Characterization of the Gene for Rat Phosphorylase Kinase Catalytic Subunit*

Kendra C. Cawley, Carol G. Akita, Karen L. Angelos, and Donal A. Walsh

From the Department of Biological Chemistry, School of Medicine, University of California, Davis, California 95616

Phosphorylase kinase, a key enzyme in glycogen metabolism, has a subunit composition of \((\alpha\beta\gamma)\), in which the \(\alpha\) and \(\beta\) subunits are regulatory, \(\delta\) is calmodulin, and the \(\gamma\) subunit is catalytic. As one segment of our studies on the regulation of the expression of phosphorylase kinase subunits, we present in this report the structure of the gene for the catalytic \(\gamma\) subunit. The gene extends over 16 kilobase pairs (kb) of DNA, and contains eight introns within the coding region plus one 3.3-kb intron upstream in the 5' untranslated region. Within this first intron, and also upstream of the transcription start site, are sequences homologous to defined regulatory elements, including some found in other muscle-specific genes. The positions of intron splice junctions for this gene have been compared with similar data for other protein kinase genes. A somewhat unexpected finding for the \(\gamma\) subunit is that two of the splice junctions fall in the midst of highly conserved strings of amino acids, both of which have been nominally defined as functional domains for the protein kinases and appear to make key contributions to substrate binding and phosphotransferase catalysis.

Phosphorylase kinase, a key enzyme in glycogen metabolism, has a subunit composition of \((\alpha\beta\gamma)\), in which the \(\alpha\) and \(\beta\) subunits are regulatory (phosphorylated by the cAMP-dependent protein kinase), \(\delta\) is calmodulin, and the \(\gamma\) subunit is catalytic and contains binding sites for the regulatory subunits (1). As one segment of our studies on the regulation of the expression of phosphorylase kinase subunits, we present in this report the structure of the gene for the catalytic \(\gamma\) subunit. The gene extends over 16 kilobase pairs (kb) of DNA, and contains eight introns within the coding region plus one 3.3-kb intron upstream in the 5'-untranslated region. Within this first intron, and also upstream of the transcription start site, are sequences homologous to defined regulatory elements, including some found in other muscle-specific genes. The positions of intron splice junctions for this gene have been compared with similar data for other protein kinase genes. A somewhat unexpected finding for the \(\gamma\) subunit is that two of the splice junctions fall in the midst of highly conserved strings of amino acids, both of which have been nominally defined as functional domains for the protein kinases and appear to make key contributions to substrate binding and phosphotransferase catalysis.

**EXPERIMENTAL PROCEDURES**

**General Procedures**

**DNA Sequencing**—Dideoxynucleotide sequencing was performed with 5'-32P-dATP (Amersham Corp.), using Sequenase (U.S. Biochemical Corp.) according to the manufacturer's recommendations. Plasmid DNA used for sequencing template was purified by CsCl gradient centrifugation (10), and when used as template, was denatured in 0.2 M NaOH and ethanol-precipitated (11) prior to the sequencing reaction.

**Polymerase Chain Reaction**—PCR was carried out in 25-μl reactions containing ~10 ng of λ DNA (clones γG6, γG23, or γG25), 50 mM Tris-HCl, pH 8.3, 1.5 mM magnesium chloride, 50 μM each dNTP, 1 μM each primer, and 0.625 unit of Taq polymerase (Perkin-Elmer Cetus). Some reactions also contained 0.25 unit of Perfect Match™ (Stratagene). One cycle of 3 min at 94°C, 1 min at the appropriate annealing temperature (based on the % GC content of primers, or empirically optimized), and 1-4 min at 72°C, was followed by 25-30 cycles of 30 s at 94°C, 1 min annealing, and 1-4 min at 72°C. Specific conditions were optimized for each primer pair to give a single major product.

**Oligonucleotide Primers**—The nomenclature and sequence of all oligonucleotide primers used in this study are given in Table I with reference to their matching location on the \(\gamma\) subunit gene.

**Overall Sequencing Strategy for the \(\gamma\) Subunit Gene**

**Genomic Clones, Mapping, and Sequencing**—A Wistar rat genomic library in λDash (Stratagene) was screened using the 1.4-kb EcoRI fragment containing the complete coding region for rat phosphorylase kinase \(\gamma\) subunit, plus 80 bp of 5'-untranslated region (UTR) and 200 bp of 3'-UTR (12, 13). Four clones which hybridized to this probe were mapped, using 32P-end-labeled oligonucleotides specific for the T3 and T7 promoter sequences (which flank the insert in the λDash vector) to probe Southern blots of partial restriction enzyme digests of the λ DNA. Specific procedures used in cloning, mapping, and subcloning specific restriction fragments into M13 mp18 or pBluescript (Stratagene) are described in Ref. 10. The four genomic clones (γG20, γG6, γG23, and γG25), which hybridized to the rat phosphorylase kinase \(\gamma\) cDNA and mapped with restriction enzymes BamHI, HindIII, and XhoI, were found to represent overlapping segments of the phosphorylase kinase \(\gamma\) gene (Fig. 1). An additional cluster of HindIII and BamHI sites upstream from the sequence indicated on Fig. 1 helped to confirm the alignment of γG20 and γG6. All four

---

*This work was supported by National Institutes of Health Grant DK13613. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M98286 and M98287.

---

**EXPERIMENTAL PROCEDURES**

**General Procedures**

**DNA Sequencing**—Dideoxynucleotide sequencing was performed with 5'-32P-dATP (Amersham Corp.), using Sequenase (U.S. Biochemical Corp.) according to the manufacturer's recommendations. Plasmid DNA used for sequencing template was purified by CsCl gradient centrifugation (10), and when used as template, was denatured in 0.2 M NaOH and ethanol-precipitated (11) prior to the sequencing reaction.

**Polymerase Chain Reaction**—PCR was carried out in 25-μl reactions containing ~10 ng of λ DNA (clones γG6, γG23, or γG25), 50 mM Tris-HCl, pH 8.3, 1.5 mM magnesium chloride, 50 μM each dNTP, 1 μM each primer, and 0.625 unit of Taq polymerase (Perkin-Elmer Cetus). Some reactions also contained 0.25 unit of Perfect Match™ (Stratagene). One cycle of 3 min at 94°C, 1 min at the appropriate annealing temperature (based on the % GC content of primers, or empirically optimized), and 1-4 min at 72°C, was followed by 25-30 cycles of 30 s at 94°C, 1 min annealing, and 1-4 min at 72°C. Specific conditions were optimized for each primer pair to give a single major product.

**Oligonucleotide Primers**—The nomenclature and sequence of all oligonucleotide primers used in this study are given in Table I with reference to their matching location on the \(\gamma\) subunit gene.

**Overall Sequencing Strategy for the \(\gamma\) Subunit Gene**

**Genomic Clones, Mapping, and Sequencing**—A Wistar rat genomic library in λDash (Stratagene) was screened using the 1.4-kb EcoRI fragment containing the complete coding region for rat phosphorylase kinase \(\gamma\) subunit, plus 80 bp of 5'-untranslated region (UTR) and 200 bp of 3'-UTR (12, 13). Four clones which hybridized to this probe were mapped, using 32P-end-labeled oligonucleotides specific for the T3 and T7 promoter sequences (which flank the insert in the λDash vector) to probe Southern blots of partial restriction enzyme digests of the λ DNA. Specific procedures used in cloning, mapping, and subcloning specific restriction fragments into M13 mp18 or pBluescript (Stratagene) are described in Ref. 10. The four genomic clones (γG20, γG6, γG23, and γG25), which hybridized to the rat phosphorylase kinase \(\gamma\) cDNA and mapped with restriction enzymes BamHI, HindIII, and XhoI, were found to represent overlapping segments of the phosphorylase kinase \(\gamma\) gene (Fig. 1). An additional cluster of HindIII and BamHI sites upstream from the sequence indicated on Fig. 1 helped to confirm the alignment of γG20 and γG6. All four

---

*The abbreviations used are: kb, kilobase pair(s); bp, base pair(s); PCR, polymerase chain reaction; MLC, myosin light chain; AchR, acetylcholine receptor; UTR, untranslated region; PIP₂, 1,4-piperazinediethanesulfonic acid.
clones were also mapped with SacI. Southern blots were hybridized with the oligonucleotides yP11, yP13, and yP23 to determine the clones were also mapped with SacI. Southern blots were hybridized to different fragments, the 3.8-kb SacI fragment to which yP11 hybridized was subcloned into a plasmid (pyG3.8wt) for additional restriction mapping.

A combination of the following methods was used to obtain the full map shown in Fig. 1. Sequencing in M13—Several fragments were subcloned directly into M13 for sequencing. The 0.23-kb EcoRI/Xhol and the 0.56-kb BamHI fragments, which hybridized to yP11 and yP13, respectively, were sequenced in entirety on both strands. The 0.56-kb BamHI/HindIII fragment and the 0.74-kb XhoI/BamHI fragments were subcloned into M13 and sequenced on one strand primer yP22 (cf. Table I and Fig. 1) was used to complete the sequence of the latter. For some fragments only partial sequence was obtained. These include the 2.0-kb SacI/NcoI fragment containing Exon 1, the 1.1-kb SacI/BamHI fragment containing the 3′ portion of Intron A, the 1.2-kb BamHI/SacI fragment which contains the unique HindIII site, and the 0.56-kb BamHI/XhoI fragment containing the 3′ portion of Intron H. Sequencing more than 350 bp from any one template was achieved with sequence-specific primers, yP12, yP26, yP29, and yP14 (Table I). These procedures identified Exons 1, 2, 3, 9, and 10.

Shotgun Cloning of MboI Fragments—Two BamHI fragments of 3.8 and 4.8 kb, derived from the central region of the gene, were isolated from a lambda vector and subcloned into the BamHI site of M13 mp18. Screening with full-length y cDNA (the 1.4-kb EcoRI fragment containing Exon 1, the 1.1-kb SacI/BamHI fragment containing the 3′ portion of Intron A, 1.2-kb BamHI/SacI fragment which contains the unique HindIII site, and the 0.56-kb BamHI/XhoI fragment containing the 3′ portion of Intron H) sequences only partial sequence was obtained. These include the 2.0-kb SacI/NcoI fragment containing Exon 1, the 1.1-kb SacI/BamHI fragment containing the 3′ portion of Intron A, the 1.2-kb BamHI/SacI fragment which contains the unique HindIII site, and the 0.56-kb BamHI/XhoI fragment containing the 3′ portion of Intron H. Sequencing more than 350 bp from any one template was achieved with sequence-specific primers, yP12, yP26, yP29, and yP14 (Table I). These procedures identified Exons 1, 2, 3, 9, and 10.

Determination of Intron Sizes—PCR was used to determine the sizes of Introns A, C, D, E, F, and H, using combinations of the flanking primers shown in Fig. 1. Each intron size was determined with at least two different sets of primers. In most cases, at least two different primer pairs were used. For example, for Intron F PCR with yP18 + yP20 provided a direct measure of the size of F, and yP18 + yP21 amplified a larger product that contained F and also contained Intron G, whose size had been obtained by direct sequencing. The length of the gene thus determined by the addition of introns and exons matched the length as determined by restriction analysis, from the 5′ EcoRI site to the unique Xhol site, within 5%. The size of Intron A, determined by PCR with yP11 and yP13, matched that determined by restriction analysis.

The combination of Sanger sequencing and primer extension identified and sized nine introns in the y subunit gene, eight within the coding sequence and one in the 5′ UTR. The sequences flanking each of the introns, along with the intron sizes and methods used for determining them, are given in Table I. The 5′ and 3′ boundaries of each were the expected g and ag sequences. The 5′ end of the gene is detailed in Fig. 2.

**Determination of Transcription Initiation Site**

**Primer Extension for Determination of 5′ End—Oligonucleotide primer yP12 was labeled at the 5′ end with [γ-32P]-ATP using T4 polynucleotide kinase (New England Biolabs) to a specific activity of 2.7 × 10^6 dpm/mmol. 5 × 10^6 dpm of [γ-32P]-yP12 was co-purified with 15 μg of rat muscle RNA, and resuspended in hybridization buffer containing 80% formamide, 0.4 M sodium chloride, 1 mM EDTA, and 40 mM PIPES, pH 6.4, incubated 10 min at 70 °C and then overnight at 30 °C. The RNA and primer were precipitated in ethanol and resuspended in 20 μl of 10 mM Tris-HCl, pH 8.0, 1 mM EDTA, and prepared for electrophoresis by adding an equal volume of gel sample buffer (95% formamide, 20 mM EDTA, 0.65% bromphenol blue, and 0.05% xylene cyanol). Using the yP12 oligonucleotide primer in this procedure a single major band was obtained (Fig. 3). Comparison with the standards in particular with the sequence obtained using plasmid pyG3.8wt and yP12 primer would place the primary transcription start site at 155 bases upstream of the 3′ end of the yP12 oligonucleotide primer (indicated as +1 in Fig. 2). This proposed position for the transcription initiation site would be supported by the sequence, in that there are TATA-like and CCAAT-like sequences at appropriate positions upstream.

**Confirmation of the Transcription Start Site by PCR**—The transcription start site, identified by the size of the product obtained by primer extension, was verified by a PCR procedure using oligonucleotides selected on each side of the proposed start site. This allowed a verification of the size and location of the upstream and downstream products. PCR reactions were examined as template either muscle cDNA, prepared by random-reverse transcription of skeletal muscle RNA, or genomic DNA (as control templates for appropriate product size). Random hexamers were used to prime reverse transcription from rat skeletal muscle total RNA. Reactions containing 200 ng of RNA, 100 pmol of random hexamers (in 20 μl), 25 mM Tris-HCl, pH 8.3, 75 mM potassium chloride, 3 mM magnesium chloride, 10 mM dithiothreitol, 6 units of RNasin (Promega), 0.5 mM each dNTP, and 200 units of Moloney murine leukemia virus reverse transcriptase (BRL), were incubated at 37 °C for 1 h, then at 68 °C for 5 min to inactivate the enzyme. Aliquots containing cDNA from 20 μl of RNA were used directly as template in PCR reactions with the indicated primer pairs. PCR conditions were as described under general procedures except for the cycling protocol: 30 cycles of 30 s at 94 °C, 1 min at 50 °C, and 1 min at 72 °C. For these PCR reactions the upstream primer was yP24, yP25–yP26, or yP27, and the downstream primer was either yP12 or yP13. The relationship of these primers to the genomic DNA sequence and the transcription start site identified by primer extension is illustrated in Fig. 4A. The results of these various PCR combinations are shown in panel B. In lanes 2–5, in which the template for the PCR reactions was genomic DNA and yP12 served as the downstream primer, the expected size amplification products were plainly evident for each of the upstream primers. Lanes 6–9 show PCR reactions using the same primer combinations but with reverse-transcribed muscle RNA as template. The expected products appear in reactions using yP24 and yP25, but no product was evident for either yP26 or yP27.
expected if the transcription start site is between the sites to which yP25 and yP26 bind, thus strongly supporting the primary location of the 5' end of the transcript as determined by the primer extension experiments. Lanes 10-13 of Fig. 4B show reactions using the same preparation of reverse-transcribed muscle RNA and the same set of upstream primers as used in lanes 6-9, but with yP13 instead of yP12 as the downstream primer. Control reactions using genomic DNA as template were not examined in this case because the products would have included the 3.3-kb Intron A and the conditions for PCR would have had to be significantly altered to obtain the larger product. The results of this experiment in general support the observations obtained with yP12; a strong signal was obtained with both yP24 and yP25 and no signal was detectable with yP27. A weak band of the appropriate size was detected with yP13/yP26, in contrast to the total absence of any band observable for the yP12/yP26 primer pair. This might suggest that there might be some alternative transcription start sites upstream of the primary site. If so, they probably represent only a small percentage of the total γ mRNA, since the amount recovered (after 30 cycles of amplification) is clearly less than that for the γP25/γP13 product of similar size. No product was detectable of a size compatible with obtaining a product with γP26 but not with γP27 in the primer extension experiments (Fig. 4F: 175-330 bp). Some minor bands of greater size (>600 bp) were detectable in the primer extension experiment (Fig. 3). If these represented true transcription start sites, it would have resulted in the formation of a PCR product with γP27, and none was observed. (The small products in lanes 8 and 9 of Fig. 4 are primer-dimer artifacts which seem to be made only when no better template for the primers is present. Adjusting PCR conditions to eliminate the artifact did not, however, allow an appropriate product from γP26 and γP12.)

RESULTS AND DISCUSSION

The Structure of the γ Subunit Gene and Potential Regulatory Elements—The gene for the γ subunit of phosphorylase kinase extends over at least 16 kb of DNA, and contains nine introns, one of which is in the 5'-untranslated region (Fig. 1). In Fig. 2 is presented a partial sequence of the upstream end of this phosphorylase kinase γ subunit gene, starting ~400 bp upstream of the beginning of Exon 1, proceeding through Exon 1, and including 343 bp of the 5' end and 1900 bp of the 3' end of Intron A, the full sequence of Exon 2, Intron B, and 101 bp of Intron C. (The control regions of a number of muscle-specific genes have been characterized and found to interact with a complex of transcription factors which seem to be made only...)

TABLE II

<table>
<thead>
<tr>
<th>5' Boundary</th>
<th>Intron</th>
<th>Size (kb)</th>
<th>Method</th>
<th>3' Boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3.3</td>
<td>RM, PCR</td>
<td>tgaatccttcttccaGGTGTC...</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>0.42</td>
<td>SEQ</td>
<td>ttcttgccccacaggGGGAGT...</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0.39</td>
<td>PCR</td>
<td>ctttcttccttcagTCAGC...</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>4.2</td>
<td>PCR</td>
<td>cttttgccccacaggGATGA...</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>2.05</td>
<td>PCR</td>
<td>gatctccccccaggAAAGAT...</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>0.86</td>
<td>PCR</td>
<td>tctcttgtatgtcaggAGGT77...</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>0.144</td>
<td>SEQ</td>
<td>cccgcctttcccaggGTTGAG...</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>0.86</td>
<td>PCR</td>
<td>ttcttgccccacaggGTAC7...</td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>0.291</td>
<td>SEQ</td>
<td>gcgtctccttgtccagGTACAT...</td>
<td></td>
</tr>
</tbody>
</table>
FIG. 2. Partial sequence of the phosphorylase kinase γ subunit gene. Nucleotides in lowercase represent intron sequences. Restriction sites for BamHI, HindIII, and EcoRI, and the ATG, are indicated in **bold** type. The TATA-like and CCAAT-like elements are shown in **bold italics**. Primers γP12, γP14, γP28, and γP29 are indicated by arrows. Single underlining indicates regions of homology shared with muscle glycogen phosphorylase. A 15-bp direct repeat is shown with double underlining. Boxes enclose two AP-2 sites (at -270 and -88), two E-box homologies (at +270 and +299), a sequence partly homologous to the AT-rich MEF-2 binding site (-213), and the heptameric AChR homology (+164).

FIG. 3. Determination of transcription initiation by primer extension. The products of reverse transcription from muscle RNA with primer γP12 (as described under "Experimental Procedures") are shown in the center lane (arrow). Dideoxynucleotide sequencing reactions (with α-32P-ATP) are for M13 mp18 with the -40 Universal Primer (right) as molecular weight standards and for γG3.8 with the γP12 primer (left). The latter gives the sequence that would correspond to the product extended from RNA using γP12, and the sequence surrounding the major product of reverse transcription is shown at the far left. Standard lanes flanking the sequencing reactions contain HaeIII-digested φX174, end-labeled by T4 polynucleotide kinase using [γ-32P]-ATP.

set of factors which cooperate to direct muscle-specific expression (reviewed in Ref. 16). Examination of the γ subunit sequence showed three regions with close homology to those of muscle-specific enhancer elements. The boxed sequences starting at +270 and +299 conform quite closely to the pair of sequences which have been identified for MyoD binding to muscle creatine kinase (5’-AACACCTGCTG and 5’-AGACATGTGGCTG, respectively; identity with γ shown by **boldface** letters), one or both of which are required for muscle-specific enhancer activity (17-19). Furthermore, the sequence of base pairs +272 to +280 of γ is identical to the last and the most important of three homologous and closely spaced sequences, which have been identified as the binding site for MyoD to the myosin light chain enhancer (MLC1/SE). Mutation of this site in the MLC gene dramatically diminishes the effects of MyoD, myogenin, and other related myogenic factors on the MLC1/3 enhancer-driven
expression of linked genes (CAT) (20). This sequence in γ (starting at +270) also conforms closely to the two nucleotide consensus sequences of 5'-AACAG(T/C)TGTT and 5'-N(A/C/G)CACCTGTT, which have been identified by PCR enhancement of oligonucleotides selected from random synthetic oligonucleotides by specific binding to myogenin (21) and E2A/MyoD dimers (22), respectively. These two sequences starting at +270 and +259 in γ conform to the minimal "E-box" consensus of 5'-CANNNTG (23), and there are in fact several places either upstream of +1 or elsewhere in the γ subunit where this motif occurs. The two selected, however, clearly have the noted additional homology with specifically defined functional enhancer elements. The E-box sequences appear to be targets for a group of factors involved in tissue-specific and developmental gene transcription. Muscle-specific factors such as MyoD, myogenin, and myf5, as well as the ubiquitous E2A proteins (such as E12 and E47, first described as factors interacting with immunoglobulin enhancers), share two characteristic domains: a helix-loop-helix motif, which is important for dimerization, and an adjacent basic region, which mediates binding of dimers to the DNA at these sequences (24-26). Paired E-boxes, as possibly seen here for γ, are found in several muscle-specific enhancers and promoters. For both muscle creatine kinase and MLC, cooperative interaction between the pairs of myoD-binding sites appears to be essential for transcriptional control (19, 20). Interestingly, the two E-box sequences of the γ subunit gene occur within the first intron; this is noteworthy because a number of other muscle-specific genes similarly have large introns within the 5'-UTR, and in at least two cases, muscle creatine kinase and troponin I, this intron has been shown to contain enhancer activity (24, 27).

In addition to these two E-box sequences, there are other similarities in the γ subunit gene with established regulatory sequences. The 7-bp sequence beginning at +164 in Intron A is identical to a 7 bp motif that is present in regulatory regions of the mouse acetylcholine receptor (AChR) δ subunit gene, and the genes for chick AChR α subunit and mouse muscle creatine kinase (28). In murine AChR this motif is repeated (imprecisely) four times within a 50-bp stretch. At +213 in the γ subunit gene there is an 8-bp sequence (5'-CTAAAAAT), which is identical to the 5' end of the A+T-rich MEF-2 binding site that has been found in a number of muscle-specific genes (29). The consensus sequence for binding MEF-2, however, was reported as 5'-CTAAAATACAG, and specifically lacking from γ is the underlined A, which by mutational analysis appeared to be essential for regulatory protein binding. The consequences of occupation of this site for muscle-specific gene expression have not been determined, but a role in conjunction with MyoD and/or myogenin binding has been proposed (29).

The γ genomic sequence has also been compared with that for muscle glycogen phosphorylase (30), both of which appear to be coordinately regulated at the level of transcription in regenerating muscle tissue (8). The results were obtained using a Pearson and Lipman homology search (31) (carried out using the FASTA program in the Wisconsin Genetics Computer Group (CGG)). Two sequences of homology between the two have so far been identified. The first (5'-TTTTTTTTTTTTTTTTGAG; upper sequence γ, lower sequence phosphorylase) is at -368 in γ and at -1076 in phosphorylase. The second sequence, 5'-GACCTTTGCGA-TTCTGCT/TCAC, starts at +105 in γ and at -1042 (in reversed orientation) in phosphorylase. Whether or not these sequence similarities have any functional significance for transcriptional expression remains to be determined. The 10-bp repeated sequence, which was found to be important for appropriate transcription in the glycogen phosphorylase gene (30), is not present in the γ sequence so far identified.

This examination of the γ subunit sequence certainly raises the possibility that several of the consensus sequences identified may play a regulatory role in γ subunit expression. Which, if any of these, is important is under current investigation.

Comparison of Intron Splice Junctions among Several Protein Kinases—The very large family of protein kinases share a considerable degree of amino acid sequence homology. Based upon this homology, and on particular groups of highly conserved amino acids, Hanks et al. (32) have aligned 65 of the protein kinases, defining in the sequences 11 domains (I-XI). In Fig. 5 the amino acid sequences of three protein kinases for which intron splice junction positions have been determined (including the data presented here) are fitted into the alignment of Hanks et al. (32). The solid bars and helices above the sequence represent β-strands and α-helix segments, respectively, as determined recently from the crystal structure of the mouse CAM-dependent protein kinase catalytic (Co) subunit (33). Junctions between the domains most frequently represent linkages between these defined structural elements, which, from the crystal structure of the CAM-dependent protein kinase, are primarily located on the exterior of the protein where variable lengths are most readily accommodated. The highly conserved amino acids among the protein kinases appear to be primarily those involved in ATP binding and phosphotransferase catalysis (33).
Although there are, at present, only a few kinase genes for which the intron positions have been determined, some interesting observations can be made by comparing the known junctions with established functional and/or structural domains for the kinases. With the cAMP-dependent protein kinases there appears to be a significant evolutionary conservation of intron splice junctions. The catalytic subunit of the cAMP-dependent protein kinase from nematode (34) has six introns, three less than mouse Ca (Ref. 35 and Fig. 5), but the six retained introns are in positions identical to those seen for the mouse enzyme (see Introns A, B, C, D, and F in Fig. 5, and Intron I at residue 399, not shown). These six introns are at exactly the same nucleotide in both genes even when the amino acids are not conserved or when the intron splits a codon. (34). This would suggest that the positions of splice junctions are well maintained on a fairly lengthy time scale, at least for this particular enzyme.

In contrast, between the different protein kinases, as evidenced by the comparison shown in Fig. 5 of the cAMP-dependent protein kinase catalytic subunit (35), casein kinase II (36), and the γ subunit of phosphorylase kinase (this study), there is no apparent conservation of intron splice sites. Examination of the casein kinase and the cAMP-dependent protein kinase sequences might suggest that the intron positions in general respect the domains defined by the highly conserved amino acids. Were this to be so, this would tend to support the notion that exon exchange may have provided evolving proteins with domain cassettes (37, 38). The positions of the intron junctions in phosphorylase kinase γ subunit would argue that the situation is clearly not that simple. By sequence, phosphorylase kinase γ subunit and the mammalian cAMP-dependent protein kinase appear to be much closer to each other in evolution than either is to casein kinase II (32, 36). However, in the γ gene two of the intron junctions interrupt highly conserved strings of amino acids. Intron B is located at Arg7 between two highly conserved glycines. This region of the protein kinase, designated also as the P-loop, is conserved in a broad range of kinases and the residues of the P-loop serve as one set of anchor points for the peptide bound by this region of the protein (2).
dependent protein kinase in the Asp adjacent to this residue is centrally involved in protein substrate binding (39).

Overall examination of the CAMP-dependent protein kinase would suggest that exon exchange is unlikely to be the mechanism from which the multiple protein kinases have arisen. In addition to the critical structural role that they contribute, the amino acids of the kinases serve four other functions, ATP recognition, catalysis, substrate recognition, and for many kinases regulatory domain recognition. The first two are the features conserved between the kinases; the latter two are not. Examination of the CAMP-dependent protein kinase structure (33, 39) shows that there is considerable interspersion between residues contributing to these different functions, often with adjacent amino acids being involved in distinct activities. This would appear to make it unlikely that exon exchange could have led to the differences in protein substrate specificities of the different protein kinases. Interesting in this regard are the structures of two homologous but clearly distinct genes for CAMP-dependent protein kinase in Drosophila (40), because for both genes kinase domains I–X are found on a single exon. If the several kinase domains had been acquired by exon shuffling, then the CAMP-dependent kinase (in Drosophila at least) must have since lost any intervening introns. In any case it seems likely that at least some of the introns in the γ gene (in particular Introns B and G) were inserted subsequently to the organization of the critical elements common to kinase structure and function.

It is interesting to note that the carboxyl terminus of the γ subunit, which contains two distinct calmodulin binding domains, is located on a single exon. The calmodulin domain between Gly302 and was defined on the basis of the duplication of an initial single exon. These two binding sites are of apparently different design; the one at residues 342–366 conforms to the amphipathic α-helix model of O’Neill and DeGrado (41), to which most calmodulin-binding proteins adhere, whereas the upstream site does not (2, 42). Furthermore, these two appear to interact with calmodulin in a quite distinct manner; the first causes the more usual compaction of the calmodulin structure, whereas the second does not (43). Functional and structural homologies between this region of γ with the two calmodulin binding sites has been proposed for a region of troponin I, which is also derived from a single exon (44). These data provide the first insights into the genomic organization of the γ subunit of phospholipase C and will now provide the base upon which to examine its transcriptional regulation and to gain greater insight into the structure/function relationships of its component parts of both the protein and the gene.

REFERENCES