the three subunits, depending on the abundance of the three subunits in a given tissue (Dunaway, 1983).
MAMMALIAN PFKS ARE HIGHLY CONSERVED

The complete primary structures of the three isozymes of mammalian PFKs have been determined. The M isozyme has been cloned and sequenced from rabbit (Poorman et al., 1984; Lee et al., 1987), mouse (Gehnrich et al., 1988), and human muscles (Sharma et al., 1989). The L isozyme has been cloned and sequenced from rat (Hotta et al., 1991), mouse (Gekakis et al., 1989), and human (Levanon et al., 1989) livers. The C isozyme has been determined from rabbit (Li et al., 1994), and rat (Gekakis et al., 1994) tissues.

Comparison of the amino acid sequences of the three types of mammalian PFK isozymes indicates that all mammalian PFKs are highly homologous: there is more than 65% amino acid sequence identity among the cloned mammalian PFK isozymes (Gekakis et al., 1994). Interestingly, there is higher sequence identity between the same type of isozymes from different species than between two types of isozymes from the same species. For example, there is only 68.9% amino acid sequence identity between rabbit M and C type PFKs, while there is 90% sequence identity between the rabbit muscle and mouse muscle PFKs (Lee et al., 1987; Gehnrich et al., 1988; Li et al., 1994). Perhaps, different PFK isozymes evolved before the mammals diversified.

TISSUE SPECIFICITY OF THE PFK ISOZYMES AT THE PROTEIN LEVEL

The levels of the three PFK subunits in mammalian tissues have been studied by SDS-polyacrylamide gel electrophoresis
and immunoprecipitation with antisera raised against each type of subunit. The tissue-specific distributions of the three types of PFK isozymes in human, rabbit, and rat are summarized in Table 1. In skeletal muscle, only the M subunit has been detected in rabbit (Tsai and Kemp, 1973), rat (Dunaway and Kasten, 1987), and human (Dunaway et al., 1988). In liver, the L type subunit is the predominant form. But there are small amounts of the M type subunit in rabbit (Tsai and Kemp, 1973), and rat liver (Dunaway and Kasten, 1987), and small amounts of the M and C type subunits in human liver (Dunaway et al., 1988). The M type PFK is the major form in the heart of all three mammals, along with L and C type subunits (Tsai and Kemp, 1973; Dunaway et al., 1988; Dunaway and Kasten, 1987). In the brain and kidney of rabbit (Tsai and Kemp, 1973), rat (Dunaway and Kasten, 1987), and human (Dunaway et al., 1988), all three types of subunits co-exist. The three subunits have also been detected in other rat tissues, such as lung, thyroid, and spleen (Dunaway and Kasten, 1987).

**THE DISTINCT KINETIC PROPERTIES OF THE THREE MAMMALIAN PFK ISOZYMES**

Kemp (1971) observed that the rabbit muscle and liver PFKS are different kinetically. He found that the liver enzyme is more sensitive to ATP and 2,3-diphosphoglycerate inhibition, less sensitive to citrate and phosphoenolpyruvate (PEP) inhibition, and less sensitive to activation by AMP, ADP, and cAMP. He concluded from these results that liver
Table 1. Tissue-specific expression patterns of the three mammalian PFK isozymes in human, rabbit, and rat tissues.

<table>
<thead>
<tr>
<th>Isozyme</th>
<th>Human(^a)</th>
<th>Rabbit(^b)</th>
<th>Rat(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M L K H B</td>
<td>M L K H B</td>
<td>M L K H B</td>
</tr>
<tr>
<td>C-type</td>
<td>+ ++ ++ +++</td>
<td>++ ++ +++</td>
<td>++ ++ +++</td>
</tr>
<tr>
<td>M-type</td>
<td>+++ + ++ +++</td>
<td>+++ + ++ +++</td>
<td>+++ + ++ +++</td>
</tr>
<tr>
<td>L-type</td>
<td>+++ + ++ ++</td>
<td>+++ + ++ ++</td>
<td>+++ + ++ ++</td>
</tr>
</tbody>
</table>

---

* a: Dunaway et al. (1988).
M: muscle; L: liver; K: kidney; H: heart; B: brain.
++++: predominant (>80%); +++: major (50-70%); ++: abundant (10-40%); +: small amount (<10%).
PFK was less suited for anaerobic energy production than muscle PFK. Gonzalez and Kemp (1978) compared the kinetics of the M₂L₂ hybrid isozyme of rabbit PFK with those of the M₄ and L₄ isozymes. They found that ATP inhibition to the hybrid enzyme is intermediate between that observed with M₄ and L₄ and is clearly not identical to a simple summing of the effects of M and L subunits. The inhibitions to the M₂L₂ hybrid by citrate, 2,3-diphosphoglycerate, and phosphocreatine are also intermediate between those observed with M₄ and L₄.

The kinetic properties of the C type PFK remained elusive for a while, partly because it co-exists with the other two isozymes in a given tissue, such as brain, which makes it more difficult to purify. Almost a decade and a half after purification of the M and L isozymes, Foe and Kemp (1985) successfully purified PFK C from rabbit brain. They found that this isozyme is kinetically distinct from the M and L isozymes. The C isozyme is more sensitive than the M isozyme but less sensitive than the L isozyme to ATP inhibition. The inhibition by citrate is opposite: the C isozyme is less sensitive than the M isozyme but more sensitive than the L isozyme to citrate inhibition. PFK M is the most sensitive isozyme to activations by Pi, AMP, ADP and F-2,6-BP, while isozyme C is the least sensitive to these activators (Foe and Kemp, 1985). They again noted that the kinetic properties of PFK hetero-tetramers are probably influenced by interactions between the different subunits, since their
kinetics cannot be predicted from the properties of the homo-tetramers.

These distinct kinetic properties of the three mammalian isozymes probably evolved according to the physiological requirement of glycolytic regulation in a given tissue. For example, PFK M is the least sensitive to ATP inhibition but the most sensitive to positive effectors such as AMP, ADP, and F-2,6-BP. This fits well with the physiology of muscle, where ATP is produced and utilized for muscle movement. Hepatocytes have very different requirements since they must be capable of carrying out gluconeogenesis in order to maintain blood glucose levels. Citrate and other Krebs cycle intermediates are also precursors in gluconeogenesis. They are oxidized to yield oxaloacetate, which in turn is converted into phosphoenolpyruvate, an intermediate in gluconeogenesis. Therefore, the L type PFK, the least sensitive to citrate inhibition, allows glycolysis and gluconeogenesis to be carried out in the hepatocytes.

**STRUCTURE OF BACTERIAL PFKS**

The structure of the bacterial PFK is reviewed here for two reasons. First, the three-dimensional structure of mammalian PFK has not been determined. Second, mammalian PFKs probably evolved by gene duplication, because the N- and C-terminal halves of the rabbit muscle PFK are homologous to each other and to bacterial PFKs (Poorman et al., 1984; Lee et al., 1987). Poorman et al. (1984) proposed a model structure of the rabbit muscle PFK based on the
structure of bacterial PFK (I will come back to this in more
detail later).

PFK from \textit{Bacillus stearothermophilus} and that from
\textit{Escherichia coli} are highly homologous, with 53% amino acid
identity. The \textit{E. coli} PFK (EcPFK) shows sigmoidal kinetics
with respect to one substrate, F-6-P, but not to the other
substrate, ATP. It is allosterically inhibited by PEP, and
allosterically activated by ADP or GDP. The \textit{B. stearother-
mophilus} PFK (BsPFK) behaves similarly, except that the
kinetics with respect to F-6-P are sigmoidal only in the
presence of PEP. In the absence of PEP, it is hypobolic
(Valdez et al., 1989a).

The three-dimensional structures of EcPFK and BsPFK have
been determined using X-ray crystallography (Evans et al.,
1981; Shirakihara and Evans, 1988; Schirmer and Evans,
1990). The structure of the enzyme bound to substrate and
effector reveals the mechanism of the catalysis and the
effector binding. Comparison of the structures of the BsPFK
in the relaxed (R) and tight (T) states shows how the
conformation is changed before and after the allosteric
transition, which provides insight into the mechanism of the
allosteric behavior of PFK (Schirmer and Evans, 1990).

Both \textit{E. coli} and \textit{B. stearothermophilus} PFKs are tetramers
of identical subunits with 319 amino acids each. Each
subunit consists of two domains: a larger one which binds
the substrate ATP, and a smaller one which binds the other
substrate F-6-P. Each domain is a 3-layered \(\alpha\beta\alpha\) sandwich
structure. The tetramer is a dimer of dimers: two subunits pack together with a large interface to form dimers (subunits AB or CD in Fig. 3) and two dimers pack together with a smaller interface to form the tetramer. The tetramer exhibits 222 symmetry: the three molecular dyad axes are p, q, and r (Fig. 3).

The active site lies between the two domains of each subunit: the ATP binding site is formed almost entirely by residues of the large domain and the F-6-P binding site consists of residues mainly from the small domain (Shirakihara and Evans, 1988). The catalytic site is located between the two domains, while the effector site lies in the larger interface between the two subunits within each dimer (Schirmer and Evans, 1990).

**STRUCTURE DIFFERENCES OF BSPFK IN THE R AND T STATES**

When the structures of the R and T states are superimposed on their 222 axes, the major difference is the rotation of the AB dimer relative to the CD dimer by about 7° (Fig. 3). The rotation axis is close to the molecular p dyad (the symmetry axis of the dimers). Not much rotation is seen between the two subunits of the dimer, because of the much larger contact area between subunits within the dimer as compared to that between dimers. Although no significant change has been noted at the large interface of the two subunits within the dimer, an additional hydrogen bond exists between the subunits in the T state. Asp 59 interacts with the ribose of the effector ADP in the R structure, and
Figure 3. Diagram depicting the general view of BsPFK and its quaternary and tertiary structure changes on the allosteric transition from R to T. a. Schematic sketch of the R-state (dashed outline) and T-state (solid outline) PFK tetramer, viewed along p. The three mutually perpendicular 2-fold symmetry axes are labeled p, q and r. The subunits are labeled A-D. Residues labeled with * belong to subunit A, all others to subunit D. b, same as in a but viewed from r and additional presentation of the ligands of the R structure. (adapted from Schirmer and Evens, 1990). c, R-state structure, subunits A and D, related by r dyad, viewed as in b. helices are presented as cylinders numbered 1-13, β-sheet strands as arrows labeled A-K. In the active site the substrate F-6-P and the product ADP are shown, in the effector site the activator ADP (Shirmer and Evens, 1990). (Reprinted with permission from Nature 343, 140-145; Copyright (c) 1990 MacMillan Magazines Limited)
is hydrogen bonded to His 215 in the neighboring subunit. On the other hand, the interface between the two dimers triggers and accommodates their relative rotation in the allostERIC transition. On transition from the R to the T state, the rigid dimers undergo a contrary rotation around their common dyad axis. This motion couples the changes in the four binding sites for the cooperative substrate F-6-P, which lie between dimers. The rotation of the dimers is accommodated by removal of the layer of water molecules that lie between the dimers in the R structure, and by restructuring the 6-F loop (residues 155-162).

The rearrangement of the 6-F loop couples the binding of F-6-P to the quaternary structure change, and to the binding of heterotropic effectors (Fig. 4). When molecules with opposite allostERIC effects bind to the effector site, most of the site remains unchanged. Only the 8-H loop (residue 213-215) moves in response to a change of effector. This movement is coupled to the rearrangement of the 6-F loop. The low-affinity T-state conformation can only be formed when the lips of the effector site are closed. This closure is induced by inhibitors, but cannot occur when the larger activator ADP is bound, hence the opposite effects (Schirmer and Evans, 1990).

STRUCTURE OF MAMMALIAN PFKS

Each subunit of Mammalian PFK has a molecular mass of 85±5 kDa, about twice that of its bacterial counterpart (37 kDa). Examination of the amino acid sequence of rabbit
Figure 4. Schematic view of the 6-F loop (residues 155-162) of BsPFK showing its different interactions across the interface between dimers. The wavy line represents part of the boundary between r-dyad-related subunits (i.e., the boundary between dimers). a, T structure, showing Glu 161 in the F-6-P site, the Arg 72-Glu 241 salt bridge. b, R structure (with bound ligands ADP and F-6-P), showing the phosphate of F-6-P bound by Arg 162 and Arg 243. (Shirmer and Evens, 1990). (Reprinted with permission from Nature 343, 140-145; Copyright (c) 1990 MaMillan Magazines Limited)
muscle PFK has revealed that its N-terminal and C-terminal halves are homologous to each other (30% identity). The two halves of the rmPFK are also homologous to bacterial PFK: there are 42% and 33% identities between the amino acid sequence of BsPFK and those of the N- and C-terminal halves of rmPFK, respectively (Fig. 5). These data suggest that the mammalian enzyme evolved from a prokaryotic progenitor by gene duplication and divergence (Poorman et al., 1984; Lee et al., 1987). Based upon the sequence data, Poorman et al. (1984) proposed that the tertiary structure of each rmPFK subunit consists of two "super" domains, each closely resembling the structure of the BsPFK monomer, joined by a connecting peptide of about 30 residues. Therefore, a tetramer of mammalian PFK is equivalent to an octomer of bacterial PFK. Yet, there are only four F-6-P binding sites in a tetramer of mammalian PFK. In addition to the four F-6-P binding sites, there are four binding sites for F-1,6-BP or F-2,6-BP (Fig. 6), but none in bacterial PFK. This is probably the structural basis why the bacterial PFKs are not allosterically regulated by either F-1,6-BP or F-2,6-BP.

Homology in the vicinity of Asp-127 in BsPFK, a residue involved in F-6-P binding, is clearly evident in the N-terminal half of rmPFK but not in the C-terminal half, where the corresponding residue is a serine. Poorman et al. (1984) postulated that the F-6-P binding site in the C-terminal half of rmPFK has evolved as an effector site for either F-1,6-BP or F-2,6-BP. This seems reasonable as Asp-127 is
Figure 5. Sequence comparison of homologous PFKs and position of introns along the amino acid sequence of rabbit muscle PFK. The amino sequence of *E. coli* PFK (Ec PFK-1), *B. stearothermophilus* PFK Bs PFK), and the N-half [RM-PFK (N)] and C-half [RM-PFK (C)] of the rabbit muscle PFK are aligned for maximum homology. Double headed arrows above the sequences indicate secondary structural elements found in BsPFK. Introns (solid triangles in the N- and C- halves of the rabbit muscle PFK are positioned between the respective amino acid residues. Underlined regions are those deduced from DNA sequence of the rmPFK gene. (modified from Hellinga and Evens, 1985 with permission)
Figure 6. Location of substrate and effector sites in the hypothetic rmPFK protein model. Residues enclosed in solid boxes are identical to those at corresponding positions in BsPFK; those enclosed in dotted boxes are different (Taken with permission from Poorman et al., 1984)
probably required as a base catalyst acting to increase the nucleophilicity of the O-1 hydroxyl of F-6-P for attack on the  γ-phosphate of ATP. Its presence would cause electrostatic repulsion of F-1,6-BP or F-2,6-BP. Replacement of Asp-127 by Ser in the C-terminal half would remove this negative charge and provide more room for binding of the F-1,6-BP and F-2,6-BP, which are known to compete for a common site (Foe et al., 1983). Poorman et al. (1984) suggested that the replacement of a catalytically essential Asp-127 has led to loss of a catalytic site and creation of an allosteric site for F-1,6-BP or F-2,6-BP activation.

**MAMMALIAN PFK GENES**

The first mammalian PFK gene cloned was the rabbit muscle PFK (Lee et al., 1987). The organization of the rabbit muscle PFK gene is shown in Figure 7. This gene has 22 coding exons, spanning 17 kp (Lee et al., 1987). When overlaid on the structure of the protein, most of the introns are located between or near the ends of the secondary structural elements. Although the two halves of the PFK are homologous to each other in amino acid sequence, the introns located in the carboxyl-terminal half of the gene are not located at positions analogous to those in the amino-terminal half (Lee et al., 1987).

Subsequently to the cloning of the rabbit muscle PFK gene, the mouse liver gene (Rongnoparut et al., 1991) and human muscle gene (Yamasaki et al., 1991) have been cloned. Comparison of these three cloned mammalian PFK genes showed
Figure 7. Exon/intron organization of the rabbit muscle PFK gene and comparison of the 5' UTRs of two rabbit and three human muscle PFK cDNAs. a. The complete exon/intron organization of the rmPFK gene. Solid boxes: coding exon I through XXII; open boxes: 5' UTR. The 1.69 kb intron in rmPFK cDNA-A is indicated by the downward bent line. The open box with a broken end indicates a UTR, the 5' end of which is not known. b. Three cDNAs for the human muscle PFK. The dashed line and box indicate an intron with unknown length and 5' UTR (Modified from Li et al, 1990).
that the exon-intron organizations of these genes are similar with respect to the exon sizes and positions, with only one exception: exon 13 in mouse liver gene is three nucleotides (1 amino acid) shorter than those in rabbit and human muscle gene. However, the sizes of the introns in the three genes are not conserved.

**TWO CDNAS OF THE RABBIT MUSCLE PFK WITH DIVERGED 5' UNTRANSLATED REGIONS (UTRS)**

Previous work in the Chang Laboratory has identified two cDNAs for rmPFK (Li et al., 1990). These two cDNAs have identical coding regions, but different 5' untranslated regions (Figs. 7 and 8; Li et al., 1990). Alignment of the cDNA sequences with that of the genomic DNA showed that there is an intron spliced out in the 5' untranslated region of one of the two cDNAs (Fig. 7). These cDNAs could be derived from one promoter and alternative splicing of the pre-mRNA. Alternatively, they could be produced by two different promoters.

**EUKARYOTIC PRE-mRNA SPlicing IS FACILITATED BY SPliceosOME**

Since the two rmPFK cDNAs could be derived by alternative splicing, the splicing of pre-mRNA and subsequently the regulation of alternative splicing are reviewed here.

Splicing of eukaryotic precursor messenger RNA (pre-mRNA) is a two-step transesterification reaction which precisely removes the introns and joins exons. The first cleavage occurs at the 5' splice site, forming cleaved exon I and
Figure 8. Sequence diversity in the 5' UTRs of two rmPFK cDNAs. Bold prints indicate identical sequences. Plain prints are divergent sequence. "TATA" sequence is underlined. Coding sequences downstream of ATG are identical in the two rmPFK cDNAs. M, methionine; T, threonine; H, histidine; E, glutamic acid. (Modified from Li et al, 1990)
lariat intron-second exon intermediates. This is followed by a second cleavage at the 3′ splice site, the concomitant joining of the two exons, and the release of the intron in a lariat form (for recent reviews, see Green, 1991; Steitz, 1992; Moore et al., 1993).

The splicing of pre-mRNA is facilitated by the spliceosome, which is formed by the assembly onto the pre-mRNA of five small ribonucleoprotein particles (U1, U2, U4, U5, and U6 snRNPs). The assembly of spliceosome in mammalian cells depends on a consensus sequence at the 5′ splice site, a branch point sequence and the adjacent polypyrimidine tract, and the 3′ AG dinucleotide. The assembly of the spliceosome starts with the binding of U1 snRNP to the 5′ splice site of the pre-mRNA. This recognition is facilitated by base-pairing between the 5′ end of U1 snRNA and the 5′ splice site (Siliciano and Guthrie, 1980). The second step is the binding of U2 snRNP to the pre-mRNA branch site. Base-pairing between U2 snRNA and the branch site of pre-mRNA is involved in this recognition (Wu and Manley, 1989). The binding of U2 auxiliary factor (U2AF) to the polypyrimidine tract near the 3′ splice site is also required for this U2 snRNP-pre-mRNA interaction (Green, 1991). The final step of the spliceosome assembly is the entry of U4/U6/U5 tri-snRNPs (Steitz, 1992; Moore et al., 1993; Green, 1991). U5 snRNA interacts with sequences at both the 5′ and 3′ splice sites (Newman and Norman, 1992; Wassarman and Steitz, 1992). U6 snRNA has also been found to
interact with the intron sequence at the 5' splice site
(Wassarman and Steitz, 1992). Furthermore, the base-pairing
between the U6 snRNA and the 5' splice site sequence has
been shown to be important and necessary in the 5' splice
site selection (Lesser and Guthrie, 1993; Kandels-Lewis and
Seraphin, 1993).

Several non-snRNP proteins are also required in the pre-
mRNA splicing (Lamm and Lamond, 1993). The cytoplasmic
extract S100 was found to be incompetent in in vitro
splicing assay, although it contains all the snRNPs required
for pre-mRNA splicing. Addition of alternative splicing
factor/splicing factor 2 (ASF/SF2) is sufficient to render
the S100 extract splicing competent (Ge and Manley, 1990;
Krainer et al., 1990). ASF/SF2 is a member of the SR protein
family, which includes SRp20, ASF/SF2, SC35, SRp40, SRp55,
and SRp70 in mammals (Zuo and Manley, 1993; Ge and Manley,
1990; Ge et al., 1991; Krainer et al., 1990, 1991; Zahler et
al., 1992; 1993; Mayeda and Krainer 1992; Mayeda et al.,
1992, 1993; Fu et al., 1992). These proteins contain one or
more RNP-type consensus sequence (RNP-CS) and sequences of
consecutive arginine (R) and serine (S) dipeptides (RS
domain). ASF/SF2 and SC35 are involved in the commitment
complex (containing pre-mRNA and U1 snRNP) formation (Kohtz
et al., 1994; Fu, 1993), by binding to the 5' splice site
(Zuo and Manley, 1994). ASF/SF2 can bind to U1 snRNP, and
more specifically, to the integral U1 70k protein (Kohtz et
al., 1994; Wu and Maniatis, 1993). This protein-protein
interaction cooperates in the binding of U1 snRNP and ASF/SF2 to the pre-mRNA (Kohtz et al., 1994). ASF/SF2 and SC35 are also shown to interact with the 35K subunit of U2AF, which binds to the polypyrimidine tract near the 3' splice site. This interaction is thought to bridge the 5' and 3' splice sites (Wu and Maniatis, 1993).

**REGULATED ALTERNATIVE SPlicing IN HIGHER EUKARYOTES**

Utilizing alternative 5' or 3' splice sites results in alternative splicing of pre-mRNAs. Alternative splicing of many genes has been reported to play important roles in developmental and cellular processes of higher organisms (for recent reviews, see Maniatis, 1991; Green, 1991; Nadal-Ginard et al., 1991; Rio, 1992). For example, the sex determination of *Drosophila melanogaster* requires sex-specific alternative splicing of at least three genes that are involved in this process (Fig. 9; Baker, 1989). Another example is the alternative splicing of HIV RNA which plays important roles in HIV proliferation (Chang and Sharp, 1989). Finally, alternatively spliced CD44, a cell adhesion molecule, potentiates tumor metastasis (Gunthert et al., 1991; Rudy et al., 1993).

Sex-specific alternative splicing of several genes plays a pivotal role in *Drosophila* somatic sex determination (Fig. 9). These genes are *Sex-lethal* (*Sxl*), *transformer* (*tra*), and *double sex* (*dxs*). First, *Sxl* protein autoregulates its own splicing. In female flies, *Sxl* protein blocks the usage of the 3' splice site of intron 2 of *Sxl* pre-mRNA. This
Figure 9. Alternative splicing in *Drosophila* somatic cell sex determination. The exon-intron organizations and the splicing pathways of *Sxl*, *tra*, and *dsx* are diagramed in the middle. The female-specific pathway depicted at the top, while the default pathway at the bottom. The products of the female-specific splicing pathways are diagramed on the left and those of the default splicing on the right. *Sxl*: sex lethal; *tra*: transformer; *tra2*: transformer 2; *dsx*: double sex. Open box with number: exon; solid line: intron. At the bottom, female form of *dsx* protein represses the functions of the male differentiation genes, while the male form *dsx* protein represses the functions of the female differentiation genes.
leads to the splicing of exon 2 onto exon 4, resulting in Sxl mRNA with exon 3 skipped. In the male flies, exon 3 is retained by default and no functional Sxl protein is produced because there is a stop codon in exon 3 (Fig. 9). The alternative splicing of tra is also regulated by Sxl protein in a similar way and functional Tra protein is only produced in female flies (Fig. 9). Tra, along with the product of transformer 2 (tra2) gene, in turn regulates the alternative splicing of dsx. Unlike Sxl and tra, dsx encodes two forms of proteins. The female form is required to suppress male differentiation in the female flies, while the male form is required to suppress female differentiation in the male flies (Fig. 9). These two forms are generated by alternative splicing of its pre-mRNA by using alternative 3' splice sites. Tra and Tra2 proteins activate the female-specific splicing of dsx pre-mRNA by forming a protein complex on a splicing enhancer located 300 bp downstream of the female-specific 3' splice site (Tian and Maniatis, 1993).

A well studied pre-mRNA alternative splicing in mammalian cells is the splicing of human immunodeficiency virus type 1 (HIV-1) precursor RNA. In HIV infected cells, the full-length 9-kb viral transcript is spliced inefficiently to yield several classes of mRNAs that include the unspliced 9-kb mRNA and singly spliced 4-kb mRNAs (which encode structural viral proteins), and doubly spliced 2-kb mRNAs. The 9-kb and the 4-kb mRNAs contain functional introns and
splice sites, and their appearance in the cytoplasm is absolutely dependent on the viral trans-activator Rev protein (Chang and Sharp, 1989). Without Rev protein, the expression of the structural viral proteins (and hence the virus proliferation) is inhibited. Rev mediates its function through a cis-acting RNA sequence termed Rev-responsive element (RRE), which is located in the env gene of HIV-1 and in the unspliced intron of the 9-kb and 4-kb class mRNAs. Until recently, it is not clear whether Rev-mediated cytoplasmic expression of the unspliced, RRE-containing RNA is due to the inhibition by Rev of the splicing of RRE-containing intron, or due to the Rev-facilitated export of RRE-containing RNA, or both. Using an in vitro splicing system, Kjem et al. (1991) tested whether Rev inhibits pre-mRNA splicing. These researchers found that Rev inhibits pre-mRNA splicing in an RRE dependent manner. On the other hand, the splicing of RRE-containing pre-mRNA is not inhibited with the absence of Rev protein. Furthermore, Kjem and Sharp (1993) have showed that Rev inhibits RRE-containing RNA splicing by blocking the entry of U4/U6/U5 tri-snRNP during spliceosome assembly. On the other hand, Fisher et al. (1994) have demonstrated that Rev directly promotes the nuclear export of unspliced RNA. Therefore, Rev functions by both inhibiting the splicing and the export to cytoplasm of RRE-containing, unspliced or singly spliced HIV RNAs.
Several mammalian SR proteins have been shown to regulate pre-mRNA alternative splicing. ASF/SF2 and SC35 promote the use of the proximal 5' splice site among two competing alternative 5' splice sites, while SRp40, SRp55, and SRp70 promote the use of the distal 5' splice site in vitro (Zuo and Manley, 1993; Ge and Manley, 1990; Ge et al., 1991; Krainer et al., 1990; Mayeda and Krainer, 1992; Mayeda et al., 1992, 1993; Fu et al., 1992; Zahler et al., 1993). The RS domain of ASF/SF2 was found essential for its splicing factor activity but not for splice site switching (Zuo and Manley, 1993). ASF/SF2 and SC35 also favor the proximal site in a pre-mRNA containing duplicated 3' splice sites (Fu et al., 1992).

In addition to SR proteins, another cellular RNA binding protein, the heterogeneous nuclear ribonucleoprotein A1 (hnRNP A1), has also been found to regulate pre-mRNA alternative splicing. This protein promotes the use of the distal 5' splice site, which is antagonistic to ASF/SF2 (Mayeda and Krainer, 1992).

**MULTIPLE PROMOTERS OF PROTEIN ENCODING GENES**

The two rmPK cDNAs (Li et al., 1990) could be produced by two separate promoters. Therefore, studies on multiple promoters of eukaryotic genes are reviewed here.

Most mammalian genes encoding proteins are transcribed from one promoter. However, an increasing number of genes have been reported to have more than one promoter (for review, see Schibler and Sierra, 1987). For example, the
gene encoding the 6-phosphofructo-2-kinase/Fructose-2,6-bisphosphatase (6PF-2-K/F-2,6-BPase) is transcribed by two tissue-specific alternative promoters: a downstream promoter transcribes the gene in the liver, while an upstream promoter expresses the gene in the muscle (Darville et al., 1989). F-6-P-2-K/F-2,6-BPase catalyzes the synthesis and degradation of F-2,6-BP, a very potent activator of mammalian PFK. Human aldolase A gene, whose product is also involved in glycolysis, is transcribed by three alternative promoters (Gautron et al., 1991). Mouse γ-glutamyltransferase gene expresses its six mRNAs from five promoters, three of which are kidney specific (Rajagopalan et al., 1993; Sepulveda et al., 1994). The human carbonic anhydrase I gene contains two promoters with different tissue specificities (Brady et al., 1991). Multiple promoter usage has also been observed in mouse glucocorticoid receptor gene (Strahle et al., 1992), rat dopamine-D2-receptor gene (Valdenaire et al., 1994), and rat liver alkaline phosphatase gene (Toh et al., 1989).

What is the advantage of having multiple promoters for a gene? In transcribing a single gene from multiple promoters, an organism gains additional flexibility in controlling the expression of the gene. In higher eukaryotes, multiple promoters are usually associated with tissue- and/or developmental stage-specific expression of the genes. As mentioned above, functional Sxl protein is required to regulate the alternative splicing of its own and tra
pre-mRNA splicing and to produce functional Sxl and Tra proteins only in female flies. Then, where does the initial functional Sxl protein come from in the female but not in the male flies? Keyes et al. (1992) showed that the Sxl gene actually has two promoters, Pe (for embryonic or establishment promoter) and Pm (for maintenance promoter). Pe is located within the intron of the Sxl-Pm transcription unit (Keyes et al., 1992). Pe is active in early embryos of the 2X/2A female but not in those of the 1X/2A male flies. In addition, the transcripts produced from Pe are spliced directly from the embryonic exon 1 to exon 4, skipping exons 2 and 3 (exon 3 has a stop codon, Fig. 9). Except the N-terminal 25 aminal acids, the early transcripts encode proteins identical to the late female transcripts. These embryonic Sxl proteins direct the female-specific splicing of the first transcripts expressed from the late, Sxl maintenance promoter (Estes et al. 1995).

There are several mechanisms that may have produced a gene with multiple promoters during evolution. First, a gene may acquire a second promoter via transposable elements or retroviral integration. For example, promoter insertion via retroviral integration has been shown to be responsible for the activation of cellular oncogenes (Hayward et al., 1981). Second, an additional promoter may be created by inversion of DNA sequence. It has been shown that the Antennapedia (Antp) phenotype (conversion of antennae into legs) in Drosophila resulted from a promoter insertion via inversion
(Fisher et al., 1986). Upstream of Antp (locus 84B) there is a transcription unit belonging to locus 84C, which is active in the head. Structural analysis of the mutant allele of Antp revealed an inversion of a DNA stretch of the 84C promoter. As a consequence, Antp transcription is brought under control of the head-specific 84C promoter, leading to Antp expression in the head. Another way to create tandem promoters may be the duplication of a promoter followed by diversification. Such example has been reported for the gene encoding pancreatic amylase in mouse strain CE (Bodary et al., 1985). Finally, multiple promoters may also be created by deletion. For example, in Drosophila melanogaster, the ADH gene contains a single gene with two promoters. In D. muelleri, however, there are two functional ADH genes in tandem, each with one promoter (Fischer and Maniatis, 1986). The D. melanogaster ADH gene and its two promoters may be produced by the deletion of the coding region of the upstream gene.

Having multiple promoters is sometimes not conserved among different species. For example, the human aldolase A gene has three promoters. However, there is no mouse counterpart for the most upstream promoter of the human gene, although homology exists in the coding regions for the human and mouse genes (Stauffer et al., 1990; Gautron et al., 1991). In rat, the aldolase A gene also lacks the most upstream promoter (Mukai et al., 1991). The most upstream promoter of the human gene may be acquired in human or
deleted in mouse and rat after humans and rodents diversified.

The regulation of a gene is more complex when the gene has multiple promoters. One aspect of this complexity is that the multiple promoters share an enhancer. In other words, the same enhancer activates the transcription from more than one promoters. For example, the most upstream, constitutive promoter of the human aldolase A gene shares a muscle-specific enhancer with the muscle-specific promoter in the muscle cells, but shares a ubiquitous enhancer with the most downstream, constitutive promoter in other cells (Concordet et al., 1993). On the other hand, having multiple promoters might result in one promoter interfering with the expression of the other promoter(s). The Drosophila melanogaster ADH gene is such an example. As mentioned above, this gene has two promoters. The proximal promoter transcribes the gene primarily in first to early third-instar larvae, while the upstream promoter is active in late third-instar larvae and adults (Corbin and Maniatis, 1989). Each promoter has its own enhancer, but the ADH larvae enhancer and the proximal promoter are separated by the ADH adult enhancer and the distal promoter. When the distal promoter is active, the proximal promoter is shut off. If the distal promoter is deleted, the proximal promoter is active in first instar larvae to adults. Therefore, the developmental switch from the proximal to distal promoter is regulated by the stage-specific activation of the distal
promoter, and the subsequent repression of the proximal promoter by transcriptional interference (Corbin and Maniatis, 1989).

Another example of promoter switch during development is the human fetal (γ-) -to-adult (β-) globin gene switch. I include this mammalian example here, although the two promoters transcribe not the same, but two closely linked genes. When either the human γ- or β-globin gene is linked separately to the β-globin gene locus activation region (LAR), both genes can be expressed in transgenic mice throughout development. However, when both genes are linked to the same LAR, the normal developmental switch of the human globin gene expression is reproduced in the transgenic mice (Enver et al., 1990). Therefore, the human fetal-to-adult globin switch is controlled through a mutually exclusive interaction between LAR and either the γ- or β-globin gene, resulting in the expression of only one of these two genes at any given moment (Enver et al., 1990). Although the detailed mechanism of this mutually exclusive expression of only one of these globin genes is not clear, several adult stage-specific transcription factors (NF-E4 and bCTF) have been shown to interact with the promoter region of the adult gene to subsequently activate this gene and indirectly repress the expression of the fetal globin gene (Foley and Engel, 1992). Another factor is human Erythroid Kruppel-like factor (EKLF), containing zinc finger domains similar to the Kruppel protein of Drosophila melanogaster.
This protein has been shown to bind to the CACCC element in human β-globin gene promoter and activate the transcription of the β-globin gene (Donze et al., 1995). This factor activates the adult gene 1000-fold, while it activates the fetal gene only 3-fold. This is because the EKLF protein binds to the CACCC box of the adult gene 8-fold more efficiently than to that of the fetal gene. In addition, this factor is expressed at a 3-fold higher level in adult erythroid tissue than in fetal erythroid tissue (Donze et al., 1995).

OBJECTIVES AND RATIONALE

In this dissertation research, I have set forth three objectives:

The first objective of this study is to characterize the 5' ends of the mRNAs of the rmPFK gene.

The rabbit muscle PFK gene (Lee et al., 1987) and then its two cDNAs (Li et al., 1990) were previously cloned. Yet, whether the two cDNAs are derived from two separate promoters or from one promoter and alternative splicing remains elusive. To further study the regulation of this gene, it is important to characterize the 5'-ends of its mRNAs and the transcription start sites.

Three methods will be used to accomplish this goal: (1) anchored polymerase chain reaction (PCR) to directly clone and sequence the 5'-end of the rmPFK mRNA; (2) primer extension; (3) S1 mapping.
There are advantages and drawbacks for each method. In primer extension and anchored PCR, secondary structure in the RNA template might cause reverse transcription to be prematurely terminated. In this case, the end of the cDNA extended from the primer would be shorter than the actual 5' end of the mRNA. 5' mapping and RNase protection assays are similar in principle: a labeled DNA or RNA probe is hybridized to the mRNA, and the 3' overhang of the probe is digested by 5' nuclease or RNase, while the double stranded mRNA-probe is protected from 5' or RNase digestion. The size of the protected probe will indicate the 5'-end of the mRNA that hybridized to the probe. However, if there is an intron spliced out in the mRNA but not in the probe, the splicing junction instead of the 5' end of the mRNA will be mapped. Since there are intrinsic problems with each method, combinations of two or three of the methods will be used to map the 5' end of the mRNA.

The second objective is to study the tissue-specificities of the different rmPFK mRNAs.

As mentioned above, the rabbit muscle PFK is highly expressed at the protein level in skeletal muscle, brain, heart, kidney, but at very low levels in liver. Since there are at least two rmPFK mRNAs, it is important to understand the tissue-specificities of different rmPFK mRNAs. Northern blot analysis and reverse-transcription (RT)-PCR will be employed to accomplish this objective. Both methods are frequently used in studying gene expression at the mRNA
level. The sensitivity of the latter method is higher than that of the former. In addition, the RT-PCR will detect alternatively spliced mRNAs.

The third objective is to optimize conditions for primer extension using a thermostable DNA polymerase.

During the course of the study, I found that the secondary structure in the rmPFK mRNAs causes the reverse transcription to prematurely terminate in primer extension and anchored PCR. Higher temperatures of the reverse transcription reaction destabilize the secondary structure and hence minimize the problem. Conventional reverse transcriptase, such as M-MLV reverse transcriptase or AMV reverse transcriptase, can only tolerate a temperature up to 50°C. At 50°C, the secondary structures in the rmPFK still cause problems. Meyers and Gelfand (1991) reported that a 94-KDa DNA polymerase from *Thermus thermophilus* possesses reverse transcriptase activity at temperatures as high as 75°C. However, the reverse transcription reaction is not efficient enough for primer extension analysis. I therefore will optimize the experimental parameters for the reverse transcription reaction using this thermostable enzyme (Section II).

This improved method will be used in the reverse transcription reactions in the primer extension and anchored PCR experiments characterizing the rmPFK mRNAs described in SECTION I of this dissertation.
SECTION I
TISSUE-SPECIFIC EXPRESSION OF SIX TYPES OF RABBIT MUSCLE PHOSPHOFRUCTOKINASE mRNAs BY MULTIPLE PROMOTERS AND ALTERNATIVE SPLICING
INTRODUCTION

Phosphofructokinase (ATP:Fructose-6-phosphate 1-phospho-
transferase, EC 2.7.1.11, PFK) catalyzes the phosphorylation
of fructose 6-phosphate to produce fructose 1,6-
bisphosphate, the rate-limiting step in glycolysis. PFK
controls glycolytic flux through its allosteric responses to
cellular metabolites. For example, mammalian PFK is
inhibited by high concentration of ATP and citrate. On the
other hand, it is activated by ADP, AMP, F-1,6-BP, and F-
2,6-BP (reviewed in Dunaway, 1983; Uyeda, 1979). In
mammalian tissues, PFK can exist as homo- or hetero-
tetramers of three types of subunits: M (muscle) type, L
(liver) type, and C for a third type of subunit in brain.
For example, in human, rabbit, and rat, the M type PFK is
the predominant isozyme in skeletal muscle, while the L type
is the most abundant form in liver. In heart, brain, and
kidney, PFK exists as homo- or hetero-tetramers of all three
subunits (Tsai and Kemp, 1973; Dunaway and Kasten, 1987;
Dunaway et al., 1988). The homo-tetramers of M, L, and C
type subunits of the rabbit PFK have been shown to possess
different kinetic properties (Foe and Kemp, 1985). Among the
three rabbit isozymes, the M PFK is the least sensitive
isozyme to ATP inhibition, while the L PFK is the most
sensitive one. To activation by Pi, AMP, ADP, and F-2,6-BP,
the M isozyme is the most sensitive one, while the C isozyme
is the least sensitive one (Foe and Kemp, 1985). The
relative levels of these three isozymes in a given tissue determine the composition of PFK tetramers and subsequently their allosteric behavior in that tissue (Dunaway and Kasten, 1987; Dunaway et al., 1988). Thus the molecular basis for the regulation of the rate of glycolytic flux lies in the control of the tissue-specific expression of the genes encoding the three types of PFK subunits.

To study the regulation of tissue-specific expression of the mammalian PFK gene, the gene for rabbit muscle PFK was cloned (Lee et al., 1987). More recently, two cDNAs of this gene with different 5' untranslated regions (UTRs) were cloned. The 5' UTR of one cDNA has an intron spliced out (Li et al., 1990). In this study, more cDNAs for the 5' UTRs of the rmPFK gene were cloned using anchored polymerase chain reaction (PCR) (Loh et al., 1989) and reverse transcription-PCR (RT-PCR) (Meyers and Gelfand, 1991). The cloning results indicate that there are six types of rmPFK mRNAs that differ in the 5' UTR. These rmPFK mRNAs are shown to be expressed in tissue-specific manners by at least three promoters and by alternative splicing. Further the transcription start sites of two promoters of this gene have been determined. Altogether, the results of this study indicate that the rabbit muscle PFK gene expresses its six mRNAs tissue-specifically by multiple promoters and alternative splicing.
MATERIALS AND METHODS

Vectors, bacterial strains, and enzymes

Restriction endonucleases, DNA ligase, and DNA kinase were obtained from GIBCO-BRL Co., New England Biolabs, or Boehringer Mannheim Corp. M-MLV reverse transcriptase (Superscript) was purchased from GIBCO-BRL Co. Taq and Tth DNA polymerases were purchased from Perkin Elmer Cetus Corp. S1 nuclease was purchased from Promega. Bluescript II KS(+) plasmid, helper phage and the XL1-Blue strain of *E. coli* were from Stratagene. The Sequenase® kit and terminal deoxynucleotidyl transferase were purchased from United States Biochemical Corp. Nucleotides labeled with $^{32}$P or $^{35}$S were obtained from duPont-NEN or ICN Biochemicals.

Oligonucleotides used in this study

The oligodeoxyribonucleotides (oligos) used in this study (listed in Table 2) include: (1) primers used in anchored PCR; (2) probes used in Northern blot analysis; (3) primers used in RT-PCR studies; and (4) primers used in primer extension studies. The oligos for the rabbit muscle PFK gene are named from A to O. The lowercase x in the name denotes an XbaI restriction site at the 5' end of the primer. The oligos were synthesized in house and purified from polyacrylamide-urea gel (Sambrook et al., 1989). Oligo(dT)-cellulose (Type III) was purchased from Collaborative Research, Inc.
Table 2. Oligonucleotides used in this study

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence (5' ----&gt; 3')</th>
<th>Position*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ax</td>
<td>GATCCACTCTAGAGGCAAGCGATGGTCAGTTAGG</td>
<td>-36, M</td>
</tr>
<tr>
<td>Bx</td>
<td>GATCTAGACCTAAATACCTTGTTGAGGA</td>
<td>-1715, M</td>
</tr>
<tr>
<td>C</td>
<td>GGCTGCATGGCTCTCTGACGATGGTCAGTTAGG</td>
<td>+3, M</td>
</tr>
<tr>
<td>Cx</td>
<td>GGATCTCTCTAGAGGCTGACATGGTCAGTTAGG</td>
<td>+3, M</td>
</tr>
<tr>
<td>Dx</td>
<td>GGGCTCTAGAGGCCGCCGCCTTTGCAGTTAGG</td>
<td>+31, H</td>
</tr>
<tr>
<td>E</td>
<td>CACCCGCTGGCCCTGGCCGACCCCCAGGGT</td>
<td>+31, M</td>
</tr>
<tr>
<td>F</td>
<td>AAGGCAAGACCTTGTCACTGAGGAACGGCTTAAGGACACA</td>
<td>-41, M</td>
</tr>
<tr>
<td>G</td>
<td>GGTGGGACTTGAACCTGCTCTTTATGACAC</td>
<td>-1736, M</td>
</tr>
<tr>
<td>H</td>
<td>CTCTGTTGCCCTTTGCAGGAGACTCTG</td>
<td>-1973, M</td>
</tr>
<tr>
<td>I</td>
<td>GTCTGCCATATCGGCTGACCTGGGAACA</td>
<td>-2570, M</td>
</tr>
<tr>
<td>J</td>
<td>GCCGCCCTTGCGCTCGTACCCATCCACAGCCT</td>
<td>-2845, M</td>
</tr>
<tr>
<td>K</td>
<td>CTAACCTGACATCGCTCTGGCCTTT</td>
<td>-13, P</td>
</tr>
<tr>
<td>L</td>
<td>CAGACTCCTGCGCCAGCCACAGGCAACAGG</td>
<td>-1949, P</td>
</tr>
<tr>
<td>M</td>
<td>GGGCTCTGCTGTTCCACCCAGGATGATT</td>
<td>-2554, P</td>
</tr>
<tr>
<td>N</td>
<td>CTGGGACTGACGGGACGCGGAGGGCCTT</td>
<td>-2818, P</td>
</tr>
<tr>
<td>O</td>
<td>CTGCCATATGACCGGCTTTATACACACA</td>
<td>+340, M</td>
</tr>
<tr>
<td>GAPDH1</td>
<td>CATGTAGACATGTGAGGGGATCTTCATG</td>
<td>M</td>
</tr>
<tr>
<td>GAPDH2</td>
<td>GATCCATTCATCGACCTCCACATACA</td>
<td>P</td>
</tr>
<tr>
<td>GAPDH3</td>
<td>GATTCGCTCCCTGGAAGATGGTGAT</td>
<td>M</td>
</tr>
</tbody>
</table>

The lower case x in the name of the oligo represents an XbaI restriction site at the 5' end. This XbaI restriction site is underlined in the sequence of primer Ax, Bx, Cx, or Dx. Oligo C and Cx are essentially the same, except that Cx has an XbaI site at the 5' end.

*: The number is the position of the nucleotide at the 3' end of each rmPFK primer or probe relative to the ATG start codon, with A being +1. P or M following the position indicates the oligo is located on the plus or minus strand, respectively.
Preparation of poly(A)$^+$ RNAs

Total RNAs from skeletal muscle, liver, kidney, heart, and brain of 10 day-old New Zealand white rabbits were extracted by homogenizing the tissue in guanidinium isothiocyanate, followed by cesium chloride gradient centrifugation (Chirgwin et al., 1979). Poly(A)$^+$ RNA was purified from total RNA by oligo(dT) cellulose column chromatography (Sambrook et al., 1989). The concentration of poly(A)$^+$ RNA was determined by measuring the absorbance at 260 nm ($A_{260}$).

Anchored PCR

Anchored PCR (Fig. 10) was performed as described (Loh et al., 1989) with the modifications given below.

Reverse transcription

Poly(A)$^+$ RNA (8 µg) was co-precipitated with 25 ng of primer C or E (Fig. 10), dissolved in 20 µl of buffer (80 mM Tris, pH 8.3; 80 mM KCl), heated at 90°C for 2 min., and allowed to anneal at 52°C for 3 hr. After annealing, 30 µl of reverse transcription mixture (20 mM DTT; 0.1 mg/ml bovine serum albumin [BSA]; 0.5 mM each of dATP, dGTP, dCTP, and dTTP; and 20 units of RNAsin [Promega]) was added to the annealing mixture. Reverse transcription was carried out in the presence of 0.03 µCi of [$\alpha^{-35}$S]dATP and 200 units of M-MLV reverse transcriptase (Ausubel et al., 1993) at 42°C for 1 hr.
Figure 10. Strategy of anchored PCR used to clone the 5' UTRs of different rmPFK mRNAs. (A) Schematic of the genomic DNA contains the 5' flanking region and the first coding exon of the rmPFK gene. (B) Diagram depicting the scheme of anchored PCR to clone the 5' portion of rmPFK cDNAs that contain the 5' UTRs of cDNA-A and cDNA-B (Li et al., 1990). Primer C was annealed to rabbit poly(A) RNA and extended with reverse transcriptase. The cDNAs were tailed with dGs with terminal deoxynucleotidyl transferase. The tailed cDNAs were amplified with primer Ax or Bx and Dx. Primers Ax and Bx are specific to the 5' UTRs of cDNA-A and cDNA-B, respectively. Primer Dx is an oligo with 15 consecutive dCs and a XbaI restriction site at the 5'. The sequences of the primers are shown in Table 2. Arrows: primers used in anchored PCR; X: XbaI restriction sit; open box: 5' UTR of cDNA-A or cDNA-B; dotted box: the first coding exon. (C) Anchored PCR strategy used to clone the 5' UTRs of different kinds of mRNAs transcribed by all the alternative promoters of the rmPFK gene. In this experiment, reverse transcription was carried out using primer E and rabbit muscle poly(A) RNA. The oligo dG tailed cDNA products were amplified with primers Cx and Dx. Primer Cx is primer C with a XbaI site at the 5' end. Open box with broken end: 5' UTR of all kinds mRNAs of the rmPFK gene.
Tailing the 3' ends of cDNAs with dGs by terminal deoxynucleotidyl transferase

The unincorporated deoxynucleoside triphosphates remaining from the previous step were removed using a Sephadex G-50 column (5 Prime to 3 Prime Co.). Fractions included in the first radioactive peak were pooled and precipitated with 2.5 volumes of ethanol in the presence of 0.3 M sodium acetate. The purified DNA-RNA duplex was dissolved in the tailing buffer (0.1 M sodium cacodylate, pH 7.1; 2 mM MnCl\textsubscript{2}; 0.1 mM DTT; 80 \mu M dGTP; 0.05 mg/ml BSA) (Deng and Wu, 1983). The dGTP polymerization reaction was carried out by the addition of 34 units of terminal deoxynucleotidyl transferase and incubation at 30°C for 1 hr. The products were precipitated once with ethanol and redissolved in 20 \mu l of TE (10 mm Tris, pH 7.6; 1 mm EDTA).

PCR amplification

One third of the DNA products generated from the tailing process was amplified by PCR (Saiki et al., 1988) using an rmPFK specific primer and primer Dx containing 15 C's and a XbaI site at the 5' end (Fig. 10; for sequence see Table 2). The conditions for denaturation, annealing and polymerization were 95, 55 and 72°C for 1.5, 1 and 2 min, respectively. The amplifications were carried out for 30 cycles.
Cloning and sequencing of the PCR products

These PCR products were inserted into the XbaI site of Bluescript plasmid which was subsequently transformed into the XL1-Blue strain of *E. coli*. Single-stranded DNA templates were prepared by coinfecting the cells with VCS-M13 helper phage according to the manufacturer's instructions. DNAs were sequenced by the dideoxynucleotide chain termination method (Sanger et al., 1977) using Sequenase kit (USB).

SI mapping

SI mapping was carried out as described (Ausubel et al., 1993). Briefly, the 5' labeled oligo F or G (Table 2) was annealed to the single-stranded DNA of plasmid RMPFK5'. This plasmid was constructed by inserting the 5' flanking region of the rmPFK gene into the SstI site of pBluescript vector. The annealed oligo was extended using Sequenase (USB). The single stranded probes were purified from an alkaline agarose gel. $2 \times 10^5$ CPM (0.2 pmole) of each probe was hybridized with 8 µg poly(A)$^+$ RNA in the aqueous hybridization buffer overnight at 50°C. SI nuclease (150 units per 300 µl reaction) digestion was carried out at 30°C for 1 hr. Final products were analyzed on a sequencing gel. The size marker is the sequencing ladder using oligo F or G as primer, and plasmid RMPFK5' as template, both of which are used to generate the single-stranded SI mapping probe.
Primer extension

Primer extension was performed as described (Ausubel et al., 1993) with 8 μg poly(A)^+ RNA from rabbit muscle and 2x10^5 CPM (0.2 pmole) of 5' end labeled primer. Reaction products were resolved on 6% polyacrylamide gels containing 8 M urea.

Primer extension and anchored PCR using thermostable Tth DNA Polymerase

Primer extension by Tth DNA polymerase was carried out as described below. The hybridization mixture (12 μl) containing 16.7 mM Tris-HCl (pH 8.3), 150 mM KCl, 2x10^5 CPM (0.2 pmole) of primer B and 8 μg of rabbit poly(A)^+ RNA was heated at 75°C for 5 minutes to denature the RNA. After annealing at 65°C for 5 minutes, extension mixture (8 μl) containing 2.5 mM MnCl2, 500 μM each of dATP, dGTP, dCTP, and dTTP, and 5 units of Tth DNA polymerase was added to the hybridization mixture. 20 μl of mineral oil was overlaid on top of the reaction mixture. Primer extension was carried out at 65°C for 1 hour. All the incubations were carried out in a Perkin Elmer Cetus thermocycler. After the extension, the mineral oil was carefully removed and residual oil was further depleted by phenol/chloroform extraction.

Along with 20 μl of H2O, 105 μl of RNase reaction mixture (100 μg/ml salmon sperm DNA, 20 μg/ml DNase-free RNase A,
Ausubel et al., 1993) were added to each of the primer extension reaction tubes. RNase digestion was carried out at 37°C for 15 minutes. The mixture was extracted with phenol/chloroform/isoamyl alcohol and precipitated and analyzed as described (Ausubel et al., 1993)

Anchored PCR using Tth DNA polymerase was carried out as described above using Superscript reverse transcriptase, except that Tth DNA polymerase instead of the Superscript reverse transcriptase was used during the reverse transcription step.

**Northern blot hybridisation**

Northern blot hybridization was carried out as described (Sambrook et al., 1989). Briefly, 8 µg of poly(A)+ RNA from each rabbit tissue was fractionated on a formaldehyde denaturing agarose gel (1%) with 1 X MOPS buffer (20 mM MOPS, pH 7.0; 8 mM NaOAc; 1 mM EDTA). The RNAs were blotted overnight onto Nytran membrane (0.45 µm). Oligonucleotide probes were end-labeled using T4 polynucleotide kinase and [γ-32P]ATP (Sambrook et al., 1989). Filters were prehybridized in hybridization solution (6 X SSPE; 10 X Denhardt’s; 50 µg/ml tRNA; 0.1% SDS; 0.05% NP-40) at 55°C for 2 hrs and for an additional 20-22 hrs after adding end-labeled probes (2 X 10^6 CPM/ml, 10^6 CPM/pmole). Hybridization was carried out in an Autoblot Micro Hybridization Oven (Bellco Glass).
Analysis of the six rmPFK mRNA expression in rabbit tissues by RT-PCR technique

RT-PCR was performed as described (Meyers and Gelfand, 1991), except that the reverse transcription reaction was carried out at 65°C for 10 min instead of at 70°C for 5 min.

RESULTS

Cloning and sequencing of cDNAs for the 5' UTRs of the rmPFK mRNAs by anchored PCR

Li et al. (1990) previously cloned two cDNAs of the rmPFK gene with divergent 5' UTRs but with identical coding regions. Alignment of the sequences of these 5' UTRs with the 5' flanking genomic sequence of the rmPFK gene revealed that cDNA-A has a 1.69 kb intron removed by splicing in the 5' UTR (Fig. 10B; Li et al., 1990).

Theoretically, cDNA-A and cDNA-B could be derived from two different promoters or from one promoter by alternative splicing. If they result from alternative splicing, the mRNAs should share a common 5' exon. In order to clone and sequence the 5' exons of the rmPFK mRNAs, anchored PCR (Loh et al., 1989) was used.

The anchored PCR is described in detail in MATERIALS AND METHODS and outlined in Fig. 10 in this SECTION. When primer Ax (Fig. 10B) was used, six clones were sequenced. Among them, three clones are identical and are therefore interpreted as one cDNA (designated as cDNA-1). The sequence of this cDNA is shown in Figure 11. The 5' end of this cDNA is located at -86 (The A of the ATG translation initiation
Figure 11. Sequences of rmPFK cDNAs cloned by anchored PCR and RT-PCR. The ATG translational start codon is in bold. The TATAAAA sequence is underlined. Proposed splice junctions are indicated by vertical lines.
codon is designated as +1). This cDNA is diagramed in Fig. 12. The 5' ends of the other three clones (sequence not shown) are 10 bp or less downstream of cDNA-1. With primer Bx, six clones were sequenced. Among them, five clones are identical in sequence and are also considered as one cDNA (designated cDNA-2, Fig. 11 for sequence). The 5' end of cDNA-2 is located at -1758. The sixth clone is 5 bp shorter than cDNA-2 (not shown). The 5' end of cDNA-2 is 45 bp shorter than that of cDNA-A of Li et al. (1990). SI mapping results (described later, Fig. 17) indicate that cDNA-2 resulted from premature termination of the reverse transcription, probably due to a strong secondary structure in the rmPFK mRNA template. Using a thermostable reverse transcriptase, several cDNAs containing the entire 5' UTR of cDNA-2 were cloned (see below). Therefore, cDNA-2 is not considered a full length cDNA of a rmPFK mRNA and is hence not diagramed in Fig. 12.

In order to identify other possible rmPFK mRNA whose 5' UTR is not complementary to either primers Ax or Bx, a different strategy was utilized (Fig. 10C). Primer E, which is downstream of primer C, was used to perform the reverse transcription. The dG-tailed cDNA products were then amplified by PCR using primer Cx and primer Dx. Primer Cx is complementary to the first 9 codons of the rmPFK open reading frame, while primer Dx is an oligo of 15 dCs with a XbaI site (Fig. 10C). In theory, any mRNA containing the
Figure 12. Schematic diagram depicting the organization of the 5' portion of the rmPFK gene and the six mRNAs identified by anchored PCR and RT-PCR. (A) Organization of the rmPFK gene. Boxes numbered IA, IB, IC, ID, and IE represent the exons in the 5' UTRs of the gene. Exon I is the first coding exon. There is no splicing event between exon IA and exon II. Only shown are the exons in the 5' UTR and the first 5 coding exons. Bent arrows indicate transcription start sites. Only start sites a and b have been established. The bent arrow in front of exon ID and IE are putative promoters, their transcription start sites have not been mapped. The positions of the oligo probes used in the Northern blot are shown under each exon. The sequences of these probes are shown in Table 2. TATA: the TATAAAA sequence found in cDNA-A (Fig. 10; Li et al., 1990); ATG: initiation codon. (B) Schematic diagram depicting the six cDNAs identified by either anchored PCR (cDNA-1, -3, -4, -5, and -6) or RT-PCR (cDNA-7). Downward bent solid line: intron with known sequence. The primers used in the RT-PCR experiments are shown on top or below the cDNA. Oligo O represents the 3' primer and is complementary to the fifth coding exon. The sequences of the primers are shown in Table 2.
first coding exon of the rmPFK gene should be cloned by this strategy. Sequence analysis of eighteen clones containing anchored PCR products (not shown) confirmed the 5' UTRs of cDNA-1 and cDNA-2 (Figure 11) cloned using Oligo Ax or Bx as downstream primer and Dx as upstream primer. These results also showed the splicing of exon IB onto exon I as in cDNA-A of Li et al. (1990).

In addition, two cDNAs with unique 5' UTRs were identified. The first one is designated cDNA-3 (Fig. 11), which has an extra 5' exon (exon IE) spliced onto the 5' end of exon IB (diagramed in Fig. 12). Comparison between the sequence of this cDNA and the genomic sequence of the rmPFK gene indicates that another intron of 1011 bp is removed by splicing in the 5' UTR of this cDNA, in addition to the spliced intron in cDNA-A of Li et al. (1990). The second clone (cDNA-4, see Fig. 11 for sequence and Fig. 12 for diagram) is a splicing variant of cDNA-3; exon IB is not retained in cDNA-4 (Fig. 12).

Cloning of cDNAs for the 5' UTRs of rmPFK mRNAs by anchored PCR using thermostable reverse transcriptase rTth DNA polymerase

In an effort to determine the transcription start sites of this gene, S1 mapping results (described later in this section, Fig. 17) suggest that the 5' end of cDNA-2 results from premature termination of reverse transcription during the anchored PCR. This premature termination is probably due
to a strong secondary structure in the 5' portion of rmPFK mRNA (Fig. 17). Therefore, anchored PCR experiments were carried out using thermostable *Tth* DNA polymerase in the reverse transcription step at high temperature (65°C). The possible secondary structure in the rmPFK RNA should be destabilized at 65°C. Using primer C in reverse transcription and primers Bx and Dx in PCR (Fig. 10B), this experiment identified three groups of cDNAs that contain exon IB. The first group (four clones) is designated cDNA-5, with the 5' end of the longest clone located at -2857 (Fig. 11 for sequence and 12 for diagram). The second group (one clone) is a new type (designated cDNA-6), with an exon of 63 bp (exon ID) spliced to exon IB (Fig. 11). Alignment of the sequences of exon ID and exon IB with that of the rmPFK genomic DNA indicates that an intron of 705 bp is removed by splicing (Fig. 12). The 5' end of this cDNA is located at -2591. The third group (eight clones) is the same as the afore mentioned cDNA-3 (Figs. 11 and 12). No clone matches cDNA-2, which is probably derived from prematurely terminated reverse transcription of the anchored PCR. This is consistent with the S1 mapping results (see later, Fig. 17). Therefore, cDNA-2 is not considered representing a rmPFK mRNA.

To summarize, these three anchored PCR experiments identified a total of five rmPFK cDNAs (Figs. 11 and 12), excluding cDNA-2. An additional cDNA, designated cDNA-7, was
identified by RT-PCR (see below). The exon sequences of the 5' UTRs of these six rmPFK cDNAs and the genomic sequences are aligned and shown in Figure 13.

**Tissue specific expression of rmPFK mRNAs containing different 5' exons as analysed by Northern blot analysis**

After the cloning of the cDNAs for the rmPFK mRNAs, the tissue-specific expression patterns of these mRNAs were examined by Northern blot analysis using oligonucleotide probes complementary to the 5' exons of the rmPFK mRNAs. The positions of the probes are shown in Figure 12 and their sequences are shown in Table 2. All the probes detected a single mRNA approximately 3 kb in length (Fig. 14). The transcript of the human muscle PFK gene is also about 3 kb (Sharma et al., 1990).

The expression of four rmPFK mRNAs are detected by Northern blot only in skeletal muscle (Fig. 4). Oligo probe F specific to mRNA-a detected the expression of this mRNA only in skeletal muscle (Fig. 14). Oligo probe H complementary to both mRNA-b1 and mRNA-b2 detected these mRNAs also only in skeletal muscle (Fig. 14). Finally, Northern blot using oligo probe I showed that mRNA-c is also skeletal muscle specific (Fig. 14).

Using oligo probe J specific to exon IE, the expression of exon IE was detected in all five tissues examined (Fig. 14). Both mRNA-d1 and mRNA-d2 contain exon IE and they are
Figure 13. Sequences of the 5' flanking region of the rmPFK gene. The A residue of the ATG (bold phase and underlined) translational initiation codon is designated +1. Sequences of exons are in capital letters and underlined. The gap in the sequence between exon IA and exon I was created to indicate the boundary of the two exons, so the sequence is continuous. The TATAAAA sequence found in cDNAs containing exon IB is italic and in bold face.
Figure 14. Northern blot analysis of different rmPFK mRNAs. Poly(A)^+ RNAs (8 μg) from different rabbit tissues were run on a 1% agarose gel with 2.2 M formaldehyde. The nucleic acids were blotted to Nytran membranes overnight by capillary force. The blot was hybridized with probes specific to the exons in the 5' UTRs of the rmPFK gene. The position of the probes are shown in Fig. 12 and their sequences are in Table 2. M: muscle; L: liver; K: kidney; H: heart; B: brain. GAPDH1: hybridized with oligo GAPDH1 (sequence shown in Table 2) complementary to the rabbit glyceraldehyde 3-phosphate dehydrogenase cDNA (Putney et al., 1983). Panels hybridized with probes F, G, and J were exposed for four days, while those hybridized with probes H and I were exposed for seven days. All the blots were hybridized at 55°C for 20-22 hrs and subsequently washed at 55°C two times with 1X SSPE and 0.1% SDS and one time with 1XSSPE. With this stringency, signals detected with probe J are unlikely non-specific (also see Fig. 15D for RT-PCR results).
alternative splicing products. Only mRNA-d1 is expressed in all five tissues examined because of the following: Oligo probe G specific to exon IB, detected rmPFK expression in skeletal muscle, heart, and brain, but not in liver and kidney (Fig. 14). Since mRNA-d2 contains this exon, therefore mRNA-d2 is not expressed in rabbit liver and kidney tissues. These results and those of RT-PCR (below and Fig. 15D) indicate that mRNA-d2 is expressed in rabbit skeletal muscle, heart, and brain, but not in liver and kidney tissues, whereas mRNA-d1 is expressed in all five rabbit tissues examined (Fig. 14 and Fig. 15D). This indicates that the promoter that produce these two mRNAs is active in all five tissues, but the alternative splicing is tissue specific. More specifically, exon IB is retained only in rabbit skeletal muscle, heart and brain tissues.

Analysis of the tissue-specific expression of multiple mRNAs of the rmPFK gene by Reverse transcription (RT)-PCR

The tissue-specific expression patterns of different rmPFK mRNAs were further examined by RT-PCR. The primers used in this study are shown in Figure 12 (for sequence, see Table 2). The common downstream primer (oligo 0) is complimentary to the fifth coding exon of the rmPFK gene. The amplified RT-PCR products include the 5' UTRs and the five coding exons of the rmPFK mRNAs. The presence of a PCR product of appropriate size in an agarose gel based upon the
sequence of the cDNA clones was considered as evidence of expression.

Results of the RT-PCR analysis are shown in Figure 15. Oligo K, the upstream primer specific to mRNA-a detected expression in skeletal muscle and heart tissue (Fig. 15A). Northern blot analysis detected the expression of this mRNA only in skeletal muscle (Fig. 14, probe F). This discrepancy might be due to the high sensitivity of RT-PCR compared to Northern blot analysis. A similar inconsistency has been found in the studies of human muscle PFK (hmPFK) gene: although this gene was shown by RT-PCR to be expressed in all seven human tissues examined, including skeletal muscle, liver, and kidney, its expression was not detected by Northern blot hybridization in liver, placenta, pancreas, stomach, and reticulocytes (Nakajima et al., 1990).

Oligo L, the upstream primer specific to exon IC showed that there are two mRNAs expressed in skeletal muscle (Fig. 15B). The larger one corresponds to mRNA-b1. In addition to the b1 product, a smaller PCR product of the skeletal muscle RNA was amplified by RT-PCR with primer L (Fig. 15B). This smaller DNA product (b2) was gel purified and cloned into pBluescript II KS(+) vector which was digested with EcoRV restriction endonuclease (Marchuk et al., 1991). The sequence of the insert (designated as cDNA-7) is shown in Figure 11. Comparison of the sequence of this clone and the rmPFK genomic sequence indicates that part of the 5' UTR of
Figure 15. Analysis of the tissue-specific expression of the six mRNAs of the rmPFK gene by RT-PCR using thermostable reverse transcriptase Tth DNA polymerase. One μg of poly(A)* RNA from each rabbit tissue was reverse transcribed with Tth DNA polymerase using the 3' primer (oligo O, see Fig. 12 for diagram and Table 2 for sequence). The reverse transcription reaction mixture was aliquoted into four tubes. Each of the four 5' primers was then added along with PCR mixture to each tube: Primer K in (A); L in (B); M in (C); and N in (D). The positions of these primers are shown in Fig. 12B and the sequences are in Table 2. PCR amplification was performed at 60°C for 1 min and 95°C for 1 min for each cycle. A total of 35 cycles were carried out. The precipitated PCR product were analyzed on 3% agarose gels (3 parts of Nusieve agarose from FMC Bioproducts and 1 part of agarose I from Amresco). For RT-PCR in panel C, the PCR product was diluted 100 times and reamplified for another 35 cycles. The expected sizes of the RT-PCR products are as following: a: 397 bp; b1: 646; c: 526 bp; D1: 503 bp; d2: 397 bp. The b2 band was cloned and sequenced and the expected size is 504 bp. (■) RT-PCR amplification of rabbit GAPDH mRNA using primers (GAPDH2 and GAPDH3, Table 2) complementary to the rabbit GAPDH cDNA (Putney et al., 1993). The expected sizes of the PCR product is 140 bp. M: skeletal muscle; L: liver; K: kidney; H: heart; B: brain; D: 100 bp DNA size ladder from GibcoBRL.
mRNA-b1 is removed by splicing as a 142 bp intron (Fig. 12). Therefore, alternative splicing gives rise to two rmPFK mRNAs, mRNA-b1 and -b2, that contain exon IC (Fig. 12).

Using oligo M, an upstream primer specific to mRNA-c, expression of this rmPFK mRNA was detected only in skeletal muscle (Fig. 15C). The RT-PCR product in panel C of Figure 15 was reamplified, indicating that mRNA-c is expressed at a relatively lower level than the other rmPFK mRNAs. This is consistent with the weak signal detected in the Northern blot analysis (Fig. 14, oligo probe I).

RT-PCR using oligo N, an upstream primer specific to exon IE, showed two predominant products corresponding to mRNA-d1 and -d2 (Fig. 15D). As mentioned above, these two mRNAs are derived from alternative splicing; exon IB is retained in mRNA-d1 but not in mRNA-d2 (Fig. 12). The expression of mRNA-d2 was detected in all five tissues examined. However, mRNA-d1 was not detected in liver and kidney, indicating that exon IB is retained in skeletal muscle, heart, and brain mRNA, but not in mRNA from kidney and liver (Fig. 15D). These results are consistent with the Northern blot analyses: rmPFK mRNAs containing exon IE (mRNA-d1 and -d2, Fig. 12) were detected in all five tissues when probe J was used (Fig. 14). However mRNAs containing exon IB including mRNA-d2 (Fig. 12) was not detected in liver and kidney tissues (Fig. 14, probe G).
The tissue-specific expression patterns of the six rmPFK mRNAs are summarized and shown in Fig. 12.

**Identification of the transcription start site of the downstream promoter**

The transcription start site of the proximal promoter (Pa) that gives rise to mRNA-a was determined by both primer extension and SI mapping (Fig. 16). The primer extension study using M-MLV reverse transcriptase mapped a major site exactly at the 5' end of cDNA-1 and several minor sites (Fig. 16A). The SI mapping experiment mapped a major start site in the vicinity of the major site mapped by primer extension, along with some minor sites (Fig. 16B). Together, these results indicate that the proximal promoter of the rmPFK gene initiates transcription at a major site 86 bp upstream of the ATG translation initiation codon, specifically at the A residue in the 5'-GCCAGTCT-3' sequence, along with several minor sites. No TATA box is found -20 to -30 bp upstream of these start sites. The sequence surrounding the major start site (GCCAGTCT) is homologous to the functional initiator element (CTCANTCT) found in the TATA-less promoter of the murine terminal nucleotidyl transferase gene (Smale et al., 1990). Multiple start sites are also characteristic of many TATA-less promoters (Dynan, 1986). For example, the human Ha-ras gene has a TATA-less promoter and initiates transcription
Figure 16. Determination of transcription start sites of the proximal promoter of the rmPFK gene. (A) Primer extension. The primer extension reaction was carried out with 8 μg of rabbit muscle poly(A)+ RNA. Primer F (panel C for position and Table 2 for sequence) used in the primer extension was 5'-end labeled with polynucleotide kinase and [γ-32P]ATP. (B) S1 mapping. Probe A was generated by annealing the labeled primer F to the single stranded pRMPFK5'. The primer was extended with Sequenase. The extended products were subsequently digested with PstI. The single stranded probes were subsequently purified with alkaline agarose gel electrophoresis (Ausubel et al., 1993). 8 μg of rabbit muscle poly(A)+ RNA and 2x10^5 CPM of probe A were hybridized in the aqueous hybridization buffer overnight at 50°C. S1 digestion (150 units/300 μl reaction mix) was carried out at 30°C for 1 hr. Both the primer extension product and the S1 mapping mixture were analyzed on a 6% acrylamide sequencing gel. Size markers are sequencing ladders generated with primer A and single stranded DNA of pRMPFK5'. The arrow indicates the position of the major band in primer extension and S1 mapping analysis. M: muscle. (C) Schematic diagram depicting the primer used in primer extension and probe used in S1 mapping. The plasmid prmPFK5' was constructed by ligation of a 2.5 kb SstI-SstI fragment of the rmPFK gene into the pBluescript II plasmid. The SstI-SstI PFK gene fragment contains 5' flanking region and the first coding exon. The single stranded S1 probe was prepared as described by Ausubel et al. (1993). Only mRNA-a will hybridize to probe A. The predicted probe fragment is diagramed under mRNA-a.
A

B

C

5' TATA ATG 3' genomic DNA

pRMPFK5'

F Primer extension

probe A for S1 mapping

mRNA-a

a
at multiple start sites spread over a 90-base pair region (Lu et al., 1994).

**Determination of transcription start site of the promoter close to the TATAAAA sequence**

Since the 5'-ends of cDNA-2 and -5 are 28 bp downstream and 233 bp upstream, respectively of the TATAAAA sequence (Fig. 11), it is possible that promoter(s) lie near this sequence. To determine the transcription start sites of such promoter(s), first S1 mapping analysis was performed. A single stranded DNA probe of 680 bp was prepared with its 5' end located 80 bp downstream of the TATAAAA sequence (S1 probe in Fig. 17E). The S1 mapping results (Fig. 17A) showed two groups of protected fragments. The largest fragment in the first group (b1) at the top of the gel is about 230 bp upstream of the TATA sequence. This position is within 5 bp of the 5' end of cDNA-5 identified by anchored PCR using Tth DNA polymerase. In the genomic DNA, a TATTTAT sequence is 20 bp upstream of the 5' end of cDNA-5 (5' end of exon IC in Fig. 13). This sequence is AT rich, but is not a perfect match to the eukaryotic TATA box consensus sequence [TATA(A/T)A(A/T)] found 20 to 30 bp upstream of the transcription start site of many RNA polymerase II promoters (Breathnach and Chambon, 1981).

The other group of fragments (x in Figs. 17A) is about 9 to 14 bp upstream of the TATAAAA sequence. This position corresponds to a splicing acceptor site (5' end of exon IB),
Figure 17. Determination of transcription start sites of the promoter close to the TATAAAA sequence of the rmPFK gene. (A) S1 mapping. Probe B (depicted in panel E) was generated similarly as probe A (see legend to Fig. 11), except labeled primer G, instead of primer F, was used. Primer G was extended with Sequenase using the single stranded pRMPFK5' as template. The extended double stranded DNA products were subsequently digested with SstI, rather than PstI restriction enzyme. The single stranded probe was then isolated from alkaline agarose gel (Ausubel et al., 1993). 8 μg of rabbit muscle poly(A)+ RNA and 2x10^5 CPM of probe B were also hybridized in the aqueous hybridization buffer overnight at 50°C. (B) Primer extension using M-MLV reverse transcriptase. The primer extension reaction was carried out with 8 μg of rabbit muscle poly(A)+ RNA and 2x10^5 CPM of 5' end labeled primer G (panel E for position and Table 2 for sequence). S1 digestion (150 units S1 nuclease/300 μl reaction mix) was carried out at 30°C for 1 hr. The S1 mapping mixture were analyzed on a 6% acrylamide sequencing gel. (C) Primer extension using thermostable reverse transcriptase Tth DNA polymerase. Primer and RNA used were the same as in (B) (for detail, see Materials and Methods) Size markers are sequencing ladder using primer B and single stranded DNA of pRMPFK5'. M: muscle. (D) Possible secondary structure in the 5' UTR of cDNA-2 as predicted by the Fold algorithm in the GCG program (Zuker and Stiegler, 1981; Frier et al., 1986). (E) Primer G (complementary to exon IB) and the S1 probe used. The construction of prmPFK5' is described in the legend of Fig. 16. The protected S1 probe fragments (b1 and x) are diagramed under individual rmPFK mRNA. mRNA-a and mRNA-d2 are not complementary to this S1 probe because they do not contain exon IB.
which is 11 bp upstream of the TATAAAA sequence. This splicing junction is common to mRNA-b2, -c, and -d1 (Fig. 17E).

The primer extension analysis also revealed multiple products (Fig. 17B). The b1 band lie about 230 bp upstream of the TATAAAA sequence, consistent with the S1 mapping results. Therefore, the rmPK gene likely contains a promoter that initiates transcription at several sites about 230 bp upstream of the TATAAAA sequence. This promoter is designated promoter b.

Primer extension generated several products between b1 and y bands in Fig. 17B. These primer extension products may represent mRNAs such as mRNA-d1, that has extra 5’ exons (exon IE in this case) spliced onto exon IB (Fig. 12). S1 mapping of these mRNAs resulted in protected fragments corresponding to the 5’ end of exon IB, the splicing junction (x in Fig. 17A). However, I can not rule out the possibility that these primer extension products are artifacts due to the secondary structure in mRNA-b1. Nevertheless, the size of these products are unlikely indicative of transcription start sites when compared to the size marker using genomic DNA sequencing ladder, because S1 analysis did not show bands at the corresponding size.

Finally, primer extension analysis showed there is a strong band (band y in Fig. 17B) about 28 bp downstream of the TATAAAA sequence. The 5’ end of cDNA-2 identified by
anchored PCR using M-MLV reverse transcriptase is also at this site. However, the S1 mapping results (Fig. 17A) showed that there is no band downstream of the TATAAAA sequence, particularly 28 bp downstream of the TATA-1 sequence.

To address this discrepancy, primer extension was carried out at 65°C using a thermostable reverse transcriptase, the rTth DNA polymerase (Fig. 17C). The results of this primer extension are similar to that using M-MLV reverse transcriptase, except that there is no strong band about 28 bp downstream of the TATAAAA sequence (Fig. 17C). This is consistent with the S1 analysis (Fig. 17A). Therefore, it is very unlikely that there is a transcription start site 28 bp downstream of the TATAAAA sequence. It is possible that the y band in the primer extension using M-MLV reverse transcriptase (Fig. 17B), and the cDNA-2 identified by the anchored PCR also using M-MLV reverse transcriptase resulted from the prematurely terminated reverse transcription reaction. This may be due to possible secondary structure in the 5' UTR of the mRNA. The fold algorithm in the GCG program (Zuker and Stiegler, 1981; Freier et al., 1986) predicted stem-loop structures in this region (Fig. 17D). In addition, anchored PCR using thermostable Tth DNA polymerase did not identify cDNA-2.

**DISCUSSION**

Previously, two cDNAs for the rabbit muscle PFK gene has been cloned (Li et al., 1990). These two cDNAs have an
identical coding region, but their 5' UTRs are different. In one of the cDNAs, an intron of 1.7 kb is removed by splicing (Li et al., 1990). These two cDNAs could be derived from two promoters, or from one promoter and alternative splicing. In addition, the tissue specificity of the rmPFK mRNAs have not been determined.

First in this report, six cDNAs for the 5' UTR and the 5' portion of the coding region of the rmPFK mRNAs were cloned (Fig. 12). These cDNAs have the same translation initiation codon as the full length cDNAs cloned by Li et al. (1990). Therefore, the cDNAs cloned here likely have the same open reading frame as the full length cDNAs of Li et al. (1990). The six cDNAs cloned are consistent with two rmPFK cDNAs cloned previously (Fig. 10; Li et al., 1990): cDNA-1 in this study contains the entire 5' UTR of cDNA-A, while cDNA-5, -7, -6, and -3 (from 3' to 5') all contain the entire 5' UTR of cDNA-2 (Fig. 12 and Fig. 10) and each has more 5' sequence than the 5' UTR of cDNA-A of Li et al. (1990).

In addition, cDNA-4 has a 5' UTR that is not present in either of the two cDNAs of Li et al. (1990). This cDNA is an splicing variant of cDNA-3: exon IB (exactly the entire 5' UTR of cDNA-A of Li et al., 1990) is not retained in cDNA-3. Two more cDNAs are derived from alternative splicing: cDNA-5 is a splicing variant of cDNA-7; more specifically, the most 5' intron in the 5' UTR of cDNA-5 is retained, but is removed by splicing in cDNA-7 (Fig. 12). Southern blot
analysis of rabbit genomic DNA digested with restriction enzymes indicates that rmPFK is a single copy gene (Knaak and Chang, unpublished data). The human muscle PFK gene has also been shown to be single copy (Sharma et al., 1990).

The transcription start sites of two downstream promoters of the rabbit gene have been mapped. Promoter a initiates transcription at a major site 86 bp upstream of the ATG translational initiation codon, along with several minor sites. The 5' end of cDNA-1 identified by anchored PCR is at the major start site mapped by primer extension and S1 mapping. The transcription start sites of a second promoter reside at -2025 upstream of the ATG codon. The 5' end of mRNA-b1 is at this site. Although the transcription start sites of the promoter(s) giving rise to mRNA-c (containing exon ID), mRNA-d1 and -d2 (both containing exon IE) have yet to be determined. Because exon ID and exon IE are upstream of promoters a and b, they can not be transcribed from these promoters and must be transcribed from promoters further upstream. Therefore, the six rmPFK mRNAs are expressed by at least three promoters and by alternative splicing from a single copy gene.

The 5' end of cDNA-2 is 28 bp downstream of the TATAAAA sequence. Primer extension using M-MLV reverse transcriptase also showed a band at this position (Fig. 17B). However, S1 mapping analysis clearly showed no band at this position (Fig. 17A). Furthermore, primer extension analysis at 65°C
using Tth DNA polymerase did not show a band at this position (Fig. 17C). Therefore, it is very unlikely that there is a transcription start site at this position. The reverse transcription reaction using M-MLV reverse transcriptase is probably prematurely terminated at this position, resulting in a band in the primer extension analysis and the 5' end of cDNA-2 at this position.

It is intriguing that there is no transcription start site 20 to 30 bp downstream of this TATAAAAA sequence, since it matches the TATA box consensus sequence found about 25 bp downstream of transcription start site of many eukaryotic protein encoding genes (Breathnach and Chambon, 1981). In addition, there are multiple putative binding site for transcription activators upstream of this sequence. For example, within 200 bp upstream of the TATA sequence there are three binding sites for the bHLH myogenic transcription activators, such as MyoD and myogenin (Weintraub et al., 1991).

There are several possibilities that there is no transcription start site downstream of this TATA sequence. First of all, there may be a specific repression sequence near the TATA sequence. Second, chromatin structure in this region might prevent transcription initiation. It has been shown in vitro that a minimal promoter containing the TATA-box is able to direct assembly of the transcription initiation complex onto the DNA template and initiate
transcription (Buratowski et al., 1989). The binding of transcription factor TFIID to the TATA box is essential for transcription initiation. However, binding of TFIID to a DNA template containing a TATA box and subsequently the transcription initiation is severely inhibited after the nucleosome assembly (Workman et al., 1991; Meinsterernst et al., 1990).

Recently, it has been shown human SWI/SNF complex can facilitate the binding of TATA-binding protein (TBP) to nucleosomal DNA, but only when the TATA box is in the favorable position (Imbalzano et al. (1994). These researchers constructed a DNA template of 149 bp containing two nucleosome positioning sequences (Shrade and Crothers, 1989) at the 5' end and the adenovirus major late promoter TATA box in the middle (PH MLT). The phasing sequence favors specific positioning of the histone core. They constructed two variant templates [PH MLT9(+3) and PH MLT(+6)] positioned the TATA box either 3 or 6 bp, respectively, closer to the phasing sequences. After nucleosome assembly, TBP can not bind to the TATA box in all three templates. Human SWI/SNF complex has been shown to facilitate the binding of TBP to the TATA box in PH MLT (+3) template. However, the complex does not facilitate the binding of TBP to the TATA box in the PH MLT and PH MLT(+6) templates (Imbalzano et al., 1994). These data imply that if a TATA box is not at a favorable rotation relative to the
nucleosome, even the SWI/SFN proteins can not facilitate the binding of TBP to the TATA sequence in a DNA template. The TATA sequence of the rmPFK gene could be in such a position. There are three sequences near the TATA sequence that match the consensus nucleosome positioning sequence. For example, sequence 5'-CCCCCAAT (28 downstream of the TATA sequence) matches the strong nucleosome positioning consensus sequence (G/C)\textsubscript{3}NN(A/T)\textsubscript{3} (Shrade and Crothers, 1989).

After cloning six rmPFK mRNAs, I then studied the tissue-specific expression of these mRNAs by Northern blot analysis and RT-PCR technique. The results showed that these mRNAs are expressed with a complex pattern by both tissue-specific transcription and alternative splicing: mRNA-b1, -b2, and -c are skeletal muscle specific (Figs. 14 and 15). The expression of mRNA-a was detected only in skeletal muscle by Northern blot analysis (Fig. 14). However, RT-PCR experiment showed that this mRNA is also expressed in heart tissue (Fig. 15A). This discrepancy might be due to the high sensitivity of RT-PCR compared to Northern blot analysis. A similar inconsistency has been found in the studies of human muscle PFK (hmPFK) gene: although this gene was shown by RT-PCR to be expressed in all seven human tissues examined, including skeletal muscle, liver, and kidney, its expression was not detected by Northern blot hybridization in liver, placenta, pancreas, stomach, and reticulocytes (Nakajima et al., 1990). The rmPFK mRNA-d1 was detected in skeletal
muscle, heart, and brain tissues, but not in liver or kidney, while mRNA-d2 was detected in all five rabbit tissues examined (Fig. 15). The pattern of the tissue-specific expression of the rmPFK mRNAs described here is consistent with that of the rmPFK protein (Table 1; Tsai and Kemp, 1973). The rmPFK has been shown to be the predominant PFK in rabbit skeletal muscle. This isozyme has also been shown to be expressed at the protein level in brain, heart, kidney, and at very low level in the liver (Tsai and Kemp, 1973). Mhaskar and Dunaway (1995) studied the subunit abundance, protein synthesis, and steady state mRNA levels of the three PFK isozymes during rat neonatal brain development. Interestingly, from zero to 23 days after birth, the level of the M-type subunit, the protein synthesis, and the steady state level of rat muscle PFK mRNA increased simultaneously, indicating the steady state level of the M-type PFK mRNA plays an important role in controlling the abundance of the muscle type PFK isozyme. This is consistent with the correlation between the tissue-specific expression pattern of the rabbit muscle PFK at the protein level (Tsai and Kemp, 1973) and at the mRNA level (this study).

There are similarities and differences in the tissue-specific expression of multiple mRNAs from the
rmPFK gene and its human counterpart. The human gene expresses three mRNAs by two promoters and alternative splicing (Fig. 7; Yamasaki et al., 1991). The downstream promoter produces two hmPFK mRNAs, -a and -b. The hmPFK mRNA-a is a splicing variant of hmPFK mRNA-b: there is an 89 bp intron spliced out in mRNA-a.

The rmPFK mRNA-a and its human counterpart (hmPFK mRNA-b) are highly conserved (80% homology in the 5' UTR). The upstream sequences of the two genes have diverged significantly: a 750 bp region of human sequence and its rabbit counterpart have only 41% homology. The rmPFK mRNA-a and its human counterparts are both skeletal muscle specific as examined by Northern blot, although the rmPFK was detected in the heart by RT-PCR. The expression of the human counterparts were not examined in this tissue. The mRNA-c transcribed by the upstream promoter of the human gene is probably the counterpart of rmPFK mRNA-d2: both mRNAs were found expressed in all the tissues examined.

In the hmPFK mRNA-c, the intron in the 5'UTR is about 3.3 kb, while in rmPFK mRNA-d2, the intron is about 2.8 kb. In the rabbit, there is a splicing variant of this mRNA, mRNA-d1, in which exon IB is retained in a tissue specific manner, more specifically retained in skeletal muscle, heart, and brain, but not in liver and kidney tissues (Fig. 12). There are no cloned human counterparts
for the rabbit mRNA-b1, -b2, and -c. In mouse, only one promoter has been characterized for the muscle PFK gene (Gekakis and Sul, 1994), which is equivalent to the downstream promoter of the human and rabbit muscle PFK genes. Such nonconservative usage of alternative promoters has also been reported in the aldolase A gene. In the human gene there are three promoters (Gautron et al., 1991), while there are only two promoters in the mouse gene, with no equivalent known human upstream promoter (Stauffer et al., 1990).

Multiple mRNAs with different 5' UTRs expressed from multiple promoters and/or alternative splicing as observed in rabbit and human muscle PFK genes have also been reported in other mammalian genes. For example, the gene from rat encoding a bifunctional enzyme 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase has two tissue specific promoters, one muscle specific and the other liver specific (Darville et al., 1989). This enzyme catalyzes both the synthesis and degradation of fructose-2,6-bisphosphate, a potent regulator of PFK (El-Maghrabi et al., 1990). The human aldolase A gene contains three promoters, including a muscle specific one (Gautron et al., 1991). Interestingly, aldolase A is also involved in glycolysis.

The mouse γ-glutamyltransferase gene expresses its six mRNAs from five promoters, three of which are kidney
specific (Rajagopalan et al., 1993; Sepulveda et al., 1994). The human carbonic anhydrase I gene contains two promoters with different tissue specificities (Brady et al., 1991). Multiple promoter usage has also been observed in mouse glucorticoid receptor gene (Strahle et al., 1992), rat dopamine-D2-receptor gene (Valdenaire et al., 1994), and rat liver alkaline phosphatase gene (Toh et al., 1989).

Multiple promoters of these genes may give an organism more flexibility to fine tune the regulation of the expression from these genes according to the physiological function of a given tissue. Multiple promoters may also serve as a safe guard to prevent shut-down of the gene expression when key promoter elements are mutated. In the case of rabbit PFKs, all three isozymes are expressed in brain and kidney, while in skeletal muscle, however, only the M type PFK is expressed. If the rmPFK gene has only one promoter, the skeletal muscle will suffer from PFK deficiency when the promoter is mutated in such a way that the expression of the gene is dramatically reduced. This might result in PFK deficiency disease like the Tarui disease. However, this can be prevented if the gene has multiple promoters. Although only one promoter of the rmPFK gene is active in rabbit kidney tissues, the down-mutation of the promoter will only result in
reduced expression from the muscle gene. The Kidney will still have PFK activity because L type and C type PFKs are also expressed in Kidney.
SECTION II
PRIMER EXTENSION AT HIGH TEMPERATURE BY THERMUS THERMOPHILUS DNA POLYMERASE
INTRODUCTION

Primer extension is a method frequently used in conjunction with S1 mapping or RNase protection analysis in determining gene transcription start sites. In this method, an RNA template is reverse-transcribed by either avian myeloblastosis virus (AMV) reverse transcriptase or Moloney-murine leukemia virus (M-MLV) reverse transcriptase. A drawback of this method is that secondary structure in the RNA template may cause the reaction to be prematurely terminated. To address this problem, Bailey and Davidson (1976) have reported the use of methylmercuric hydroxide to destabilize the RNA secondary structure. This compound is extremely toxic and hence inconvenient to handle. An alternative method is to raise the temperature of the extension reaction (Huibregste and Engelke, 1986; Shimimaye and Salvato, 1989). Elevated temperature would destabilize the secondary structure in the RNA template and increase primer extension specificity. Yet the upper temperature limit is about 50°C for both the M-MLV and AMV reverse transcriptases, due to the thermal inactivation of these enzymes. In my studies of the transcription start sites of the rabbit muscle phosphofructokinase (PFK) gene, I found that primer extension was prematurely terminated, even at 50°C (see SECTION I). As predicted by the Fold algorithm in the GCG program (Zuker and Stiegler, 1981; Freier et al., 1986), there is a putative secondary structure in the rabbit muscle PFK mRNA consisting of two stem-loops, including four...
consecutive G-C basepairs immediately upstream of where the primer extension stopped (Fig. 17 D). This prompted me to search for a thermostable reverse transcriptase which could perform primer extension at temperatures higher than 50°C.

Recently, thermostable DNA polymerase from *Thermus aquaticus* (Taq) has been shown to possess reverse transcriptase activity in reverse transcription-polymerase chain reaction (RT-PCR) (Jones, 1993; Jones and Foulkes, 1989; Tse and Forget, 1990; Shaffer et al., 1990). The 94-kDa DNA polymerase from *Thermus thermophilus* (Tth) has been reported to possess even higher reverse transcriptase activity in RT-PCR experiment than Taq DNA polymerase (Myers and Gelfand, 1991). However, primer extension products extended from cellular RNA by either Taq or Tth DNA polymerase have not been fully characterized other than by RT-PCR. Under the conditions described by Myers and Gelfand (1991), the Tth DNA polymerase was found not efficient enough for primer extension analysis of cellular RNA. Therefore, several conditions for primer extension with this DNA polymerase were studied using rabbit heart poly(A)^+ RNA and a primer complementary to the rabbit glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene. Under the optimized conditions, primer extension with Tth DNA polymerase at 65°C was as efficient as that with M-MLV reverse transcriptase at 50°C. More importantly, primer extension of the rabbit muscle PFK mRNA using Tth DNA polymerase at 65°C solved the problems caused by the RNA secondary structure (see SECTION
I). This study established a protocol of primer extension to map transcription start sites at elevated temperature using the thermostable Tth DNA polymerase. It would also be useful for the first strand cDNA synthesis in procedures such as cDNA library construction, anchored-PCR (Loh et al., 1989), and rapid amplification of cDNA ends (RACE) (Frohman et al., 1988).

MATERIALS AND METHODS

Enzymes

Tth DNA polymerase (2.5 units/μl) was purchased from Perkin Elmer Cetus Co. Superscript II reverse transcriptase (RNase H deficient, genetically engineered M-MLV reverse transcriptase) was purchased from Life Science Technologies, Inc. T4 polynucleotide kinase was purchased from New England Biolabs, Inc. Shrimp alkaline phosphatase was purchased from United States Biochemical Co.

Preparation of the oligonucleotide primer and size marker

Primer RabGAPDH (5' OH-CATGTAGACCATGTAGTGGAGGTCAATG-OH 3') was complementary to the 5' sequence of rabbit GAPDH cDNA (Putney et al., 1983). This oligonucleotide was prepared by DNA International, Inc. and purified by polyacrylamide gel electrophoresis as described (Sambrook et al., 1989). The purified oligo was 5' end-labeled with [γ-32P]ATP and T4 polynucleotide kinase (Sambrook et al., 1989). Size markers were generated by digesting pBR322 plasmid DNA
with HpaII restriction endonuclease, followed by
dephosphorylation with alkaline phosphatase and
5' end-labeling (Sambrook et al., 1989).

**Poly(A)^+ RNA preparation**

Total RNA from skeletal and cardiac muscle of 10 day old
New Zealand white rabbits was extracted by homogenizing the
tissue in guanidinium isothiocyanate, followed by cesium
chloride gradient centrifugation (Chirgwin et al., 1979).
Poly(A)^+ RNA was purified from the total RNA by oligo(dT)
cellulose column chromatography (Sambrook et al., 1989). The
concentration of poly(A)^+ RNA was determined by measuring
the absorbance at 260 nm (A_{260}).

**Primer extension by Tth DNA polymerase**

Primer extension by Tth DNA polymerase was carried out as
described below or as otherwise stated in the individual
figure legend. The hybridization mix (12 µl) containing 16.7
mM Tris-HCl (pH 8.3), 150 mM KCl, 5x10^4 CPM of primer
RabGAPDH and 3 µg of rabbit poly(A)^+ RNA, was heated at 75°C
for 5 min to denature the RNA. After annealing at 65°C for 5
min, the extension mix (8 µl) containing 2.5 mM MnCl_2, 500
µM each of dATP, dGTP, dCTP, and dTTP, and 5 units of Tth
DNA polymerase was added, and the reaction was overlaid with
20 µl of mineral oil. Primer extension was carried out at
65°C for 1 hr in a Perkin Elmer Cetus thermocycler. After
extension, the mineral oil was carefully removed, and the
residual oil was further extracted with phenol/chloroform
before precipitation (see below).
Along with 20 μl of H₂O, 105 μl of RNase reaction mix (100 μg/ml salmon sperm DNA, 20 μg/ml DNase-free RNase A) (Ausubel et al., 1993) was added to each primer extension reaction tube. RNase digestion was carried out at 37°C for 15 min, followed by the addition of 15 μl of 3 M sodium acetate, pH 5.2. This mixture was extracted once with phenol/chloroform/isoamyl alcohol and precipitated with ethanol. Half of the primer extension products were analyzed as described (Ausubel et al., 1993).

**Primer extension by M-MLV reverse transcriptase**

Primer extension was performed basically as described (Ausubel et al., 1993), except that the extension reaction was carried out with 100 units of Superscript II reverse transcriptase at 50°C instead of 42°C. Three μg of poly(A)⁺ RNA from rabbit heart and 5x10⁴ CPM of primer were used. Half of the reaction products were resolved on 6% polyacrylamide gels containing 7 M urea.

**RESULTS AND DISCUSSION**

In order to conduct primer extension at elevated temperature, it is necessary to find a thermostable reverse transcriptase. Tth DNA polymerase is a very attractive enzyme because it is thermostable and possesses higher reverse transcriptase activity than Taq DNA polymerase (Myers and Gelfand, 1991). These researchers have reported successful results using thermostable Tth DNA polymerase in RT-PCR to detect cellular RNA expression when 400 pg of total RNA has been used. They have optimized several
reaction parameters for the reverse transcription using Tth DNA polymerase. For example, they found that Mn$^{++}$ is a better cation than Mg$^{++}$. However, using these conditions, only a very faint band was visible when the primer extension reaction was analyzed on a polyacrylamide gel (Fig. 19, lane 1). These results suggest that under these conditions, the reverse transcription efficiency by Tth DNA polymerase is low and is useful only when the reverse transcription is coupled to the amplification by PCR.

In order to explore the possibility of using Tth DNA polymerase in mapping transcription start sites, I optimized the experimental conditions which improved the reverse transcription efficiency of this enzyme. First, I studied the performance of Tth DNA polymerase as a reverse transcriptase at different temperatures using GAPDH specific primer and rabbit heart poly(A)$^+$ RNA. Although Tth DNA polymerase has been shown to possess reverse transcriptase at 70°C (Myers and Gelfand, 1991), this activity was reduced at increased temperature (Fig. 18). At 70°C and 75°C, only a very faint band is visible (Fig. 18, lanes 1 & 2). On the other hand, shorter bands were also present in addition to the band of 203 bp when the reaction was performed at 55°C and 60°C, albeit the reverse transcriptase activity of Tth DNA polymerase was higher at 60°C than at 65°C (Fig. 17, lane 4 vs lane 3). When the reaction was carried out at 65°C, a predominant band of 203 bp was produced. Therefore,
Figure 18. Performance of Tth DNA polymerase in primer extension at different temperatures. Primer extension was carried out at 75, 70, 65, 60, and 55°C for 60 min (lanes 1, 2, 3, 4, and 5, respectively). In each reaction, $5 \times 10^4$ CPM of 5'-end labeled oligonucleotide primer complementary to the rabbit GAPDH cDNA and 3 μg of poly(A$^+$) RNA from rabbit heart is used, and half of the primer extension product is loaded. M: DNA size marker generated by digesting pBR322 DNA with HpaII restriction endonuclease, followed by dephosphorylation and 5' end labeling with [γ-$^{32}$P]ATP and T4 polynucleotide kinase. Arrow indicates the primer extension product about 203 bp.
subsequent primer extension reactions with Tth DNA polymerase were carried out at 65°C.

Next, I examined the effect of the extension time on the amount of the product extended. The results showed that longer extension times enhanced the intensity of the band, the one extended for one hr being the strongest (Fig 19, lanes 2-4).

Actinomycin D is routinely used in primer extension to prevent primers to form hairpin when AMV or M-MLV reverse transcriptase is employed. However our results showed that this antibiotic reduced the reverse transcription efficiency of Tth DNA polymerase (lane 1 vs lane 2, Fig. 20). Hence, this antibiotic is not recommended in the primer extension reaction when Tth DNA polymerase is employed.

Finally, I compared the products extended by M-MLV reverse transcriptase with those by Tth DNA polymerase. The results showed that there was no visible difference in intensity and mobility of the products extended by these two polymerases (Fig. 21, lane 1 vs lane 2). Primer extensions were carried out for 60 min with M-MLV reverse transcriptase at 50°C and for 60 min with Tth DNA polymerase at 65°C. Extension for one hour has been recommended in primer extension protocols for AMV or M-MLV reverse transcriptase (Sambrook et al., 1989; Ausubel et al., 1993). The results under conditions described here demonstrate that primer extension with Tth DNA polymerase was as efficient as that with M-MLV reverse transcriptase. Primer extension by
Figure 19. Effect of extension time on the primer extension by Tth DNA polymerase. Extension reaction was performed at 65°C for 10, 30, and 60 min. (lanes 2, 3, and 4, respectively). Lane 1: primer extension under the conditions (at 70°C for 10 min) used by Myer and Gelfand (1991). M: DNA size marker and see legend to Figure 18 for detail.
Figure 20. Effect of actinomycin on primer extension by Tth DNA polymerase. Extension reaction was carried out at 65°C for 60 min with (lane 1) and without (Lane 2) actinomycin (0.15 mg/ml). M: DNA size marker, same as in Figure 18.
Figure 21. Comparison of primer extension of GAPDH RNA by Tth DNA polymerase and that by MMLV reverse transcriptase. Lane 1: primer extension of rabbit heart poly(A⁺) RNA by MMLV reverse transcriptase; lane 2: primer extension of rabbit heart poly(A⁺) RNA by Tth DNA polymerase; lane 3: primer extension of rabbit skeletal muscle poly(A⁺) RNA by Tth DNA polymerase. In each reaction, 5x10⁴ CPM of primer and 3 µg of poly(A⁺) RNA from rabbit heart, except lane 3 of panel B from skeletal muscle, is used. Half of the primer extension product is loaded in each lane. M: size marker, see Fig. 18 for detail. Arrow indicates the primer extension product about 203 bp.
Tth DNA polymerase using poly(A)+ RNA isolated from rabbit skeletal muscle as template also produced a band similar in mobility to those produced using poly(A)+ RNA from rabbit heart (Fig. 21).

In conclusion, I have optimized conditions to perform primer extension of cellular RNA at elevated temperatures using thermostable Tth DNA polymerase. Before optimization, only a faint band was visible (Fig. 19, lane 1). After optimization, a strong band was produced (Fig. 19 lane 4). Furthermore, primer extension using Tth DNA polymerase at 65°C and M-MLV reverse transcriptase at 50°C produced bands similar in size and intensity (Fig. 21, lane 2 vs lane 1). This protocol might be useful not only in mapping transcription start sites, but also in the first strand cDNA synthesis for cDNA library construction, for anchored-PCR (Loh et al, 1989), or for rapid amplification of cDNA ends (RACE) (Frohman et al., 1988), especially when secondary structure of the template RNA may be a potential problem in the procedure.
CONCLUSIONS AND FUTURE STUDIES
PFK catalyzes the ATP dependent phosphorylation of F-6-P, the rate-limiting step in glycolysis. This enzyme is very interesting to biochemists due to its importance in the glycolytic pathway and also its allosteric kinetic behavior responding to various metabolites (Dunaway, 1983). In addition, this enzyme is physiologically important. The loss of function of the muscle PFK gene results in Tarui disease (Tarui et al., 1965).

In mammalian tissues, PFK exists as homo- or heterotetramers of three distinct subunits. The three isozymes exhibit different kinetic properties (Foe and Kemp, 1985). The relative levels of these isozymes in a given tissue determine the composition of PFK tetramers and subsequently their allosteric behaviors in that tissue (Dunaway, 1983). Thus the molecular basis for the regulation of glycolysis in a given tissue lies in the control of the tissue-specific expression of the genes encoding these three isozymes.

To study the regulation of tissue specific expression of rabbit muscle PFK gene, the Chang laboratory has previously cloned the rabbit muscle PFK gene (Lee et al., 1987) and recently the two cDNAs of this gene (Li et al., 1990). These two cDNAs have an identical coding region, but divergent 5' untranslated regions. However, 1) we do not know whether these two cDNAs are derived from one or two promoters. If these cDNAs came from one promoter, their mRNAs would share the 5' end of the pre-mRNA. 2) the tissue specificity of these two cDNAs remains elusive.
IDENTIFICATION OF SIX cDNAs FOR THE rmPFK GENE

I have addressed the first question by cloning the 5' ends of the rmPFK mRNAs using anchored PCR technique. Five rmPFK cDNAs have been cloned by this technique. Using RT-PCR, an additional rmPFK cDNA has been cloned. Therefore, there are total of six rmPFK cDNAs (SECTION I). The sequences of the six cDNAs are consistent with the two previously cloned cDNAs (Li et al., 1990). For example, cDNA-1 identified in this study contains the entire 5' UTR of cDNA-B of Li et al. (1990), while the 5' UTR of cDNA-A identified by Li et al. (1990) is present in four of the six cDNAs cloned in this study.

THE SIX rmPFK mRNAs ARE EXPRESSED FROM AT LEAST THREE PROMOTERS OF A SINGLE GENE

Southern blot analysis of rabbit genomic DNA digested with restriction enzymes indicates that rmPFK is a single copy gene (Knaak and Chang, unpublished data). The human muscle PFK has also been shown to be a single copy gene (Sharma et al., 1990). Therefore, the six cloned rmPFK cDNAs are produced from a single gene.

We have mapped the transcription start sites of two downstream promoters of this gene. Promoter a initiates transcription at a major site 86 bp upstream of the ATG translational initiation codon, along with several minor sites (Fig. 16). The 5'-end of cDNA-1 identified by anchored PCR is at the major site. The transcription start sites of a second promoter initiate transcription from several sites
about 2 kb upstream of the ATG codon (Fig. 17). The 5'-end of cDNA-5 is at one of these transcription start sites. The transcription start sites of the promoter(s) giving rise to mRNA-c, mRNA-d1 and -d2 have yet to be determined. Since the 5' ends of these mRNAs are upstream of promoters a and b, they can not be transcribed from these promoters. Therefore, the rmPFK gene expresses these six mRNAs from at least three and possible four promoters. The promoter(s) that gave rise to mRNA-c, -d1, and -d2 can be mapped by both primer extension and S1 mapping, techniques that successfully mapped the transcription start sites of promoter a and promoter b.

**FOUR OF THE SIX RMPFK mRNAs ARE PRODUCED BY ALTERNATIVE PRE-mRNA SPlicing**

Alignment of the sequence of the cDNAs with that of the rmPFK genomic DNA showed that two alternative splicing events of two pre-mRNAs resulted in four of the six rmPFK mRNAs. First, mRNA-b1 is a splicing variant of mRNA-b2; more specifically, the first intron in the 5' UTR was spliced out as an intron in mRNA-b2, but retained in mRNA-b1 (Fig. 12). Also mRNA-d1 is a splicing variant of mRNA-d2; exon IB was not retained in mRNA-d2 (Fig. 12). Alternative splicing occurred in the 5' UTRs. The human muscle PFK gene has been shown to express two of its three mRNAs by alternative splicing (Fig. 7; Yamasaki et al., 1991). However, the alternative splicing events occurred in different pre-mRNAs in the muscle PFK genes of the two species. In the human
gene, the pre-mRNA transcribed by the downstream promoter is
alternatively spliced, while in the rabbit gene, the
equivalent (mRNA-a) is not alternatively spliced, although
5' UTR of the rmPFK mRNA-a and its human counterpart is
highly conserved (82% homology in a region of 450 bp,
including the first coding exon and 300 bp upstream of the
ATG codon). The alternatively spliced mRNAs of the rmPFK
gene are transcribed from promoter b and a promoter further
upstream.

IT IS UNLIKELY THAT THERE IS A TRANSCRIPTION START
SITE 20 TO 30 BP DOWNSTREAM OF THE TATAAAA SEQUENCE

The 5' end of cDNA-2 is 28 bp downstream of the TATAAAA
sequence. Primer extension using M-MLV reverse transcriptase
also showed a band at this position (y in Fig. 17B). However, S1 mapping analysis clearly showed no band at this
position (Fig. 17A). Furthermore, primer extension analysis
at 65°C using Tth DNA polymerase did not show a band at this
position (Fig. 17C). Therefore, it is very unlikely that
there is a transcription start site at this position. The
reverse transcription reaction using M-MLV reverse
transcriptase is probably prematurely terminated at this
position, resulting in a band in the primer extension
analysis and the 5'-end of cDNA-2 at this position.

It is intriguing that there is no transcription start
site 20 to 30 bp downstream of this TATAAAA sequence, since
it matches the TATA box consensus sequence. In addition,
there are multiple putative binding site for transcription
activators upstream of this sequence. For example, in the 190 bp DNA between promoter b and the TATAAAA sequence, there are three binding sites for the bHLH myogenic transcription activators, such as MyoD and myogenin (Weintraub et al., 1991).

There are several possibilities to explain the apparent lack of a transcription start site downstream of this TATAAAA sequence. First of all, there may be a specific repression sequence near this sequence. This can be investigated by making a construct fusing the chloroamphenical acetyltransferase (CAT) reporter 3' to the rmPFK genomic DNA sequence 180 bp upstream and 80 bp downstream of the TATAAAA sequence, since the promoter b is about 200 bp upstream of the TATAAAA sequence and the splice site at the 3' end of exon IB is about 88 bp downstream of the TATAAAA sequence. Linker scanning mutants can be created based on this construct. Transfection of these constructs to myoblast C2C12 cells and subsequent CAT assay will probably reveal the effect of certain mutations at a specific sequence on the repression of transcription initiation 20 to 30 bp downstream of the TATAAAA sequence, if there is a sequence-specific repressor of transcription initiation in this region of the rmPFK gene. In other words, certain mutation in this region will leads to transcription initiation, when a sequence-specific repressor is mutated. Of course, such mutant should be examined to find out that gain of transcription initiation is not due to the creation
of a positive element by the mutation. This can be done by mutate the putative repressor to other sequences.

Secondly, chromatin structure in this region might prevent transcription initiation. It has been shown in vitro that a minimal promoter containing the TATA-box is able to direct assembly of the transcription initiation complex onto the DNA template (Buratowski et al., 1989). The first step is the binding of transcription factor IID (TFIID) to the TATA-box, followed by the TFIIA, TFIIIB, RNA polymerase II, and TFIIE. This complex is able to initiate transcription in the presence of four dNTPs. The binding of TFIID to the TATA box is essential for transcription initiation. However, binding of TFIID to a DNA template containing a TATA box and subsequently the transcription initiation is severely inhibited after the nucleosome assembly (Workman et al., 1991; Meinterernst et al., 1990). However, the binding remains and potentiates transcription initiation if the nucleosome is assembled after the TFIID binds to the DNA (Workman and Roeder, 1987). Recently, it has been shown that purified yeast SWI/SNF protein complex can potentiate the binding of transcription activator GAL4 derivative to nucleosomal DNA (Cote et al., 1994). Imbalzano et al. (1994) have examined whether human SWI/SNF complex can facilitate the binding of TBP to nucleosomal DNA. These researchers constructed a DNA template of 149 bp containing two nucleosome positioning sequences (Shrade and Crothers, 1989) at the 5' end and the adenovirus major late promoter TATA
The phasing sequence favors specific positioning of the histone core. They constructed two variant templates (PH MLT9(+3) and PH MLT(+6)) positioned the TATA box either 3 or 6 bp, respectively, closer to the phasing sequences. After nucleosome assembly, TBP can not bind to the TATA box in all three templates. Human SWI/SNF complex has been shown to facilitate the binding of TBP to the TATA box in PH MLT (+3) template. However, the complex does not facilitate the binding of TBP to the TATA box in the PH MLT and PH MLT(+6) templates (Imbalzano et al., 1994). These data imply that if a TATA box is not at a favorable rotation relative to the nucleosome, even the SWI/SFN proteins can not facilitate the binding of TBP to the TATA sequence in a DNA template. The TATATAAA sequence of the rmPFK gene could be in such a position. There are three sequences near the TATATAAA sequence that match the consensus nucleosome positioning sequence. For example, sequence 5'-CCCCCAAT (28 downstream of the TATATAAA sequence) matches the strong nucleosome positioning consensus sequence (G/C)_3NN(A/T)_3 (Shrade and Crothers, 1989).

If the nucleosome is repressing transcription, a naked DNA template of the above mentioned construct can be used which has a CAT reporter fused 3' to an rmPFK gene fragment about 180 bp upstream and 80 bp downstream of the TATATAAA sequence. This template should initiate transcription in in vitro transcription system using histone-depleted nuclear
extract from rabbit muscle, since there are three putative MyoD binding sites in this rmPFK gene fragment.

THE SIX RMFPK MRNAS ARE NOT EXPRESSED IN SAME TISSUE-SPECIFIC MANNER

Northern blot and RT-PCR analysis revealed that the six rmPFK mRNAs are expressed in a tissue-specific manner: mRNA-b1, -b2, and -c are skeletal muscle specific (Figs. 14 and 15). mRNA-a was also found expressed only in skeletal muscle by Northern blot analysis (Fig. 14). However, RT-PCR experiment showed that this mRNA is also expressed in heart (Fig. 15A). This discrepancy might be due to the high sensitivity of RT-PCR compared to Northern blot analysis. mRNA-d1 was only detected in skeletal muscle, heart, and brain tissues, but not in liver and kidney, while mRNA-d2 was detected in all five rabbit tissues examined (Fig. 15). The hmPFK gene has also been shown by RT-PCR to be expressed in all seven human tissues examined (Nakajima et al., 1990).

The pattern of the tissue-specific expression of the rmPFK mRNAs described here is consistent with that of the rmPFK protein (Tsai and Kemp, 1973). The rmPFK has been shown to be the predominant PFK in rabbit skeletal muscle. This isozyme has also been shown to be expressed in brain, heart, kidney, and minimally in liver tissues (Tsai and Kemp, 1973). A similar pattern of expression of the muscle type PFK has also been shown in humans (Dunaway et al., 1988) and in rats (Dunaway and Kasten, 1987).
THE DIFFERENT 5' UTRS OF THE SIX RMFFK mRNAs MAY PLAY A ROLE IN THE STABILITY AND TRANSLATIONAL EFFICIENCY OF THESE mRNAs

The six different mRNAs of the rmPFK gene have different 5' UTRs. There is one short open reading frame (ORF) of 72 nucleotides in the 5' UTR of mRNA-c (for sequence, see Fig. 11). In the 5' UTR of mRNA-b1, there are three ORFs of 15, 12, and 60 nucleotides. Four short ORFs in the 5' UTR of the yeast GCN4 mRNA have been shown to regulate the translation of the GCN4 ORF (Hinnebusch et al., 1993). Whether the short ORFs in the 5' UTRs of the rmPFK mRNA-b1 and mRNA-c play any role in the stability and translational efficiency of the PFK ORF is not known. This could be addressed by the following experiments.

The 5' UTRs of the different rmPFK cDNAs will be fused 5' to the CAT coding region. These chimeras will be driven by a common promoter, such as adenovirus major late promoter. The stabilities of the chimeric mRNAs can be deduced by examining the levels of these chimeric mRNAs before and at different times after the addition of Actinomycin D, which stops transcription by RNA polymerase II. The effects of these different UTRs on the translational efficiency can also be studied using these constructs. The CAT enzyme activity of the different constructs and the steady state level of the mRNAs would reveal the translational efficiency of the different chimeric mRNAs.
PRIMER EXTENSION AT HIGH TEMPERATURE USING THERMOSTABLE TTH DNA POLYMERASE

In this dissertation research, I have developed a protocol for primer extension at high temperature. Primer extension at 50°C of rmPFK around the TATAAAA sequence resulted in prematurely terminated products when M-MLV reverse transcriptase was used. Higher temperature should destabilize the secondary structure and reduce the artifact. However, both AMV and M-MLV reverse transcriptases have been found inactivated at temperatures higher than 50°C (Gibco BRL catalog). Tth DNA polymerase has been reported to possess reverse transcriptase activity and can tolerate high temperature (Myers and Gelfand, 1991). Primer extension under the conditions in the existing protocol (Myers and Gelfand, 1991) resulted in a very faint band. After optimization, primer extension using Tth DNA at 65°C produced a band similar in size and intensity to that using M-MLV reverse transcriptase at 50°C. Using the optimized protocols developed in this study (SECTION II), primer extension at 65°C using Tth DNA polymerase resolved the problem caused by the secondary structure in the rmPFK mRNA (SECTION I). The improved method may be useful for other procedures such as cDNA library construction and anchored PCR. The usefulness of this protocol has been shown in anchored PCR experiment described in SECTION I of this dissertation.
REFERENCES


APPENDIX

RESTRICTION MAP OF THE 5' FLANKING REGION OF THE RABBIT MUSCLE PHOSPHOFRUCTOKINASE GENE

This map was generated using the GCG software (Zuker and Stiegler, 1981).

<table>
<thead>
<tr>
<th>C</th>
<th>N</th>
<th>Bc</th>
<th>S</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Av</td>
<td>B1</td>
<td>sMc</td>
<td>v</td>
<td>Mv</td>
</tr>
<tr>
<td>li</td>
<td>a</td>
<td>aRwi</td>
<td>i</td>
<td>sI</td>
</tr>
<tr>
<td>uJ</td>
<td>nI</td>
<td>J10F</td>
<td>J</td>
<td>pJ</td>
</tr>
<tr>
<td>II</td>
<td>iV</td>
<td>III</td>
<td>I</td>
<td>I I</td>
</tr>
</tbody>
</table>

TTAGCTCACTGCCTCACTCATTGGCACCCTCCAGGCTTTATACCTTTATGCCTCCGCTCGTA

AATCGAGTGCAGAAGTCCAGAAATCGGTGGGTCGCAAATGTGAATAGCAAGGGCAGACAT

TGTTGTGGAAATTGTGACGCAGATAATTGTGAGCGCATAAACATTTACACACAGAAACAG

ACAACACACCTTAACACTCAGCCTAATTACACACTCGCGCTCTAATGTTCTCTTTTGTC

CTAGTACATGATTGGAAATTGGCAGCTCCTGTAACCGCAGCCGAGCCGAGGAGAGACTGGG

GATACTGTACATATGCTAAGCTGACCCATGCTGCTCCGGCTCGAAATCACCGCCACCC

ACTGACGCGAGCGCCCGGCTGCCGCTGCTGCTCGCCTGCTCGCCTGCTGACCC

D a sAsHzvmt u vselMMvBNC ve chVoc BM
d c mcohIwh 9 iaIpsiscr ii enIR gw
e 8 FifaJ10a 6 JIhlpJ11F JI flJ1F lo
i I IIIIII III II II II

ACTGACGCGAGCGCCCGGCTGCCGCTGCTGCTGAGCGCCGAGGAGCTGGAG

TGACTCCGCTCGCCTCGGAGCCACTCGTCCCGGCGCCGACCCTCCGCTGCTGACCCTC
```
<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>D</td>
<td>MAV</td>
<td>MT</td>
<td>j</td>
<td>1</td>
<td>ep</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>d</td>
<td>nii</td>
<td>hh</td>
<td>e</td>
<td>1</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>e</td>
<td>luJ</td>
<td>aa</td>
<td>P</td>
<td>I</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>I</td>
<td>III</td>
<td>II</td>
<td>I</td>
<td>I</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```

```
```
### Enzymes that do cut:

<table>
<thead>
<tr>
<th>AccI</th>
<th>AcII</th>
<th>AflIII</th>
<th>AgeI</th>
<th>AluI</th>
<th>AlwI</th>
<th>Alw21I</th>
<th>Alw9I</th>
</tr>
</thead>
<tbody>
<tr>
<td>ApaI</td>
<td>ApaBI</td>
<td>AvoI</td>
<td>AvoII</td>
<td>AvaI</td>
<td>AvrII</td>
<td>BanI</td>
<td>BanII</td>
</tr>
<tr>
<td>BbsI</td>
<td>BbvI</td>
<td>BccI</td>
<td>BceII</td>
<td>Bcfl</td>
<td>BciI</td>
<td>BglI</td>
<td>BgiI</td>
</tr>
<tr>
<td>BpmI</td>
<td>Bpu101I</td>
<td>BsaAI</td>
<td>BsaHI</td>
<td>BsmAI</td>
<td>BsmBI</td>
<td>BsmAI</td>
<td>BsmBI</td>
</tr>
<tr>
<td>BscGI</td>
<td>BseRI</td>
<td>BsiI</td>
<td>BsiEI</td>
<td>BsiI</td>
<td>BsmI</td>
<td>BsmI</td>
<td>BsmAI</td>
</tr>
<tr>
<td>BamFI</td>
<td>BsoFI</td>
<td>Bsp21I</td>
<td>Bsp1207I</td>
<td>BspEII</td>
<td>BspGI</td>
<td>BspMI</td>
<td>BspMI</td>
</tr>
<tr>
<td>BarBI</td>
<td>BarDI</td>
<td>BarFI</td>
<td>BstXI</td>
<td>Bsu36I</td>
<td>Cac8I</td>
<td>CjeI</td>
<td>CjePI</td>
</tr>
<tr>
<td>CviJI</td>
<td>CviRI</td>
<td>DdeI</td>
<td>DpnI</td>
<td>DraI</td>
<td>DraIII</td>
<td>DrdI</td>
<td>DrdII</td>
</tr>
<tr>
<td>DsaI</td>
<td>EaeI</td>
<td>EalI</td>
<td>Eam105I</td>
<td>EarI</td>
<td>EciI</td>
<td>EcoN1</td>
<td>EcoG109I</td>
</tr>
<tr>
<td>EcoRI</td>
<td>EcoRII</td>
<td>FauI</td>
<td>FokI</td>
<td>PspI</td>
<td>QdI</td>
<td>HaeI</td>
<td>HaeIII</td>
</tr>
<tr>
<td>HaeIII</td>
<td>HglI</td>
<td>HhaI</td>
<td>HincII</td>
<td>Hinfl</td>
<td>HpaI</td>
<td>Hphi</td>
<td>KpnI</td>
</tr>
<tr>
<td>MaeII</td>
<td>MaeIII</td>
<td>MboII</td>
<td>MnlI</td>
<td>MseI</td>
<td>MspI</td>
<td>MspAl</td>
<td>MspAl</td>
</tr>
<tr>
<td>MwoI</td>
<td>NciI</td>
<td>NcoI</td>
<td>NheI</td>
<td>NlaIII</td>
<td>NlaIV</td>
<td>NruI</td>
<td></td>
</tr>
<tr>
<td>NapI</td>
<td>PflMI</td>
<td>PfrI</td>
<td>PmlI</td>
<td>PshAl</td>
<td>Psp5II</td>
<td>PstI</td>
<td>PvuII</td>
</tr>
<tr>
<td>RcaI</td>
<td>Rsal</td>
<td>SacI</td>
<td>SacII</td>
<td>Sapi</td>
<td>Sau96I</td>
<td>Sau3AI</td>
<td></td>
</tr>
<tr>
<td>SclFI</td>
<td>SfaNI</td>
<td>SfiI</td>
<td>SfaI</td>
<td>Stfl</td>
<td>StyI</td>
<td>TaqI</td>
<td></td>
</tr>
<tr>
<td>TaqII</td>
<td>TfiI</td>
<td>Thal</td>
<td>TesI</td>
<td>Tsp45I</td>
<td>Tsp509I</td>
<td>Tth111I</td>
<td>XbaI</td>
</tr>
<tr>
<td>XmnI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Enzymes that do not cut:

<table>
<thead>
<tr>
<th>AatII</th>
<th>AflIII</th>
<th>Alw44I</th>
<th>AscI</th>
<th>BaeI</th>
<th>BanHI</th>
<th>BclI</th>
<th>BglII</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bpu102I</td>
<td>BsaBI</td>
<td>BsmBI</td>
<td>BspLUII</td>
<td>BsrGI</td>
<td>BssHI</td>
<td>Bst107I</td>
<td>BstEII</td>
</tr>
<tr>
<td>BstYI</td>
<td>ClaI</td>
<td>Eco47III</td>
<td>Eco57I</td>
<td>EcoRV</td>
<td>FaiI</td>
<td>HgiII</td>
<td>HindIII</td>
</tr>
<tr>
<td>MluI</td>
<td>MmeI</td>
<td>MacI</td>
<td>MnlI</td>
<td>NdeI</td>
<td>NgoAI</td>
<td>NotI</td>
<td>NsiI</td>
</tr>
<tr>
<td>NapV</td>
<td>PacI</td>
<td>Pfl100I</td>
<td>PmeI</td>
<td>Psp106I</td>
<td>PvuII</td>
<td>RsrII</td>
<td>Sall</td>
</tr>
<tr>
<td>ScaI</td>
<td>SexAI</td>
<td>SfiI</td>
<td>SfiI</td>
<td>SgrAI</td>
<td>SnaBI</td>
<td>SpeI</td>
<td>Sphi</td>
</tr>
<tr>
<td>Sse8387I</td>
<td>SspI</td>
<td>SunI</td>
<td>Swal</td>
<td>Tth111I</td>
<td>VspI</td>
<td>XcmI</td>
<td>XhoI</td>
</tr>
</tbody>
</table>
VITA

Shouhua Xiao was born on February 10, 1963 in Nanchang County, Jiangxi Province, China. He attended Hoshan High School and received an Outstanding High School Student honor by the Nanchang City, Jiangxi Province. He attended Jiangxi Agricultural University and majored in Horticulture. He graduated with the highest GPA in the class of 40 students receiving a B.S. degree in July of 1983. In the Fall semester of the same year, he entered a three-year master program in the Graduate School of Beijing Agricultural University with a specialty in Plant Genetics and Breeding. Since the completion of his graduate studies with a M.S. degree in July of 1986, he had worked for three years in the Institute of Botany, Chinese Academy of Sciences, Beijing, China.

In August, 1989, Shouhua started his Ph.D. studies at Louisiana State University first in the Department of Plant Pathology and Crop Physiology. In January, 1991, he switched to the Ph.D. program in the Department of Biochemistry and continued his Ph.D. studies under the direction of Dr. Simon H. Chang. Upon the completion of his degree requirement he will receive a Ph. D. degree. He has accepted a postdoctoral position in Dr. James L. Manley’s laboratory at Columbia University in New York City.
Candidate: Shouhua Xiao

Major Field: Biochemistry

Title of Dissertation: The Rabbit Muscle Phosphofructokinase Gene: Six Types of mRNAs Are Tissue-Specifically Expressed by Multiple Promoters and Alternative Splicing

Approved:

[Signature]
Major Professor and Chairman

[Signature]
Dean of the Graduate School

EXAMINING COMMITTEE:

[Signature]

[Signature]

[Signature]

Patrick J. DiMeio