

Characterization of the Gene Encoding the Human Type II cGMP-Dependent Protein Kinase (PRKG2)

Oliwia Witczak,^{*,1} Sigurd Ørstavik,^{*} Vasanti Natarajan,^{*} Eirik Frengen,[†] Tore Jahnsen,^{*} and Mårten Sandberg^{*,‡}

^{*}Institute of Medical Biochemistry, University of Oslo, N-0317 Oslo, Norway; [†]The Biotechnology Centre of Oslo, University of Oslo, N-0316 Oslo, Norway; and [‡]Department of Anesthesiology, Central Hospital of Akershus, N-1474 Nordbyhagen, Norway

Received February 23, 1998

The type II cGMP-dependent protein kinase (cGK) plays a pivotal role in the regulation of intestinal fluid balance in man. Furthermore, mice carrying a null mutation for the gene encoding the type II cGK develop as dwarfs indicating that this enzyme has other less characterized roles. The present report describes the isolation and characterization of bacterial artificial chromosome (BAC)- and P1-derived artificial chromosome (PAC)-clones containing the gene encoding the human type II cGK. The gene was estimated to cover at least 125 kb and consisted of 19 exons separated by introns of various lengths. The splice junctions of the type II cGK gene corresponded well with the structure of the gene encoding human type I cGK and with the splice junctions observed in the *Drosophila melanogaster* DG2 gene. 5'-rapid amplification of cDNA-ends established the presence of a non-translated exon. © 1998 Academic Press

The intracellular level of the second messenger cyclic guanosine monophosphate (cGMP) is influenced by nitric oxide (NO) as well as a number of drugs and hormones (1,2). The increased level of cGMP results in activation of various receptors like cGMP-regulated ion channels, cyclic nucleotide dependent phosphodiesterases and cGMP-dependent protein kinases (cGKs) (3,4). The role of cGK is controversial, although the major effects of cGMP have been attributed to the activation of this enzyme (4-6). It has been suggested that some of the cGMP effects may be mediated by cross-activation of the cyclic adenosine monophosphate (cAMP) dependent protein kinases (cAKs) (6), a family of enzymes homologous to the

cGKs (7). The picture is further complicated by the existence of two major classes of the cGKs, type I (8-11) and type II (12-15), cGK, encoded by separate genes, in mammalian tissues. Type I cGK is a soluble protein and vascular smooth muscle cells and blood platelets contain high concentrations of this enzyme. Functionally, it has been shown that activation of the type I cGK leads to relaxation of vascular smooth muscle (3) as well as to inhibition of blood platelet activation and aggregation (16). In man, two differentially regulated isoforms of the type I cGK, type I α and type I β , have been described, but little is known about possible differences in their physiological roles (10,11,17). In contrast to the type I cGK, the type II cGK is membrane-bound (18). This enzyme is found in various human tissues, but the physiological role of the type II cGK is not fully understood, although it has been shown to be involved in the regulation of intestinal fluid secretion (19,20). Recently, mice deficient in the type II cGK were generated and it was shown that they were resistant to *Escherichia coli* STa, an enterotoxin that stimulates cGMP accumulation and intestinal fluid secretion (21). Interestingly, the mice also developed as dwarfs due to a defect in endochondral ossification at the growth plates.

The type I cGK exists as a dimer (22), and the elucidation of the primary sequence of the type I α cGK made it possible to assign specific functions to various domains of the protein (23). The amino terminal part of the molecule includes a dimerization domain followed by two cGMP binding sites, while the ATP-binding as well as the catalytic domains are located in the carboxyl end of the protein. Complementary DNA clones for the type II cGK have been isolated from several species (12-15) and also this enzyme can be divided into similar domains as the type I cGK. Although original studies indicated that the type II cGK existed as a monomer, a recent report has shown that this enzyme is also a dimer (24).

In order to characterize the gene encoding the type II cGK, we have determined its exon structure by ana-

¹ To whom correspondence should be addressed at Institute of Medical Biochemistry, PO Box 1112, Blindern, N-0317 Oslo, Norway. Fax: +47-22851497. E-mail: oliwia.witczak@basalmed.uio.no.

Sequence data from this article have been deposited in GenBank/EMBL Data Libraries under Accession Nos. Y16105-Y16123.

lyzing bacterial artificial chromosomes (BAC) and P1-derived artificial chromosomes (PAC) clones. The present report presents evidence that the gene in question is at least 125 kb in length and consists of 19 exons. We have earlier presented the exon structure of the type I cGK gene (11), and the exon structure of the two genes is very similar. However, the two genes differ in their 5'-ends and in contrast to what we have demonstrated for the gene encoding type I cGK, we have so far found no indications for the existence of several isoforms of the type II cGK.

MATERIALS AND METHODS

Screening of a human bacterial artificial chromosome (BAC) library. A BAC library containing human genomic DNA was screened by polymerase chain reaction (PCR) (Research Genetics Inc., Huntsville, AL). A sequence tagged site (STS) marker that corresponded to exon 2 of the type II cGK was designed. DNA pools were screened in 25- μ l reactions containing 10 pmol of primer 1 (5'-ATG-CTCTGCGGAACAAGG-TGACA-3', complementary to nucleotides 71 to 93 in the type II cGK) and 10 pmol of primer 2 (5'-CTCCTTCAA-ATGGTACTCCC-3', complementary to nucleotides 159 to 140 in the type II cGK), 200 μ M concentrations of each dNTP, 50 mM KCl, 10 mM Tris-HCl, pH 8.0, 1.5 mM MgCl₂, and 2 units of *Taq* DNA polymerase (Perkin Elmer, Foster City, CA). Thirty cycles were performed (94°C for 45 s; 66°C for 45 s; 72°C for 45 s) with a final 10-min extension period at 72°C. The resulting 89-bp reaction product was an indication of the presence of the type II cGK gene.

Screening of human P1-derived artificial chromosome (PAC) library. The vector pPAC4 (E. Frengen, B. Zhao, D. Weichenhan, E. Gjernes, J. Jessee, H. Prydz, and P. J. de Jong, unpublished results) has been used in the construction of a 4-fold redundant human library, RPC16. This library was prepared using previously described methods (25), and the PAC clones have been picked into 384-well microtiter dishes. High density colony membranes containing this library were screened using the type II cGK cDNA. The hybridizations were carried out according to Church and Gilbert (26). The clones RPC16-73B10, -79K18, -124A1, -132L8, and -203G17 were confirmed as positive by hybridization to colony filters containing two replicas of the clones from the library screening, and by using PCR. DNA from the PAC clone was isolated by a modified alkaline extraction protocol (27).

Preparation and analysis of cloned DNA. Bacteria containing the BAC or PAC clones were grown in liquid culture overnight and the DNA was isolated using ion-exchange columns as described by the manufacturer (Qiagen, Hilden, Germany). The resulting DNA was digested with *Eco*RI, *Hind*II or *Hind*III, separated electrophoretically and blotted onto nylon membranes. The resulting Southern blots were then successively hybridized using radioactively labeled oligonucleotides corresponding to the relevant exons of the type II cGK. Hybridizing fragments of suitable lengths were subcloned into a pUC vector and sequenced. The lengths of the BAC and PAC clones were determined by digesting DNA from the clones with *Nof*I, before the fragments were separated by pulsed field gel electrophoresis using Biometra Rotaphor R22 (Biometra Biomedizinische Analytik GmbH, Göttingen, Germany) and visualized by ethidium bromide staining.

DNA sequencing and analysis of sequences. The subcloned fragments were sequenced by the dideoxy chain termination method (28) using Sequenase DNA polymerase (Amersham, Buckinghamshire, UK) and a combination of insert- and vector-specific primers. Some of the DNA sequencing was performed by Eurogentec (Seraing, Belgium) using the original BAC or PAC clones as templates. Nucleotide and amino acid sequence data were analyzed using the GCG program

package (Program Manual for the Wisconsin Package, Version 8, September 1994, Genetics Computer Group, WI).

Determination of intron lengths. The isolated plasmids or human genomic DNA were used as templates in PCR reactions to determine the lengths of the introns of the gene. Oligonucleotides flanking each intron were synthesized and the introns amplified using the Expand long template PCR amplification mix as described by the manufacturer (Boehringer-Mannheim, Mannheim, Germany). The resulting PCR products were analyzed by agarose gel electrophoresis and visualized by ethidium bromide staining.

5' Rapid amplification of cDNA ends (5' RACE). The 5'-end of the type II cGK cDNA was amplified from human prostate Marathon RACE-ready cDNA (Clontech, Palo Alto, CA) using the Advantage KlenTaq Polymerase Mix as described by the manufacturer. The first amplification was performed using adapter primer 1 and primer 3 (5'-AATTGTATAAAATGGCAAGCTCCCCAA-3', complementary to nucleotides 692-718 of the human type II cGK). The primary reaction mixture was initially heated to 94°C for 30 s, followed by 5 cycles of 60 s at 94°C and 2 min at 72°C, 5 cycles of 45 s at 94°C and 2 min at 70°C, and finally 25 cycles of 45 s at 94°C and 2 min at 68°C. The resulting product was diluted (1:50) and used as template for a new amplification as above using adapter primer 2 and primer 4 (5'-AAAGATATGGTTTCCTGG-TTCTCCTTG-3', complementary to nucleotides 592-618 of the human type II cGK). The resulting PCR product was cloned into a pCR2.1 vector using the Invitrogen TA cloning kit as instructed by the manufacturer (Invitrogen, San Diego, CA).

Radiolabeling of DNA probes and oligonucleotides. Complementary DNA was labeled with (α -³²P)dCTP using a random priming protocol (29). Oligonucleotides were synthesized by National Biosciences (Plymouth, MN) or Eurogentec and radiolabeled with (γ -³²P)ATP using T4 polynucleotide kinase as described by the manufacturer (Bethesda Research Laboratories, Bethesda, MD).

RESULTS

Identification of the Exons Encoding the Type II cGK

Based on the cDNA sequence of the human type II cGK (14,15), a pair of PCR primers was designed and expected to give rise to an 89 bp PCR-fragment when human genomic DNA was used as the template. A human BAC library (30,31) was screened with this primer pair resulting in one positive clone (320I1), approximately 110 kb in length (Fig. 1). Characterization of this clone revealed the presence of the 10 most 5' exons of the type II cGK gene including 1253 nucleotides of the 2289-nucleotide coding region (Table 1).

In order to obtain genomic sequence information corresponding to the remaining part of the open reading frame of the cDNA (nucleotides 1254-2289), a human PAC library was screened using an *Eco*RI-fragment from the 3' end of the human type II cGK cDNA as a probe. This resulted in a total of 5 positive clones out of which a 120-kb clone (RPC16-132L8) was further characterized (Fig. 1). This clone was shown to contain exons 2-18, corresponding to nucleotides (-13) to 2193 in the cDNA sequence, where the start codon ATG is represented as nucleotides 1-3. However, no BAC or PAC clone containing the most 3' exon had been isolated. In order to obtain information about the 3' splice junction of exon 18 as well as the 5' splice junction

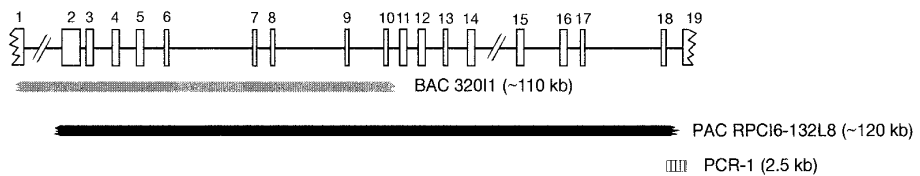


FIG. 1. Schematic representation of the human type II cGK gene. The open boxes depict the exons, and the lines between the boxes represent the introns. The scale used for the exons and the introns are different. However, the sizes of the boxes and the lengths of the lines indicate the relative sizes of the exons and introns, respectively. Clones used to determine the gene structure are indicated below. BAC: bacterial artificial chromosome, PAC: P1-derived artificial chromosome, PCR-1: PCR-generated clone (see Materials and Methods for details).

of exon 19, oligonucleotides from the two exons were constructed and a PCR reaction with total human genomic DNA as the template was performed. The resulting PCR product (Fig. 1) was characterized and gave sequence information about exon 19 and the remaining part of the open reading frame (Table 1).

The previously published human cDNA sequence extended only 15 nucleotides upstream of the start of translation (11). However, the murine type II cGK cDNA (12), contained 442 nucleotides of the 5' untranslated region. This led us to investigate if an additional 5' untranslated exon could exist in man. The 5' untranslated region of

TABLE 1
Exon/Intron Organization of the Human Type II cGK Gene

Exon no	Exon size (bp)	5' intron sequence	Nuc. no ¹	Exon sequence	Nuc. no ²	3' intron sequence	Intron phase	Intron size (kb)
1				CGGCTGTCTG	−14	gtaggttcgt		ND
2	474	gccttctcag	−13	GTCCCTGAGC	461	gtaagaaatt	2	0.6
3	167	attataccag	462	E K K TGAGAAGAAG	628	gtgggtttca	1	3.0
4	114	tgcttgatag	629	E G R L AGGGTCGACT	742	gtaataaaag	1	2.4
5	106	ttttgtaaag	743	A I T N CTATTACCAA	848	gtaagaataa	2	2.8
6	64	aattttgaag	849	V S S TGTATCCTTG	912	gtaagaaatt	0	12
7	78	atcccaacag	913	E Y Y GAATACTATG	990	gtaactataa	0	1.8
8	95	cattgttttag	991	V K V GTAAAAGTAA	1085	gtgaattcat	2	10
9	69	ttgtttgcag	1086	D D V TGATGATGTC	1154	gtgagtatac	2	4.5
10	99	atctattttag	1155	T F N AACATTCAAC	1253	gtaagtggag	2	1.3
11	154	tcttgtagacag	1254	R S M GCGGTCCATG	1407	gtaagggggtt	0	1.8
12	137	tcttacatag	1408	V K V GTAAAGTAA	1544	gtaagtgacc	2	2.8
13	90	gtcttcacag	1545	L Y R ATTATATCGT	1634	gtaatgaaaa	2	2.5
14	142	ttcttgcaag	1635	G S F AGGCAGCTTT	1776	gtaagacaac	0	ND
15	164	cgccgacaag	1777	V D F GTTGACTTTG	1940	gtatgtaccg	2	5.2
16	123	ctctatctag	1941	P P F CCCACCTTTT	2063	gtgagaatga	2	1.9
17	63	attgtttttag	2064	Q N P GCAAAATCCA	2126	gtacataatt	2	11
18	67	catttctcag	2127	W L N GTGGTTAAAT	2193	gtattgtatt	0	2.5
19		ggtccttgag	2194	L K G CTCAAGGGAC				

Note. Nuc. nos indicate the ¹first and ²last nucleotides of the exons, which are numbered according to the published type II cGK cDNA sequence (15). The nucleotides 1-3 represent the start codon ATG and the one-letter code for amino acids are indicated above the second base in each triplet. ND—not determined.

```

-1481  TTCTCACTCCTTCTCCTCTAATTTTTTTTTTTTGGAGGTCCTTATATAAATGTACAGACTCTGGGATAGCTTCATGAAAAGGTACTTTGTGTACTTTCGCG
-1379  ATTTTCCTCCAGTTAGCACAAAACAAACAGGACTATATTCCTCCTCCTCTGAAACTTAATGGTCTCATTCAACATATTTTACATGGGCCTCATTTTAAATA
-1277  TAAGTTTATGTATGTAACAGTGTATACAAATGTTACATGAGCCCCACTGCCACCATGTGGTTTAAAAACCAAGAGAGTTGGTTGTAAACATTCTCTAGATGC
-1175  ACAGTGAGGAATAGACCCCTCACTATACTTTCCATCTCTCAACACAGAAACTGCCCTTGCTCCTCGTGAATTTGCTGGTGAAATTCCTCTGGGGGCAGAAGGA
-1073  AAGGGAAAAAGCCACATGGAGATCTCACAACCTCAGAGCACAGAAATAATGCGGTGGGGGTGTACATATATATTAGGAAACCATATTGCACCCCTTCAGGAA
-971   AAAAAATGTACTCAAAGGTGACGAGAGAGCATTATCTTCTGGGTGATCAAGAAAGCTCCCTTCCCAAAGACCTGCAGATTGCGAAGGTATAATCTCTGT
-869   CCCTCGAAACTCGATTCAACCCCTGCCCCAGCATCAGTGAACCTGGCAAAGTGTAGTGAAGTTATGGCCCCAGAGTCCGGGAGGGGCGCGTCGACGAGCTG
-767   CTTGGGGGTGACATGTCTTCGTAAAAAGGCATTGGTGTGACAGAGAGGGCATGTCAGGGTACGTGCGCTGTTTGTAGCTCTGGGAGGGAAGGCGAAGTGG
-665   AGAGAGCCTGCAGTGAAAACAGACTGGAAGGATGGAGGAAAGGGAGGGGAAGGAAAAATGAACACCTCCCTGAACTTTCTCTCCCTCGTCCCACACACT
-563   TGAAATAGAAATAATAACATGAACAAAAATGCTTACAACAGGATGTTATCGCCGTACTTAAGTTGCTCAGTATGGAGGGTATGGGGTATAAATGCAGAC
-461   TGGCGCTGTGTGCGGAGTAAAGGTTTCGGGCTAGTGTGCGTGCAGGCGCGGGTGACCCACCTTTAGGGAAGCCGGCGCCCAACTATCCAGGCAGTAGCCCCG
-359   GCTGACCCCTCTCCTCCTCTCCCTCCTCTCTCCCTCCCTCCCTGCGTGTCTCTCTGCACCGGCCCGCGCGGAGCCAGGCGGCAGCACCCGACGCGG
-257   CCGCAGCTCCTGCCGCATCGCGCTAACCCGCGCCCCGGCTCACACGTAGCCGGAGCGACTGACCCGGCAGCCCCGGGCCCCAGCTGCGCACCGTGTGCAC
-155   GCGCGGCTGTGGCTGAGCTGACCTGCGCTGGGCTCCGGCGCCGAGGCGGCTGGGGAAGGGCGGCTGGGAGAGGACGCTGAGACGCGAGGGCGGCTCCGCGCGA
-53    GGTGACAAGCTGCGCACCTGGAAGTTACCCGGCTGTCTGgtgaggttcg.....gccttctcagTCCCTGAGCAAA
1      ATGGGAAATGGTTCAGTGAAACCTAAACATTCTAAGCACCCAGATGGACACTCTGGGAACCTCACCCTGATGCTCTGCGGAACAAGGTGACAGAGCTGGAG
MetGlyAsnGlySerValLysProLysHisSerLysHisProAspGlyHisSerGlyAsnLeuThrThrAspAlaLeuArgAsnLysValThrGluLeuGlu
103   AGAGAGTTGAGGAGGAAGGATGCTGAGATCCAGGAGCGGGAGTACCATTGGAAGGAGCTGCGGAGCAGCTGTGCAAGCAGACTGTGGCCATTGTGTAACCTC
ArgGluLeuArgArgLysAspAlaGluIleGlnGluArgGluTyrHisLeuLysGluLeuArgGluGlnLeuSerLysGlnThrValAlaIleAlaGluLeu
205   ACAGAGGAGCTCCAGAACAAGTGATCCAGCTGAACAAGCTGCAGGATGTGGTGCATATGCAGGGAGGAAGCCCGCTTCAGGCCTCTCCAGATAAAGTGCTT
ThrGluGluLeuGlnAsnLysCysIleGlnLeuAsnLysLeuGlnAspValValHisMetGlnGlyGlySerProLeuGlnAlaSerProAspLysValPro
307   CTTGAGGTCCACCGAAGACCTCTGGATTGGTCTCTCTCCATAGCAGAGGGGAGCAAAAGGCTGGCGTGTCTGCTGAGCCAACAACCCGGACCTATGACCTG
LeuGluValHisArgLysThrSerGlyLeuValSerLeuHisSerArgArgGlyAlaLysAlaGlyValSerAlaGluProThrThrArgThrTyrAspLeu
409   AACAAACCCCTGAATTTTCTTTGAGAAAGCAAGAGTCAGAAAAGACTCCAGGtaagaaatttcc
AsnLysProProGluPheSerPheGluLysAlaArgValArgLysAspSer

```

FIG. 2. Nucleotide sequence of the untranslated exon 1, exon 2 and the 5' upstream region of the human type II cGK. Nucleotide "A" in the translation initiation codon is numbered +1. Lower case letters represent intron sequences, while upper case letters represent the exons and the 5' untranslated sequences. Intron 1 is represented with dots. The translation of the exon is shown in the three-letter amino acid code. A potential CpG island is underlined, Sp-1 binding sites are boxed and potential TATA boxes are encircled. The start of the longest 5'-RACE clone is indicated with an arrowhead.

the murine type II cGK cDNA was compared to all sequences in GenBank using the Blast search program (32), and was shown to be similar to a human genomic sequence with accession no. z58461 (33). This sequence was derived from DNA isolated using a column binding methylated DNA intended to identify 5' ends of genes. In order to test the hypothesis that the human type II cGK gene contained an untranslated exon, an oligonucleotide corresponding to nucleotides (–410)–(–391) in the murine type II cGK cDNA was used as a probe. This oligonucleotide was also identical with nucleotides 87–106 of the genomic sequence z58461 and was shown to hybridize to BAC clone 320I1, demonstrating the existence of homologous sequences in the human type II cGK gene. This region of the clone was sequenced (Fig. 2) and

the human and murine sequences were shown to be 80 % identical in this region. In comparison, the identity between the coding regions of the type II cGK from the two species was 89 %.

Amplification of the 5'-Ends of the Type II cGK cDNA

To investigate whether exon 1 was expressed in human tissues, 5'-RACE was performed using cDNA derived from human prostate using an anchor primer and an oligonucleotide from exon 2 of the type II cGK. Several products of varying lengths were identified. The longest RACE-product extended 56 nucleotides upstream of the translation initiation site in the type II cGK cDNA sequence. Since the 3'-end of exon 1 corresponded to position (–14), the 43-nucleotide overlap

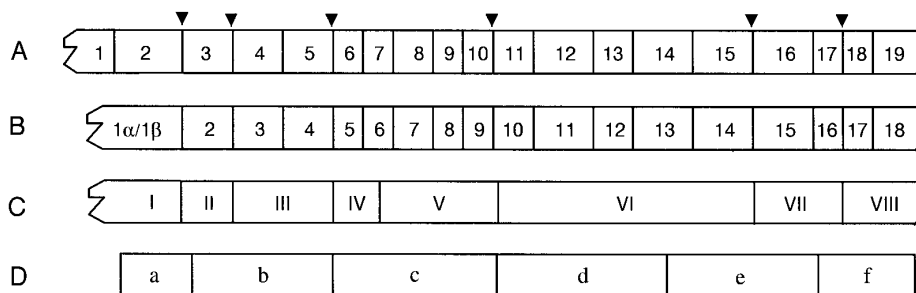


FIG. 3. Comparison of exon organization of the genes for human type II cGK (A), human type I cGK (B) and the *D. melanogaster* cGK; DG2:T1 (C) and their relation to functional domains of the bovine type I α cGK protein (D). The different boxes represent the exons encoding the respective genes and are given in arabic numerals for the human genes and in roman numerals for the *D. melanogaster* gene. The functional domains are illustrated according to that determined by Takio *et al.* (23). The N-terminal domain corresponds to the dimerization domain (a), followed by the two cGMP-binding domains (b and c). C-terminal to these binding sites are the catalytic domain (d), the ATP-binding domain (e) and a carboxyl-terminal domain (f). Arrows indicate the conserved splice sites between the human and *D. melanogaster* genes. All sequences were aligned with gaps introduced to obtain maximum homology.

with 100 % identity between the two sequences proved that exon 1 is indeed expressed in human tissues and not only in mouse.

The Length of the Gene Encoding Type II cGK

The lengths of the introns were determined using PCR with oligonucleotide primer pairs based on the exons flanking the various introns. This strategy was successful for all introns except introns 1 and 14. The shortest intron was intron 2 which was 0.6 kb in length, whereas the longest characterized intron was the 12-kb long, intron 6. The 120-kb PAC clone contained exons 2-18 and introns 2-17, as well as parts of introns 1 and 18. Using PCR, it was determined that intron 18 is 2.5 kb in length. It can therefore be concluded that the total length of introns 1 and 14 put together, was at least 53.5 kb (the length of the PAC clone minus the length of the exons and the length of introns 2-13 and 15-18). The fact that the total length of introns 1 and 14 exceed 50 kb can also explain why we were unable to determine their lengths using PCR, since it is difficult to amplify fragments of this size using PCR. The gene itself must span more than 125 kb as this is the total length of the PAC clone and exons 1 and 19.

DISCUSSION

This study presents the structure of the gene encoding the human type II cGK and is the first report on the structure of this gene from any mammal. The gene is at least 125 kb in length and is divided into 19 exons. The cDNA sequence encoding human type II cGK has previously been published (14,15) as well as the corresponding cDNA sequences from mouse (12) and rat (13). When these cDNA sequences are compared, a striking difference is the fact that the reported murine cDNA sequence contains 442 nucleotides 5' of the start codon while the reported human and rat sequences con-

tain only 14 and 47 nucleotides 5' of the start codon, respectively. Using an oligonucleotide corresponding to the 5'-end of the murine type II cGK cDNA sequence as a probe, we have been able to show that an untranslated exon 1 exists in man. This exon is homologous to the long 5' untranslated end of the murine type II cGK. Interestingly, out of the 47 published nucleotides from the 5' untranslated region in the rat type II cGK cDNA, the 13 nucleotides located most 3' corresponded well with the murine and human sequences, while the most 5' 34 nucleotides, in contrast, showed no similarity to the two other species. This discrepancy can be interpreted in several ways. Since the human exon 2 started at nucleotide (-13), the discrepancy between the two species can be due to incomplete splicing out of introns in the published rat sequence. This is conceivable, as nucleotides (-15)-(-14) in the rat sequence were "AG", the consensus splice acceptor sequence (34). However, it cannot be ruled out that the differences between the rat sequence and the other sequences represent species differences. An attractive explanation is that alternative exons exist for type II cGK and that the isolated rat sequence represented one splice variant, and the murine and human cDNA sequences represented another, alternatively spliced, mRNA. The presence of alternative, untranslated exons have been shown for several genes like the human gene encoding the RI α subunit of the (35). Although we do not have any evidence of alternative splicing of the type II cGK, alternative splicing and/or closely related isoforms have been found for type I cGK as well as for all the regulatory and catalytic subunits of the closely related cAKs (36). It is therefore a likely hypothesis that a multiplicity of isoforms also exist for the type II cGK.

Several independent amplifications of the 5' end of the human type II cGK cDNA were performed in order to determine the start of transcription of the human type II cGK using 5'-RACE. The ends of the cDNA

clones produced were located in the interval (–56) to (–3) relative to the start of translation. However, based upon the extensive size of the 5' non-translated region of the murine type II cGK cDNA and our present data, it seems unlikely that the start of transcription is located as close to the start of initiation as the 5' RACE products suggest. Furthermore, it must be emphasized that our results only demonstrate that exon 1 is transcribed in human prostate and offer no information about other human tissues. Studies of other genes have shown that the expression of untranslated exons can vary in a tissue-specific manner (37,38). This may also be the case for type II cGK. As indicated earlier, there is also a possibility that more than one untranslated exon exist.

The gene encoding the human type I cGK has been characterized and shown to consist of 19 exons encompassing at least 220 kb. The type I cGK gene is alternatively spliced and two isoforms, designated type I α and type I β cGK, exist. The type I and type II cGK have the same functional domains and their primary structures are similar (Fig. 3). At the cDNA level, the sequences have a 63 % identity when the most 5' regions, encoding the dimerization domain, have been excluded from the comparison. In the genomic regions corresponding to these parts of the cDNAs, both genes are divided into 17 exons and all splice junctions coincide between the two genes. However, in the remainder 5'-end of the cDNA sequences, the sequences representing the two isoforms differ substantially. The genomic region corresponding to this part of the cDNA consists of two exons that are both expressed in the type II cGK. In the type I cGK, the corresponding region is represented by two alternatively spliced exons giving rise to two different type I cGK cDNAs. In all instances, this variable part of the proteins confer the same property, namely enabling them to exist as dimers *in vivo*. In addition, as shown in Table 1, most of the introns are in phase 1 or 2, meaning that they do interrupt the coding triplets. This indicates that the gene probably has not arisen by shuffling of mini genes each encoding proteins with distinct properties.

The 5' region of the gene contains three putative TATA-boxes (Fig. 2). In addition, it contains G/C-rich sequences and several copies of the hexanucleotide G/C-rich motif (GC-box) that binds the mammalian transcription factor SP-1. Both the TATA-boxes and SP-1 sites are located in the interval (–60) to (–360) relative to the start of translation (ATG). An SP-1 site and a site for TFIID are located within the potential CpG island and it is a typical feature for CpG islands to contain both the promoter and one or several exons within the region (33). Two putative TATA-boxes have been identified more than 1400 nucleotides upstream of ATG. It is not possible to say if these putative TATA-boxes are real, as we do not know the entire length of

exon 1. It is also possible that the gene for the human type II cGK contains more than one promoter.

Jarchau and co-workers (1994) have earlier published a phylogenetic tree based on the cDNA sequences of mammalian cGK types I and II and *D. melanogaster* cGK (DG1 and DG2), and they concluded that the mammalian type II cGK and *Drosophila* DG1 genes diverged at an earlier time in evolution compared to the mammalian type I cGK and the *Drosophila* DG2. Based on the correlation of the splice sites of the human types I and II cGK and the *D. melanogaster* DG2 genes (Fig. 3), we conclude that these three genes are more closely related to each other than to the DG1 gene of the *D. melanogaster*. In addition, the genomic structure of the mammalian cGKs resemble each other more than any of the *D. melanogaster* cGKs. This suggests that the two human cGKs are the results of a gene duplication that occurred later in evolution than the separation of vertebrates from insects.

The isolation of the gene for the human type II cGK allows studies directed towards the understanding of the regulation of expression of this gene and opens the possibility to study, in detail, the regulatory mechanisms involved in tissue-specific expression.

ACKNOWLEDGMENTS

We acknowledge the technical assistance of Marianne Nordahl and Liv Paltiel. We thank Dr. Erik Dissen (Institute of Anatomy, University of Oslo) for advice concerning pulsed field gel electrophoresis. This work was supported by the Research Council of Norway (NFR), The Norwegian Cancer Society, Anders Jahre's Foundation for the Promotion of Science, Novo Nordisk Foundation and Johan Ekebergs legat.

REFERENCES

- Garbers, D. L., and Lowe, D. G. (1994) *J. Biol. Chem.* **269**, 30741–30744.
- Schmidt, H. H., and Walter, U. (1994) *Cell* **78**, 919–925.
- Lincoln, T. M., and Cornwell, T. L. (1993) *FASEB J.* **7**, 328–338.
- Francis, S. H., and Corbin, J. D. (1994) *Annu. Rev. Physiol.* **56**, 237–272.
- Forte, L. R., Thorne, P. K., Eber, S. L., Krause, W. J., Freeman, R. H., Francis, S. H., and Corbin, J. D. (1992) *Am. J. Physiol.* **263**, C607–C615.
- Chao, A. C., de Sauvage, F. J., Dong, Y. J., Wagner, J. A., Goeddel, D. V., and Gardner, P. (1994) *EMBO J.* **13**, 1065–1072.
- Shabb, J. B., and Corbin, J. D. (1992) *J. Biol. Chem.* **267**, 5723–5726.
- Wernet, W., Flockerzi, V., and Hofmann, F. (1989) *FEBS Lett.* **251**, 191–196.
- Sandberg, M., Natarajan, V., Ronander, I., Kalderon, D., Walter, U., Lohmann, S. M., and Jahnsen, T. (1989) *FEBS Lett.* **255**, 321–329.
- Sandberg, M. (1991) Molecular Cloning and Primary Structure of Mammalian Cyclic Nucleotide-Dependent Protein Kinases. University of Oslo, Norway.
- Ørstavik, S., Natarajan, V., Taskén, K., Jahnsen, T., and Sandberg, M. (1997) *Genomics* **42**, 311–318.

12. Uhler, M. D. (1993) *J. Biol. Chem.* **268**, 13586–13591.
13. Jarchau, T., Häusler, C., Markert, T., Pohler, D., Vanderkerckhove, J., de Jonge, H. R., Lohmann, S. M., and Walter, U. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 9426–9430.
14. Fujii, M., Ogata, T., Takahashi, E., Yamada, K., Nakabayashi, K., Oishi, M., and Ayusawa, D. (1995) *FEBS Lett.* **375**, 263–267.
15. Ørstavik, S., Solberg, R., Taskén, K., Nordahl, M., Altherr, M. R., Hansson, V., Jahnsen, T., and Sandberg, M. (1996) *Biochem. Biophys. Res. Commun.* **220**, 759–765.
16. Eigenthaler, M., Ullrich, H., Geiger, J., Horstrup, K., Hönig-Liedl, P., Wiebecke, D., and Walter, U. (1993) *J. Biol. Chem.* **268**, 13526–13531.
17. Tamura, N., Itoh, H., Ogawa, Y., Nakagawa, O., Harada, M., Chun, T. H., Suga, S., Yoshimasa, T., and Nakao, K. (1996) *Hypertension* **27**, 552–557.
18. de Jonge, H. R. (1981) *Adv. Cyclic Nucleotide Res.* **14**, 315–333.
19. Markert, T., Vaandrager, A. B., Gambaryan, S., Pöhler, D., Häusler, C., Walter, U., de Jonge, H. R., Jarchau, T., and Lohmann, S. M. (1995) *J. Clin. Invest.* **96**, 822–830.
20. French, P. J., Bijman, J., Edixhoven, M., Vaandrager, A. B., Scholte, B. J., Lohmann, S. M., Nairn, A. C., and de Jonge, H. R. (1995) *J. Biol. Chem.* **270**, 26626–26631.
21. Pfeifer, A., Aszódi, A., Seidler, U., Ruth, P., Hofmann, F., and Fässler, R. (1996) *Science* **274**, 2082–2086.
22. Lincoln, T. M., Dills, W. L. J., and Corbin, J. D. (1977) *J. Biol. Chem.* **252**, 4269–4275.
23. Takio, K., Wade, R. D., Smith, S. B., Krebs, E. G., Walsh, K. A., and Titani, K. (1984) *Biochemistry* **23**, 4207–4218.
24. Gamm, D. M., Francis, S. H., Angelotti, T. P., Corbin, J. D., and Uhler, M. D. (1995) *J. Biol. Chem.* **270**, 27380–27388.
25. Ioannou, P. A., Amemiya, C. T., Garnes, J., Kroisel, P. M., Shizuya, H., Chen, C., Batzer, M. A., and de Jong, P. J. (1994) *Nat. Genet.* **6**, 84–89.
26. Church, G. M., and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1991–1995.
27. Ioannou, P. A., and de Jong, P. J. (1996) in *Current Protocols in Human Genetics* (Dracopoli, N. C., Haines, J. L., Korf, B. R., Moir, D. T., Morton, C. C., Seidman, C. E., Seidman, J. G., and Smith, D. R., Eds.), pp. 5.15.1–5.15.24, Wiley, New York.
28. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
29. Feinberg, A. P., and Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–267.
30. Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 8794–8797.
31. Kim, U. J., Birren, B. W., Slepak, T., Mancino, V., Boysen, C., Kang, H. L., Simon, M. I., and Shizuya, H. (1996) *Genomics* **34**, 213–218.
32. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) *Nucleic Acids Res.* **25**, 3389–3402.
33. Cross, S. H., Charlton, J. A., Nan, X., and Bird, A. P. (1994) *Nat. Genet.* **6**, 236–244.
34. Shapiro, M. B., and Senapathy, P. (1987) *Nucleic Acids Res.* **15**, 7155–7174.
35. Solberg, R., Sandberg, M., Natarajan, V., Torjesen, P. A., Hansson, V., Jahnsen, T., and Taskén, K. (1997) *Endocrinology* **138**, 169–181.
36. Taskén, K., Solberg, R., Foss, K. B., Skålhegg, B. S., Hansson, V., and Jahnsen, T. (1995) *The Protein Kinase FactsBook: Protein-Serine Kinases*, Academic Press, San Diego.
37. Yamada-Mouri, N., Hirata, S., and Kato, J. (1996) *J. Steroid Biochem. Mol. Biol.* **58**, 163–166.
38. Hirata, S., Koh, T., Yamada-Mouri, N., and Kato, J. (1996) *Biochem. Biophys. Res. Commun.* **225**, 849–854.