The GRK4 Subfamily of G Protein-coupled Receptor Kinases

ALTERNATIVE SPlicing, GENE ORGANIZATION, AND SEQUENCE CONSERVATION*

(Received for publication, March 19, 1999, and in revised form, June 21, 1999)

Richard T. Premont†, Alexander D. Macrae‡§, Samuel A. J. R. Aparicio¶, Humphrey E. Kendall‡, Jeffrey E. Welch†, and Robert J. Lefkowitz‡**

From the †Departments of Medicine (Cardiology) and Biochemistry, Howard Hughes Medical Institute, Duke University Medical Center, Durham, North Carolina 27710, §Wellcome/CRC Institute, Tennis Court Road, Cambridge CB2 1QR, United Kingdom, and the ¶Regulatory Toxicology Division, National Health and Environmental Effects Research Laboratory, MD-72, United States Environmental Protection Agency, Research Triangle Park, North Carolina 27711

G protein-coupled receptor kinases (GRKs) desensitize G protein-coupled receptors by phosphorylating activated receptors. The six known GRKs have been classified into three subfamilies based on sequence and functional similarities. Examination of the mouse GRK4 subfamily (GRKs 4, 5, and 6) suggests that mouse GRK4 is not alternatively spliced in a manner analogous to human or rat GRK4, whereas GRK6 undergoes extensive alternative splicing to generate three variants with distinct carboxyl termini. Characterization of the mouse GRK 5 and 6 genes reveals that all members of the GRK4 subfamily share an identical gene structure, in which 15 introns interrupt the coding sequence at equivalent positions in all three genes. Surprisingly, none of the three GRK subgroups (GRK1, GRK2/3, and GRK4/5/6) shares even a single intron in common, indicating that these three subfamilies are distinct gene lineages that have been maintained since their divergence over 1 billion years ago. Comparison of the amino acid sequences of GRKs from various mammalian species indicates that GRK2, GRK5, and GRK6 exhibit a remarkably high degree of sequence conservation, whereas GRK1 and particularly GRK4 have accumulated amino acid changes at extremely rapid rates over the past 100 million years. The divergence of individual GRKs at vastly different rates reveals that strikingly different evolutionary pressures apply to the function of the individual GRKs.

The G protein-coupled receptor kinases (GRKs) are a family of six serine/threonine protein kinases characterized by their ability to phosphorylate and desensitize agonist-occupied cell surface G protein-coupled receptor proteins (1, 2). There is increasing functional evidence that these kinases can be segregated into three distinct subfamilies: GRK1 (rhodopsin kinase), GRK2-like (GRKs 2 and 3), and GRK4-like (GRKs 4, 5, and 6).

Although the GRK4-like kinases were initially identified through their similarity to GRK2, close examination reveals that the highest sequence similarity is clustered within the central protein kinase catalytic domain. There is considerable sequence divergence among the six GRKs in the regulatory amino- and carboxyl-terminal domains, which correlates with differences in regulation among the three kinase subfamilies. Thus GRK2-like kinases are translocated to the plasma membrane following receptor activation by interaction of the carboxyl-terminal pleckstrin homology domain with G protein βγ-subunits and membrane phosphatidyl inositol 4,5-bisphosphate (3, 4). In contrast, GRK4-like kinases exhibit high constitutive membrane association (5, 6). Although both GRK2-like and GRK4-like kinases are regulated by membrane lipids, the site and mechanism of action differ between the two groups (7). Recently, GRK2 and GRK5 were shown to be regulated differentially by protein kinase C phosphorylation and by Ca2+/calmodulin (2). Nevertheless, both GRK2-like and GRK4-like GRKs are capable of phosphorylating agonist-occupied G protein-coupled receptors. In the case of rhodopsin and the β2-adrenergic receptor, many individual sites of GRK phosphorylation by GRK2 and GRK5 are the same (8, 9), indicating that these distinct kinases can exert similar effects on the receptor. However, it is also evident that GRK2-like and GRK4-like GRKs can have distinct functional effects on other G protein-coupled receptors (2).

We have cloned the cDNAs encoding mouse GRKs 4, 5, and 6, and the genes for mouse GRK5 and 6. In this paper, we show that GRK4 alternative splicing is not conserved across species, whereas GRK6 alternative splicing is conserved in several species. Further, the conservation in gene organization within the GRK4 subfamily strongly supports the functional classification of these GRKs into a distinct subgroup. However, the amino acid sequences of the three individual GRKs in this subgroup appear to be evolving at vastly different rates.

* This work was supported by National Institutes of Health Grant HL16037 (to R. J. L.). 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®/EBI Data Bank with accession number(s) AF040745 (mouse GRK4 cDNA), AF040746 (mouse GRK5 cDNA), AF040747, AF040748, and AF040749 (mouse GRK6A, B, and C cDNAs), AF040750 (rat GRK6C cDNA), AF040751 and AF040752 (human GRK5 and C cDNAs), AF040753 (human GRK6 intron O), AF040755, AF040756, AF040757, AF040758, and AF040759 (mouse GRK5 gene), and AF040754 (mouse GRK6 gene).

§ Supported in part by a Wellcome Trust Advanced Training Fellowship from the Department of Medicine, Royal Postgraduate Medical School, Du Cane Road, London, UK. Present address: Neuroscience Therapeutics Unit, SmithKline Beecham Pharmaceuticals, 3rd Ave., Harlow, Essex CM19 5AW, UK.

¶ Investigator of the Howard Hughes Medical Institute. To whom correspondence should be addressed: Dept. of Medicine, Howard Hughes Medical Inst., Box 3821, Duke University Medical Center, Durham, NC 27710. Tel.: 919-684-2974; Fax: 919-684-8875; E-mail: lefkow001@mc.duke.edu.

** The abbreviations used are: GRK, G protein-coupled receptor kinase; EST, expressed sequence tag; G protein, heterotrimeric G protein guanine nucleotide-binding regulatory protein; bp, base pair(s); kb, kilobase pair(s).
EXPERIMENTAL PROCEDURES

Materials—Restriction enzymes and thermostable polymerases were from Promega. TA cloning kit was obtained from Invitrogen. Random-prime label reagents were from New England Biolabs. Nucleotides were from Roche Molecular Biochemicals, and radioisotopes from NEN Life Science Products. General laboratory chemicals were from Sigma.

Genomic Library Screening—A mouse 120/SVJ genomic library in AFXII obtained from Stratagene was hybridized with full-length human GRK4, bovine GRK5 and human GRK6 cDNA probes under standard conditions (10). Phage were sorted into three classes based on relative signal intensity with the three full-length GRK cDNA probes and were further sorted by hybridization to amino-terminal and carboxyl-terminal region probes. XbaI-digested phages from each class were subcloned into pBSII for sequencing. Fragments were ordered using restriction mapping and Southern blotting, DNA sequencing, or polymerase chain reaction amplification of the phage DNA from phages derived from partial sequences of subcloned fragments, in which case XbaI junctions were confirmed by direct sequencing of the resulting DNA bands. In addition, some intron regions were amplified from phage templates using flanking exon primers, and the resulting DNA bands were subjected to DNA sequencing directly. Amplification reactions were performed for 15 cycles of 95 °C for 1 min, 60 °C for 1 min, and 72 °C for 10 min using reagents as described below. Internal regions were sequenced from specific primers, and all regions were sequenced on both strands of DNA. Sequencing was performed using dye terminator cycle sequencing using ABI Prism AmpliTaq DNA polymerase. Sequencing reactions were visualized with ABI 373 or ABI 377 instruments. Raw sequence data were edited using EditView and AutoAssembler software (Perkin-Elmer) and analyzed using GeneWorks (Intelligenetics).

Cloning of Mouse GRK cDNAs—Mouse GRK cDNA sequences were obtained by amplification of mouse tissue first strand cDNA using degenerate oligonucleotide primers (10). Oligo(dT)-primed first strand cDNA was synthesized from mouse brain and testis poly(A) RNA (CLONTECH) as described previously (11). Amplification reactions were performed for 15 cycles of 95 °C for 1 min, 60 °C for 1 min, and 72 °C for 10 min using reagents as described below. Internal regions were sequenced from specific primers, and all regions were sequenced on both strands of DNA. Sequencing was performed using dye terminator cycle sequencing using ABI Prism AmpliTaq DNA polymerase. Sequencing reactions were visualized with ABI 373 or ABI 377 instruments. Raw sequence data were edited using EditView and AutoAssembler software (Perkin-Elmer) and analyzed using GeneWorks (Intelligenetics).

Mouse Genomic Localization—Genomic localization was determined by single strand conformational polymorphism analysis (13) using genomic DNA from the European Collaborative Interspecific Mouse Backcross panel (14). Primer pairs were chosen to amplify intronic regions containing potentially polymorphic sequences. The primer pair for GRK4 (5'-GGTCTTCTATGGCTACGGACG-3', 5'-TGCCACAACTCTGACCTAAAGGATG-3') amplified a 410-bp region within intron F. The primer pair for mouse GRK5 (5'-ACCTATGCTCTCCTTACTTTCTC-3', 5'-TTACAACCTTCTCCATCCACGATTG-3') amplified a 174-bp region of intron D. The primer pair for mouse GRK6 (5'-CATCCCCGGTACCTGGGAG-3', 5'-CTCAACGGCCACCCACCTGTAC-3') amplified a 350-bp region including exon IX and intron K which contains a GT repeat. 10 ng of individual Mus musculus (C57Bl/6) x Mus spretus backcross mouse genomic DNA (14) was amplified in a 20-μl reaction containing buffer and nucleotides as above, but with 1 μl of each primer and 2 μl of [α-32P]dCTP. Polymerase chain reaction products were denatured and separated on 6% nondenaturing polyacrylamide gels containing 10% glycerol. The gel was dried and exposed to x-ray film, and polymorphic bands were scored for species origin. Polymorphisms were analyzed and locus order was determined using the Map Manager program (15). Absolute map positions on mouse chromosomes as well as nearby mouse genes and syntenic human chromosome regions were identified using the chromosome maps of the mouse genome data base through the Encyclopedia of the Mouse Genome at the Jackson Labs.

Phylogenetic Analysis—Protein sequences corresponding to all known mammalian GRKs, and to those from Drosophila melanogaster and Caenorhabditis elegans, were obtained from the GenBank™ and SwissProt data bases using BLAST (12) and aligned using ClustalX. The resulting alignment was used as the framework to align the cDNA sequences encoding these protein sequences. These two alignments were then used for subsequent inferences of phylogenetic relationships. For proteins, trees were calculated using the NJ method (with bootstrap resampling of 1000 subreplicates) and, independently, using maximum parsimony (PROTPARS) in the PHYLIP (16) suite of programs. For DNA sequences, the alignment was used directly in the NJ method, a maximum likelihood method (DNAML) and was also used to calculate a distance matrix (DNADIST), which was used in the Fitch-Margoliash least squares methods (with and without clock assumptions). All of the trees produced were rooted. Topology trees were drawn with DRAWGRAM. Pairwise identities between proteins or nucleic acid sequences were computed during multiple alignment.

RESULTS

Mouse GRK4, 5, and 6 cDNAs—To examine the GRK4 subfamily in a single species, the mouse GRK4, GRK5, and GRK6 cDNA and gene sequences were determined. The deduced amino acid sequences of mouse GRK4, GRK5, and GRK6 are shown in Fig. 1A.

Mouse GRK4 is 574 amino acids in length, with a predicted mass of 66.8 kDa. This mouse GRK sequence is only 77% identical to human GRK4, a degree of homology low enough to raise the possibility that this might represent a novel GRK4-like sequence. However, gene mapping and Northern blotting (see below) indicate that this mouse gene does represent the mouse GRK4 homolog. Mouse GRK4 retains a single carboxy-terminal domain cysteine residue (Cys660) as a potential site for palmitoylation, in contrast to the two potential sites in human GRK4 (6). Human GRK4 mRNA and protein has been shown to exist in four forms, which arise from the alter-
native splicing of exons II and XV (6), whereas rat GRK4 has been reported to undergo distinct alternative splicing of exons VI, VII, and XIV (17). Amplification of the amino-terminal, central, and carboxyl-terminal domains of the GRK4 cDNA from mouse testis cDNA and isolated mouse germ cell cDNA yielded only single product bands (data not shown), evidence that alternative splicing of the mouse GRK4 mRNA is not prominent. Mouse GRK4 appears to exist as only a single form, equivalent to the longest (α-variant) of human and rat GRK4 (6, 17).

Mouse GRK5 is 590 amino acids in length, with a predicted size of 67.6 kDa. This mouse GRK5 sequence is 99% identical to rat and 95% identical to human and bovine GRK5. Mouse GRK5 retains the major autophosphorylation sites (Ser503-Thr504) localized in bovine and human GRK5 (5, 17), as well as the presumed protein kinase C sites (Ser566 and Ser572) found in human GRK5 (19). Mouse GRK5 contains an amino-terminal phosphatidylinositol binding polybasic region (K22RRKK) (7), as well as a carboxyl-terminal region rich in basic and serine residues (5).

GRK6 has recently been reported to exist as two splice variants in the rat, called GRK6A and GRK6B (20). Rat GRK6A and B differ in the sequence of the extreme carboxyl-terminal region, where the addition of 2 bp in the type B mRNA leads to a shift in the reading frame. Amplification of mouse GRK6 cDNA sequences using both an extreme carboxyl-terminal primer and a 3′-untranslated region primer led to the identification of these two variants of the mouse GRK6 cDNA as well (Fig. 1B). In addition, searches of the GenBank EST database confirmed the presence of mouse EST sequences encoding GRK6A and GRK6B variants, as well as a third GRK6 variant we call GRK6C (Fig. 1B). Mouse GRK6C appears to arise from the use of alternative splicing of exons II and XV (6), whereas rat GRK4 has been reported to undergo distinct alternative splicing of exons VI, VII, and XIV (17). Amplification of the amino-terminal, central, and carboxyl-terminal domains of the GRK4 cDNA from mouse testis cDNA and isolated mouse germ cell cDNA yielded only single product bands (data not shown), evidence that alternative splicing of the mouse GRK4 mRNA is not prominent. Mouse GRK4 appears to exist as only a single form, equivalent to the longest (α-variant) of human and rat GRK4 (6, 17).

Mouse GRK5 is 590 amino acids in length, with a predicted size of 67.6 kDa. This mouse GRK5 sequence is 99% identical to rat and 95% identical to human and bovine GRK5. Mouse GRK5 retains the major autophosphorylation sites (Ser503-Thr504) localized in bovine and human GRK5 (5, 18), as well as the presumed protein kinase C sites (Ser566 and Ser572) found in human GRK5 (19). Mouse GRK5 contains an amino-terminal phosphatidylinositol binding polybasic region (K22RRKK) (7), as well as a carboxyl-terminal region rich in basic and serine residues (5).

GRK6 has recently been reported to exist as two splice variants in the rat, called GRK6A and GRK6B (20). Rat GRK6A and B differ in the sequence of the extreme carboxyl-terminal region, where the addition of 2 bp in the type B mRNA leads to a shift in the reading frame. Amplification of mouse GRK6 cDNA sequences using both an extreme carboxyl-terminal primer and a 3′-untranslated region primer led to the identification of these two variants of the mouse GRK6 cDNA as well (Fig. 1B). In addition, searches of the GenBank EST database confirmed the presence of mouse EST sequences encoding GRK6A and GRK6B variants, as well as a third GRK6 variant we call GRK6C (Fig. 1B). Mouse GRK6C appears to arise from the use of alternative splicing of exons II and XV (6), whereas rat GRK4 has been reported to undergo distinct alternative splicing of exons VI, VII, and XIV (17). Amplification of the amino-terminal, central, and carboxyl-terminal domains of the GRK4 cDNA from mouse testis cDNA and isolated mouse germ cell cDNA yielded only single product bands (data not shown), evidence that alternative splicing of the mouse GRK4 mRNA is not prominent. Mouse GRK4 appears to exist as only a single form, equivalent to the longest (α-variant) of human and rat GRK4 (6, 17).

Mouse GRK5 is 590 amino acids in length, with a predicted size of 67.6 kDa. This mouse GRK5 sequence is 99% identical to rat and 95% identical to human and bovine GRK5. Mouse GRK5 retains the major autophosphorylation sites (Ser503-Thr504) localized in bovine and human GRK5 (5, 18), as well as the presumed protein kinase C sites (Ser566 and Ser572) found in human GRK5 (19). Mouse GRK5 contains an amino-terminal phosphatidylinositol binding polybasic region (K22RRKK) (7), as well as a carboxyl-terminal region rich in basic and serine residues (5).
of a novel last exon (see below) that encodes only a single amino acid, arginine, before the termination codon. Mouse GRK6A has 576 residues and a predicted size of 65.9 kDa, mouse GRK6B has 589 residues and a size of 67.0 kDa, and mouse GRK6C has 560 residues and a calculated size of 64.2 kDa. The mouse GRK6 sequences are 99% identical to rat and 97% identical to human GRK6. All variants of mouse GRK6 retain the Thr^{504} residue in the autophosphorylation region, although this site appears not to autophosphorylate in human GRK6 (21). The sequence of the mouse GRK6A carboxyl-terminal region (17 residues) is identical to those of rat and human and contains the three carboxyl-terminal cysteine residues that are the sites of palmitoylation (22). The sequence of the GRK6B variant carboxyl-terminal region (30 residues) has only one conservative difference with that of rat (Arg^{770} to Lys) and has five differences with the equivalent region of human GRK6B (Fig. 1B). Mouse GRK6B contains a consensus protein kinase A phosphorylation site at Ser^{573}, which is also conserved in rat and human sequences. Like mouse GRK6C, rat and human GRK6C also end with a single arginine residue (Fig. 1B). The type C variant has a truncated carboxyl-terminal region compared with GRK6 A or B and lacks both the sites of palmitoylation in GRK6A and the consensus cAMP-dependent protein kinase site found in GRK6B.

Tissue Distribution of GRK mRNAs in the Mouse—Because of the quite low similarity of the mouse GRK4 sequence with human GRK4, we were unsure whether this represented a novel GRK4-like kinase. Therefore, we documented the pattern of the expression of this mRNA in adult tissues and during embryonic development in the mouse, along with those of the more highly conserved GRK5 and GRK6 homologs (Fig. 2A). Hybridization of mouse Northern blots with probes specific for each of GRKs 4–6 shows the distinct patterns of expression for each previously described in other species (1, 2). GRK6 is widely expressed at high levels, GRK5 is highly expressed in the heart and lung with widespread low expression, and GRK4 expression is quite high but confined to the testis, as expected from the distribution of human GRK4 mRNA (6, 23, 24). Further, probing the mouse tissue Northern blot with a human GRK4 probe revealed hybridization only to a testis mRNA species of the same size (data not shown). In the developing embryo, GRK6 mRNA is highly expressed at all stages of development. GRK4 mRNA expression is not detected in the developing embryo. Finally, although it appears that GRK5 mRNA is most prevalent in the day 7 embryo, these samples contain placenta, a tissue that has previously been shown to express high levels of GRK5 (25).

Because the GRK4 mRNA is readily apparent in testis RNA from adults but is not detected in whole embryo RNA, RNAs isolated from mouse testes at various ages after birth were blotted and hybridized to mouse GRK4 probe (Fig. 2B). No GRK4 hybridization is apparent until after post-partum day 16. GRK4 mRNA is expressed at high levels from day 18 after birth, and expression continues through adulthood. In contrast, GRK2 mRNA is present at a low level in the testis at all times examined (data not shown). In the development of the testis in mouse, day 18 marks the first appearance of secondary spermatocytes and round spermatids (26). From the timing of the onset of GRK4 mRNA accumulation, it appears that GRK4 gene transcription is initiated in the developing male germ cells during the late pachytene stage in spermatocytes, just prior to the first meiotic division (26, 27). The GRK4 protein has been reported to be present in mature sperm cells, and the GRK4 mRNA has been observed in a spermatogonial cell line by Northern blotting and in spermatocytes and spermatids by RNA in situ hybridization (17, 23).

The tissue distribution of mRNA encoding the GRK6C variant was examined using a probe specific for this form (Fig. 2C). Specific hybridization to a 3.8-kb mRNA band was observed in all tissues, but at a quite low level that required a much longer exposure compared with the “common” GRK6 probe (compare Fig. 2A). Interestingly, the levels of GRK6C mRNA appear relatively high in the day 11 and day 15 mouse embryo, suggesting that this variant may be developmentally regulated. Because of the overlapping nature of exons 16A and 16B (Ref. 20 and see below), mRNAs encoding the GRK6A and B variants cannot be distinguished by Northern blotting.

GRK Gene Localization in Mouse—To further confirm that the mouse GRK4 sequence represents the homolog of human GRK4, this gene was mapped to 20 centimorgans from the centromere of chromosome 5, near the D5Mit75 marker (Fig.
This is consistent with the known location of the GRK4 gene, because GRK4 was identified by positional cloning at human chromosome band 4p16.3 in the search for the Huntington's disease locus, and the GRK4 gene lies immediately adjacent to the Huntington's disease gene in both the human and mouse genomes (24, 28, 29).

The mouse GRK5 gene maps to 52 centimorgans from the centromere of chromosome 19, in a position indistinguishable from the d19MIT1 marker (Fig. 3). This is within a region of synteny with human chromosome 10, band q25, and in agreement with the localization of the GRK5 gene to human chromosome interval 10q24-qter (30). The GRK5 locus is near the genes for the mouse α2A- (50 centimorgans) and β1-adrenergic receptors (51 centimorgans) and the vesicular monoamine transporter 2 (53 centimorgans).

Mouse GRK6 gene maps to 45 centimorgans from the centromere of chromosome 13, between the d13MIT21 and d13MIT47 markers (Fig. 3). Most genes within this area of the mouse genome map to human band 5q13 (adenylyl cyclase 2 at 41 centimorgans, sodium/hydrogen exchanger 3 at 43 centimorgans, dopamine transporter at 46 centimorgans, and rasA at 47 centimorgans). However, the sodium/hydrogen exchanger 2 gene also maps to 45 centimorgans in the mouse and to human band 5q35. Thus the GRK6 and NHE2 genes appear to be in a small region of synteny with human chromosome 5, band 5q35, that is within a larger region of synteny with human region 5q13. The GRK6 gene in human has been previously localized to human chromosome band 5q35 (30). A human GRK6-like pseudogene has also been localized to human chromosome band 5q35 (30, 31). Rodents appear to have no GRK6-like pseudogene (data not shown and Ref. 20).

**Comparative Organization of the GRK4/5/6 Genes**—The mouse GRK5 gene consists of 16 exons separated by 15 introns (Fig. 4A and Table I). All splice junctions conform to the GT-AG consensus. The lengths of the four largest introns (A, B, C, and D) have not been determined, because phage containing the flanking exons did not overlap; however, the GRK5 gene extends over more than 80 kb in the genome. For the 11 fully sequenced introns within the mouse GRK5 gene, only three are less than 1 kb in length.

The GRK6 gene consists of 16 exons separated by 15 introns, as summarized in Fig. 4B and Table II. Only two GRK6 introns are significantly greater than 1 kb (introns A and M), and none are less than 350 bp. Thus, the GRK6 gene is comparatively compact and extends over less than 20 kb of genomic DNA.

![Fig. 3. Chromosomal localization of mouse GRK4, GRK5, and GRK6 genes.](image)

![Fig. 4. Organization of the mouse GRK5 (A) and GRK6 (B) genes.](image)
Conserved Structure and Evolution of GRK4 Subfamily Genes

Table I

| Number | Inton/exon junctions | Exon size | Name | Intron size |
|--------|----------------------|-----------|------|-------------|
| 1      | GAAAGTGTTGAG...292 bp | A         | 3' UTR | >8 kb       |
| 2      | ttccatcttcgag        | B         | 5' UTR | >15 kb      |
| 3      | tacatcttccttgag      | C         | 3         | >3 kb       |
| 4      | ttttgttttttag         | D         | 4         | >12 kb      |
| 5      | attcttcctccag        | E         | 5         | 659 bp      |
| 6      | gtcacccctggctag      | F         | 6         | 3042 bp     |
| 7      | tgtatataattcag       | G         | 7         | 891 bp      |
| 8      | tccccctccacag        | H         | 8         | 3116 bp     |
| 9      | ctctttgtttctcag      | I         | 9         | 2533 bp     |
| 10     | gttttcatctctcag      | J         | 10        | 1498 bp     |
| 11     | tactgtcccccaacag     | K         | 11        | 2123 bp     |
| 12     | ttctttctgtcag        | L         | 12        | 2817 bp     |
| 13     | acactgcttcagag       | M         | 13        | 3701 bp     |
| 14     | ttctgcttgtag         | N         | 14        | 451 bp      |
| 15     | tttttcttctccag       | O         | 15        | 1255 bp     |
| 16     | catctcggttcctcag     | P         | 16        |            |

Alternatively spliced exons 16A, 16B, and 16C are all found in the proper region of the GRK6 locus (Fig. 4, B and C). Exon 16A uses a canonical 3′ splice site and leads to the addition of 17 amino acid residues prior to termination. Exon 16B uses a nonconventional 3′ splice site (underlined in Table II) that is shifted 2 bp upstream from the canonical 3′ splice site of exon 16A, leading to an alteration of the reading frame and the addition of a further 30 amino acids, as noted for rat GRK6B (20). All other splice junctions in the GRK6 gene conform to the GT-AG consensus. Exon 16C arises from the use of a canonical 3′ splice site adjoining exon 16C, 451 bp upstream from exons 16A/B (Fig. 4C). Analysis of the rat GRK6 gene intron O sequence amplified by Firsov and Elalouf (20) and of the human GRK6 gene intron O sequence amplified from human genomic DNA using primers in the flanking exons 15 and 16A/B indicates that a potential exon 16C encoding a single arginine residue before a stop codon resides within a 150-bp region of high sequence conservation in the cognate positions in the rat and human GRK6 genes (data not shown). The use of this putative exon 16C in spliced mouse and human GRK6 mRNAs was confirmed by amplification across the junction of exons 15 and 16C from several cDNA libraries, followed by direct sequencing of the product bands (data not shown).

Comparison of the mouse GRK6 and GRK5 genes reveals that their exon/intron organization is identical, with each intron located in the equivalent position in these two genes. Further, the GRK5 and GRK6 organization is identical to the organization of the human GRK4 gene (6) and the partially characterized mouse GRK4 gene (data not shown).

Amino acid sequence comparisons had previously indicated that GRKs 4, 5, and 6 comprise a distinct subfamily within the GRK family (1). This relationship clearly also extends to the genomic level.

Evolution of GRK Subfamilies—The high degree of dissimilarity of the mouse and human GRK4 sequences prompted a closer examination of GRK sequence conservation. GRK sequences from several mammalian species have now been deposited in the sequence data bases. Comparisons were based on protein sequence alignment and subsequent cDNA sequence alignment as described under “Experimental Procedures.” The global alignment of GRK protein sequences reveals overall higher conservation of the central kinase catalytic domains of these proteins, compared with the amino- and carboxyl-terminal regions (data not shown).

Pairwise comparison of all available mammalian GRK amino acid sequences indicates that there is a strikingly variable degree of sequence conservation for individual GRK subtypes. GRK2 is the most highly conserved member of the family, with 96% of amino acid residues identical among mammalian homologs. Only 19 out of 689 amino acids vary among rodent, bovine, and human GRK2 enzymes, and no more than 16 residues change between any two species. The GRK2 protein kinase catalytic domain is almost completely conserved. Similarly, GRK6 from mouse, rat, and human exhibits 96% identity, whereas GRK5 from mouse, rat, bovine, and human remains 94% identical. GRK3 is less well conserved, with 89% identical residues among rat, bovine, and human enzymes. Surprisingly, GRK1 and GRK4 appear to be very poorly conserved. Comparison of the rat, bovine, and human GRK1 sequences reveals that only 78% of amino acid residues are identical among these species. Similarly, GRK4 retains only 72% identity from mouse or rat to human. Mouse and rat GRK4 sequences retain a surprisingly low 90% identity.

To examine further the relationships between GRK proteins and their evolution, phylogenetic trees were constructed based on aligned protein and cDNA sequences. All of the independent methods gave topologically equivalent trees for protein or DNA sequence (with one minor exception; see below). A representative protein tree with branch lengths scaled proportionally to
Conserved Structure and Evolution of GRK4 Subfamily Genes

TABLE II
Exon (intron structure of the mouse GRK6 gene

| UTR, untranslated region; aa, amino acid(s). | Exon size | Intron size | Name | Size bp |
|---------------------------------------------|-----------|-------------|------|---------|
| Number | Intr/exon junctions | Exon | | |
| 1 | GGTCCGGGCCGGG..100 | bp | | A | 3750 |
| 2 | tgatgtccttcagcag | 18 aa | A | 3750 |
| 3 | ggttcggccggaaag | G | 3750 |
| 4 | cccctcaccctgtag | G | 3750 |
| 5 | tggctatcttcag | G | 3750 |
| 6 | aatctctctcaacag | G | 3750 |
| 7 | tactctgctctag | G | 3750 |
| 8 | tctgctctctccag | G | 3750 |
| 9 | gctctccgagcacag | G | 3750 |
| 10 | ttcctccttcctag | G | 3750 |
| 11 | cataccctctcag | G | 3750 |
| 12 | ctctctctctccag | G | 3750 |
| 13 | actgctctctcctag | G | 3750 |
| 14 | cctctctctctcctag | G | 3750 |
| 15 | cctctgctctccag | G | 3750 |
| 16 (A) | tccctctctctcag | G | 3750 |
| 16 (B) | ctctctctctctcag | G | 3750 |
| 16 (C) | ctctctctctctcag | G | 3750 |

characterization of the mouse GRK4 subfamily cDNAs and genes has revealed several interesting features. We determined that GRK4 alternative splicing is not conserved among mammalian species, whereas GRK6 alternative splicing is conserved in rodents and primates. We identified a novel GRK6 splice variant, GRK6C, which is expressed during embryonic development but is found only at low levels in adult tissues. The GRK4, 5, and 6 genes are located in the equivalent locations in the human and mouse genomes, indicating that these genes diverged and were distributed in the mammalian genome prior to the divergence of rodents and primates 100 million years ago. However, the GRK4, 5, and 6 genes share an identical exon organization that is totally distinct from the genes of other GRKs, further supporting their assignment as a distinct GRK subfamily with a common ancestry. Finally, comparison of individual GRKs from several species revealed that some GRKs are very highly conserved (GRK2 and GRK6), whereas others are very poorly conserved (GRK4 and GRK1).

**GRK Alternative Splicing**—Alternative splicing generates four forms of the human GRK4 mRNA that differ in the presence or absence of exon II in the amino-terminal region and exon X in the carboxyl-terminal region (6, 23). Rat GRK4 was recently reported to undergo alternative splicing as well, but of exons VI, VII, and XIV (24). In examining the mouse GRK4 cDNA, we have found no evidence for either pattern of alternative splicing. Thus GRK4 alternative splicing may not be a
common feature among mammalian species.

In contrast, GRK6 appears to undergo extensive alternative splicing of its extreme carboxyl-terminal to yield three distinct variants, GRK6A, B, and C, and these variants can be found in mouse, rat, and human. Mouse GRK6A and B differ by the retention of 2 bp during the splicing of intron O, presumably because of the use of a secondary 3’ splice site, as noted for rat GRK6 (20). This 2-bp difference leads to a shifted reading frame for the last 17 or 30 amino acids, respectively. The GRK6C variant arises from the use of a distinct exon (exon 16C) that encodes a single amino acid residue. This alternative splicing should have significant functional consequences for GRK6, because these final residues are known to be important for the membrane localization and activity of the enzyme. The palmitoylation of GRK6 within the GRK6A-specific carboxyl-terminal region is associated with membrane localization of the enzyme, and palmitate-modified GRK6A is substantially more active than GRK6A protein that has no lipid modification (22, 32, 33). Therefore, the GRK6B and GRK6C variants, which are predicted to lack palmitoylation, would require some other mechanism for membrane localization. The GRK6B variant carboxyl-terminal sequence resembles and aligns with the cognate domain of GRK5, which is basic and rich in proline and hydroxyl groups. This region of GRK5 is thought to be important for constitutive association of GRK5 with the membrane (5). Interestingly, this region of GRK6B contains a consensus cAMP-dependent protein kinase site that might allow phosphorylation-regulated association of the GRK6B protein with the membrane. By contrast, GRK6C has a truncated carboxyl-terminal domain lacking both palmitoylation sites and GRK5-like membrane localization domains. This form of GRK6 might be expected to associate with membranes poorly and be a poor regulator of G protein-coupled receptors. These multiple GRK6 variants may underlie the heterogeneity of GRK6 in tissue immunoblots (34).

**Organization of GRK Genes—** Direct comparison of the human GRK4 gene with the mouse GRK5 and GRK6 genes reveals that all three genes share an identical gene organization, in which all 15 introns are found in the equivalent positions within the coding regions. However, the length of the cognate introns varies considerably among these three genes. This conservation of gene structure within the GRK4 family supports the previous classification of these three sequences as a distinct subfamily of GRKs based on primary sequence similarity (1, 2). The C. elegans genomic sequencing project has identified a single GRK4-like gene, F19C6.1 (35), which has 13 exons. This gene shares six exon/intron boundaries (mammalian introns A, E, H, K, N, and O) with the mammalian GRK4-like genes, whereas four other introns are located within 6 bp of the mammalian position (data not shown).

GRK2 and GRK3 also form a subfamily of GRKs based on amino acid sequence and functional similarities (1). The GRK2 gene has 21 exons separated by 20 introns (36, 37). The portion of the mouse GRK3 gene that has been cloned and characterized appears identical in organization to the human and mouse GRK2 genes, although the introns are much larger (38). A single C. elegans sequence for a GRK2-like gene, W02B3.2, has eight introns and shares four exon/intron boundaries (introns C, D, H, and L) with the mammalian GRK2-like genes (35). Three other introns are shifted (4, 15, and 9 bp) relative to human introns P, R, and S. The final intron is within the equivalent of the last exon of human GRK2 (data not shown).

No single exon/intron boundary is shared between the GRK2-like and the GRK4-like GRK gene families. This is true in mammals, in C. elegans, and even when comparing mammalian and C. elegans GRK genes. The complete dissimilarity of the exon/intron structure of both the mammalian and the C. elegans GRK2-like and GRK4-like genes indicates that these two gene families have not arisen recently from a common ancestor but that the common GRK ancestor gene diverged prior to the divergence of protostomes and deuterostomes 1 billion years ago (39). The complete conservation of gene structure within each GRK subfamily in mammals provides evidence that the members of each subfamily arose more recently by gene duplication.

GRK1 forms a third distinct subfamily among the mammalian GRKs (1). The human GRK1 gene has seven exons separated by six introns (40). The GRK1 gene shares no exon/intron boundaries with either the GRK2-like or GRK4-like genes. The complete dissimilarity of the GRK1 gene structure with those of the GRK2-like and GRK4-like subfamilies argues that GRK1 is also an ancient gene lineage and is not recently derived from a GRK2 or GRK4 subfamily member. It will be of interest to examine the rate of divergence of this subfamily in lower vertebrates and to search for its existence in sister groups to the vertebrates, such as amphioxus.

**Conservation of GRK Sequences—** The relationships among family members and their relative rates of divergence are illustrated by the phylogenetic reconstruction of relationships. This analysis emphasizes and confirms these differences and supports the existence of three ancient groups of GRK se-
quences. Individual GRKs clearly are subject to a variable degree of pressure for sequence conservation, because there are GRKs that have an unusually slow rate of amino acid substitution (GRKs 2, 5, and 6), and others (GRK1 and GRK4) with a 10-fold higher apparent rate of amino acid substitution. Given the variable rate of evolution for these genes, what relationships can be inferred among the GRK family groupings and their purported functions that might account for these selective pressures? Targeted deletion of the GRK2 gene leads to embryonic lethality because of malformation of the heart during development (36). GRK2 is expressed ubiquitously throughout the body and appears to play a role in the regulation of many distinct receptor proteins in diverse cell types (2). These widespread roles may in part underlie the strong evolutionary pressure to conserve the GRK2 sequence. GRK3 appears to play a subsidiary role to the highly similar GRK2 protein, with few tissues expressing more GRK3 than GRK2 (2). It is therefore not surprising that targeted deletion of the GRK3 gene does not alter viability (38). Although their global functions remain unknown, the relatively high conservation of GRK5 and GRK6 argues that these may also be fundamental to conserved processes in several systems.

In marked contrast, GRK1 and GRK4 are known to have very limited tissue distributions. For GRK1, restriction to the rod and cone outer segments (41) limits the functional role of the enzyme to that of a “rhodopsin kinase,” because other potential substrates are unavailable. GRK4 expression is primarily in cells of the male germ line (6, 17, 23), and expression does not begin until after day 16 post-partum. The natural substrates of GRK4 remain unknown, but the limited expression of GRK4 suggests that there can be very few of them. It is possible that these limited functional roles for GRK1 and GRK4 allow greater flexibility for the enzymes, which in turn permits a very high rate of sequence substitution. Once GRK sequences from additional nonmammalian species are available, a more comprehensive analysis of the relationship of expression pattern, function, and evolutionary history and divergence of the GRKs will be possible.

Acknowledgments—We thank Millie McAdams and Judy Phelps of the HHMI/DUMC Biopolymer Facility for ABI 377 DNA sequencing and Drs. Helena Abushamaa and Andy Peterson of the Genome Center at the DUMC Comprehensive Cancer Center for the single-strand conservation of GRK4 will be possible.

GRK4 allow greater flexibility for the enzymes, which in turn possible that these limited functional roles for GRK1 and suggest that there can be very few of them. It is does not begin until after day 16 post-partum. The natural substrates of GRK4 remain unknown, but the limited expression of GRK4 suggests that there can be very few of them. It is possible that these limited functional roles for GRK1 and GRK4 allow greater flexibility for the enzymes, which in turn permits a very high rate of sequence substitution. Once GRK sequences from additional nonmammalian species are available, a more comprehensive analysis of the relationship of expression pattern, function, and evolutionary history and divergence of the GRKs will be possible.

Acknowledgments—We thank Millie McAdams and Judy Phelps of the HHMI/DUMC Biopolymer Facility for ABI 377 DNA sequencing and Drs. Helena Abushamaa and Andy Peterson of the Genome Center at the DUMC Comprehensive Cancer Center for the single-strand conservation of GRK4 will be possible.

GRK4 allow greater flexibility for the enzymes, which in turn possible that these limited functional roles for GRK1 and suggest that there can be very few of them. It is does not begin until after day 16 post-partum. The natural substrates of GRK4 remain unknown, but the limited expression of GRK4 suggests that there can be very few of them. It is possible that these limited functional roles for GRK1 and GRK4 allow greater flexibility for the enzymes, which in turn permits a very high rate of sequence substitution. Once GRK sequences from additional nonmammalian species are available, a more comprehensive analysis of the relationship of expression pattern, function, and evolutionary history and divergence of the GRKs will be possible.

Acknowledgments—We thank Millie McAdams and Judy Phelps of the HHMI/DUMC Biopolymer Facility for ABI 377 DNA sequencing and Drs. Helena Abushamaa and Andy Peterson of the Genome Center at the DUMC Comprehensive Cancer Center for the single-strand conservation of GRK4 will be possible.

GRK4 allow greater flexibility for the enzymes, which in turn possible that these limited functional roles for GRK1 and suggest that there can be very few of them. It is does not begin until after day 16 post-partum. The natural substrates of GRK4 remain unknown, but the limited expression of GRK4 suggests that there can be very few of them. It is possible that these limited functional roles for GRK1 and GRK4 allow greater flexibility for the enzymes, which in turn permits a very high rate of sequence substitution. Once GRK sequences from additional nonmammalian species are available, a more comprehensive analysis of the relationship of expression pattern, function, and evolutionary history and divergence of the GRKs will be possible.

Acknowledgments—We thank Millie McAdams and Judy Phelps of the HHMI/DUMC Biopolymer Facility for ABI 377 DNA sequencing and Drs. Helena Abushamaa and Andy Peterson of the Genome Center at the DUMC Comprehensive Cancer Center for the single-strand conservation of GRK4 will be possible.