

# Multiple genes encode the human Na<sup>+</sup>,K<sup>+</sup>-ATPase catalytic subunit

(multiple isoforms/genomic library)

MARCIA M. SHULL AND JERRY B LINGREL\*

Department of Microbiology and Molecular Genetics, University of Cincinnati, College of Medicine, 231 Bethesda Avenue, ML 524, Cincinnati, OH 45627-0524

Communicated by I. S. Edelman, February 27, 1987

**ABSTRACT** A human genomic library was constructed and screened with hybridization probes derived from sheep and rat cDNAs encoding the  $\alpha$  and  $\alpha(+)$  isoforms, respectively, of the Na<sup>+</sup>,K<sup>+</sup>-ATPase catalytic subunit. Genomic sequences spanning 150 kilobases were isolated. Four genes, designated  $\alpha A$ ,  $\alpha B$ ,  $\alpha C$ , and  $\alpha D$ , each 20–25 kilobases in length, were identified by restriction mapping, Southern blot hybridization analysis, and limited DNA sequencing. We present evidence that two of these genes,  $\alpha A$  and  $\alpha B$ , encode the  $\alpha$  and  $\alpha(+)$  isoforms, respectively. The other genes,  $\alpha C$  and  $\alpha D$ , one of which is physically linked to the  $\alpha(+)$  gene, exhibit nucleotide and amino acid homology to Na<sup>+</sup>,K<sup>+</sup>-ATPase catalytic subunit cDNA sequences but do not correspond to any previously identified isoforms.

The Na<sup>+</sup>,K<sup>+</sup>-ATPase is an integral membrane protein responsible for establishing and maintaining the electrochemical gradients of Na<sup>+</sup> and K<sup>+</sup> across the plasma membrane of animal cells. As these gradients are essential for osmoregulation, Na<sup>+</sup>-coupled transport of a variety of organic and inorganic molecules, and electrical excitability of nerve and muscle, the enzyme plays a central role in cellular physiology. It is composed of two subunits, a large catalytic subunit ( $\alpha$ ) and a smaller glycoprotein subunit ( $\beta$ ) of unknown function. Biochemical studies have demonstrated the existence of two isoforms of the catalytic subunit,  $\alpha$  and  $\alpha(+)$  (for a review, see ref. 1). Kidney contains predominantly the  $\alpha$  form, while both  $\alpha$  and  $\alpha(+)$  are found in brain (2), adipose tissue (3), and skeletal muscle (3, 4). Besides differences in tissue distribution, the two isoforms exhibit differential expression during development and in response to hormones (5, 6).

In addition to  $\alpha$  and  $\alpha(+)$ , yet a third isoform ( $\alpha III$ ) has recently been identified in rat brain by use of molecular cloning techniques (7), raising the possibility that isoforms other than the three identified to date may exist in certain tissues. Although cDNAs encoding the catalytic-subunit isoforms from a number of species and tissues have been characterized (7–12), the corresponding genomic sequences have not been isolated. It seems clear from cDNA cloning studies that the three known isoforms,  $\alpha$ ,  $\alpha(+)$ , and  $\alpha III$ , arise from separate genes (7), but the total number of genes has not been determined. Also, it is not known whether other mechanisms, such as the use of alternative exons within a single gene, are responsible for generating additional diversity in Na<sup>+</sup>,K<sup>+</sup>-ATPase isoforms. In order to address the question of the genetic basis for the multiple isoforms and to examine the possibility that additional Na<sup>+</sup>,K<sup>+</sup>-ATPase isoforms exist, we constructed and screened a human genomic library. The data presented in this paper demonstrate that the human Na<sup>+</sup>,K<sup>+</sup>-ATPase catalytic subunit is encoded by multiple genes. Separate genes encoding the  $\alpha$  and  $\alpha(+)$  isoforms were identified. In addition, we isolated two other

genes that exhibit high sequence similarity to Na<sup>+</sup>,K<sup>+</sup>-ATPase catalytic-subunit sequences.

## MATERIALS AND METHODS

**Construction and Screening of Human Genomic Library.** Human leukocyte nuclei were prepared according to Bell *et al.* (13). Genomic DNA was isolated from the nuclei, partially digested with *Sau3A1* and fractionated by sucrose gradient centrifugation (14). DNA fragments 15–23 kilobases (kb) in length were ligated to  $\lambda$ EMBL4 arms (15) and packaged using a commercial packaging extract (Stratagene Cloning Systems, San Diego, CA). The unamplified library was screened by filter hybridization (16) using a nick-translated 2.7-kb *Xho*I–*Bgl*II fragment of the sheep kidney Na<sup>+</sup>,K<sup>+</sup>-ATPase  $\alpha$ -subunit cDNA (8) as a probe. Filters were prehybridized for 16 hr at 63°C in 6 $\times$  SET (1 $\times$  SET is 0.15 M NaCl/0.02 M Tris-HCl, pH 7.8/1 mM EDTA) containing 10 $\times$  Denhardt's solution (1 $\times$  Denhardt's solution is 0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% bovine serum albumin), 0.1% NaDodSO<sub>4</sub> and 100  $\mu$ g of salmon sperm DNA per ml and then were hybridized at 63°C for 40 hr in the same solution containing 3  $\times$  10<sup>6</sup> cpm of probe per ml. Filters were washed successively at 63°C in 6 $\times$  SSC/0.1% NaDodSO<sub>4</sub> (1 $\times$  SSC is 0.15 M NaCl/0.015 M sodium citrate, pH 7.0), 3 $\times$  SSC/0.1% NaDodSO<sub>4</sub>, 1.5 $\times$  SSC/0.1% NaDodSO<sub>4</sub>, and 1 $\times$  SSC/0.1% NaDodSO<sub>4</sub>. After autoradiography, the same filters were rescreened with a nick-translated 4.9-kb insert of the rat brain  $\alpha(+)$  cDNA (7).

**Characterization of Genomic Clones.** Phage DNA was isolated; digested with the restriction enzymes *Eco*RI, *Bam*HI, and *Hind*III; fractionated by electrophoresis in 0.8% agarose gels; and transferred to nylon membranes (17). The identity and order of restriction fragments was determined by successively hybridizing DNA on the filters to a series of <sup>32</sup>P-labeled restriction fragments derived from the sheep kidney  $\alpha$ -subunit cDNA (8) and to isoform-specific probes derived from rat brain catalytic-subunit cDNAs (7) as described below.

**cDNA Probes.** Seven restriction fragments of the sheep kidney  $\alpha$ -subunit cDNA were used in the mapping studies (see Fig. 1 legend for a description of these probes). The fragments were labeled by nick-translation to a specific activity >2.0  $\times$  10<sup>8</sup> cpm/ $\mu$ g (18). Filters were prehybridized at 60°C in 6 $\times$  SSC/5 $\times$  Denhardt's solution/0.1% NaDodSO<sub>4</sub> containing 100  $\mu$ g of salmon sperm DNA per ml. After overnight hybridization in the same solution containing 10<sup>5</sup> cpm of probe per ml, the filters were washed at 60°C in 3 $\times$  SSC/0.1% NaDodSO<sub>4</sub>.

**Isoform-Specific Probes.** In pairwise comparisons of cDNA sequences of  $\alpha$ ,  $\alpha(+)$ , and  $\alpha III$  from rat brain (7), the N-terminal regions exhibited the greatest divergence. Based on this information, probes specific for the  $\alpha$ ,  $\alpha(+)$ , and  $\alpha III$

isoforms were prepared as follows. Restriction fragments selected from the regions of least similarity were isolated from the rat brain cDNA clones. The fragments were labeled to a specific activity of  $3 \times 10^9$  cpm/ $\mu$ g essentially as described by Feinberg and Volgelstein (19), except that specific oligonucleotides corresponding to unique regions near the translation start site of each cDNA were used as primers. For the  $\alpha$ -specific probe, an *Nco*I-*Dde*I fragment corresponding to nucleotides (nt) 237-373 was used as a template and an oligonucleotide corresponding to nt 265-279 was used as a primer. For  $\alpha(+)$ , a *Bam*HI-*Nco*I fragment (nt 43-208) was used as template and a 15-mer corresponding to nt 121-135 was used as primer. For  $\alpha$ III, the template fragment was an *Ava*I fragment (nt 65-328) and the primer was an oligonucleotide corresponding to nt 156-170. Hybridization to cloned DNAs on Southern blots was performed as described for the cDNA fragment probes; filters were washed at 65°C for 30 min in  $0.5 \times$  SSC/0.1% NaDodSO<sub>4</sub>.

**Genomic Southern Blot Analysis.** Human genomic DNA was digested to completion with *Eco*RI, fractionated by electrophoresis in 0.8% agarose, and transferred to nylon membranes. Blots were prehybridized overnight at 45°C in  $6 \times$  SSC/10 $\times$  Denhardt's solution/1% NaDodSO<sub>4</sub> containing 100  $\mu$ g of salmon sperm DNA per ml and then were hybridized at 42°C to the isoform-specific probes ( $10^6$  cpm/ml) in  $6 \times$  SSC/50% (vol/vol) formamide/5% dextran sulfate/1% NaDodSO<sub>4</sub> containing salmon sperm DNA at 100  $\mu$ g/ml. Filters were washed at 60°C in  $3 \times$  SSC/0.1% NaDodSO<sub>4</sub> for 30 min.

**DNA Sequencing.** Restriction fragments were subcloned into M13mp19 (20) and sequenced by the dideoxy chain-termination method of Sanger *et al.* (21), using 2'-deoxyadenosine 5'-[ $\alpha$ -<sup>35</sup>S]thio]triphosphate.

## RESULTS

**Isolation of Human Na<sup>+</sup>,K<sup>+</sup>-ATPase Sequences.** Approximately  $1.2 \times 10^6$  recombinant bacteriophage were screened

at low hybridization stringency, using probes derived from cDNAs encoding the  $\alpha$  and  $\alpha(+)$  isoforms of the catalytic subunit. The first probe contained most of the coding region of the sheep  $\alpha$  isoform (8) and yielded 26 positive clones. After gentle removal of the probe ( $0.1 \times$  SSC/0.1% NaDodSO<sub>4</sub>, 63°C, 45 min), the filters were reprobated with a rat brain cDNA corresponding to the entire coding region and the 5' and 3' untranslated regions of the mRNA for the  $\alpha(+)$  isoform (7). This probe identified 21 positive clones, 17 of which had been isolated during the first screening.

**Four Genes Identified by Restriction Endonuclease Mapping and Hybridization Analysis.** DNAs from the 30 positive clones were analyzed by single and double restriction enzyme digestions and Southern blot hybridization using cDNA probes. The sheep kidney  $\alpha$ -subunit cDNA fragments (see Fig. 1 legend) used as hybridization probes, each 300-400 base pairs (bp), spanned the entire  $\alpha$ -subunit coding region and included 175 bp of 5' untranslated sequence and 193 bp of 3' untranslated sequence. Based on these analyses, the clones could be placed into four classes designated  $\alpha$ A,  $\alpha$ B,  $\alpha$ C, and  $\alpha$ D (Fig. 1). Since the entire coding region was represented in the cDNA probes and all of the probes hybridized with each class of clones (Fig. 1), it seems likely that each class represents a complete gene.

The presence of a large number of overlapping clones in the  $\alpha$ A and  $\alpha$ C genes aided restriction mapping. In order to confirm the map in regions containing only a few overlapping clones, restriction fragments that appeared to represent overlaps were isolated from their respective phage clones and digested in both single and double digests with a series of restriction enzymes having 4-bp recognition sequences. The resulting fragments were end-labeled using [ $\gamma$ -<sup>32</sup>P]ATP and polynucleotide kinase, fractionated in adjacent lanes of 8% polyacrylamide sequencing gels, and exposed to x-ray film. Since fragments differing by as little as 1 bp could be resolved in these gels, overlaps could be confirmed by comparing the complex autoradiographic patterns of the digested fragments.

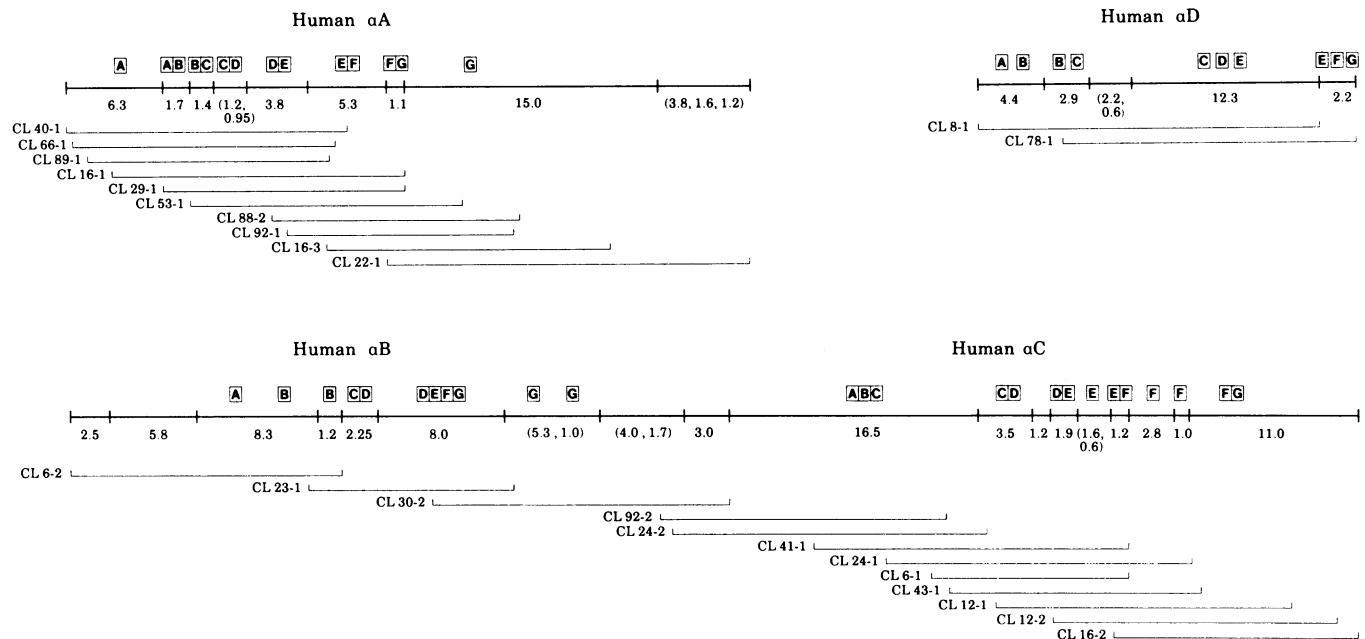


FIG. 1. Restriction maps of four classes of human sequences that hybridize to Na<sup>+</sup>,K<sup>+</sup>-ATPase catalytic-subunit probes. Composite *Eco*RI restriction maps for each of the four classes are shown above the phage clones from which they were derived. Locations of *Eco*RI sites are indicated by vertical bars; sizes of resulting fragments are given in kb. Numbers in parentheses indicate that the order of the fragments was not determined. Hybridization analysis was performed using restriction fragments from a sheep kidney  $\alpha$ -subunit cDNA. Boxed letters above the maps indicate to which cDNA fragments the genomic fragments hybridized. Sheep kidney  $\alpha$ -subunit cDNA (8) fragments used as probes were as follows: A, *Ava*I-*Ava*I, nt 106-454; B, *Ava*I-*Xba*I, nt 455-965; C, *Xba*I-*Bam*HI, nt 965-1359; D, *Bam*HI-*Bam*HI, nt 1360-1795; E, *Nco*I-*Nco*I, nt 1827-2456; F, *Nco*I-*Bgl*II, nt 2457-3112; G, *Bgl*II-*Ava*I, nt 3113-3523.

This procedure was used to verify overlaps in the  $\alpha$ B,  $\alpha$ C, and  $\alpha$ D gene classes. In this manner, for example, the 3.0-kb fragments of clones 30-2, 92-2, and 24-2 were shown to be identical, demonstrating that the  $\alpha$ B and  $\alpha$ C genes are physically linked.

**Identification of Genes Encoding Known  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase Isoforms.** To determine whether any of the genes encoded the  $\alpha$ ,  $\alpha(+)$ , or  $\alpha$ III isoforms, *Eco*RI-digested cloned DNAs were hybridized with isoform-specific probes derived from the N-terminal coding regions of rat brain cDNAs. Representative clones from each gene class are shown in Fig. 2. The  $\alpha$ A gene appears to encode the  $\alpha$  isoform, since clone 40-1, representative of  $\alpha$ A, contains two *Eco*RI fragments, 1.7 and 1.4 kb long, that hybridize only with the  $\alpha$ -specific probe. Clone 6-2, representative of  $\alpha$ B, contains an 8.3-kb fragment that hybridizes strongly to the  $\alpha(+)$ -specific probe, suggesting that the  $\alpha$ B gene encodes the  $\alpha(+)$  isoform. The  $\alpha(+)$ -specific probe also hybridizes weakly to clone 24-2 ( $\alpha$ C), which is linked to  $\alpha$ B sequences. However, clone 24-2 ( $\alpha$ C) does not exhibit strong hybridization with any of the isoform-specific probes. Clone 8-1, representative of  $\alpha$ D, did not hybridize to any of the isoform-specific probes.

In order to confirm that the isoform-specific fragments observed in *Eco*RI digests of cloned DNAs correspond to unique fragments present in the genome, *Eco*RI-digested human genomic DNA was hybridized with the  $\alpha$ -,  $\alpha(+)$ -, and  $\alpha$ III-specific probes. The  $\alpha$ -specific probe hybridizes to two fragments, 1.7 and 1.4 kb in length (Fig. 2A), which are the same sizes as those observed when clone 40-1, representative of the  $\alpha$ A gene, is hybridized with this probe. In addition, an 8.3-kb band is observed. Hybridization with the  $\alpha(+)$ -specific probe (Fig. 2B) suggests that this band represents the  $\alpha(+)$  gene, which cross-hybridizes with the  $\alpha$  probe under the low-stringency conditions used for genomic analysis. The

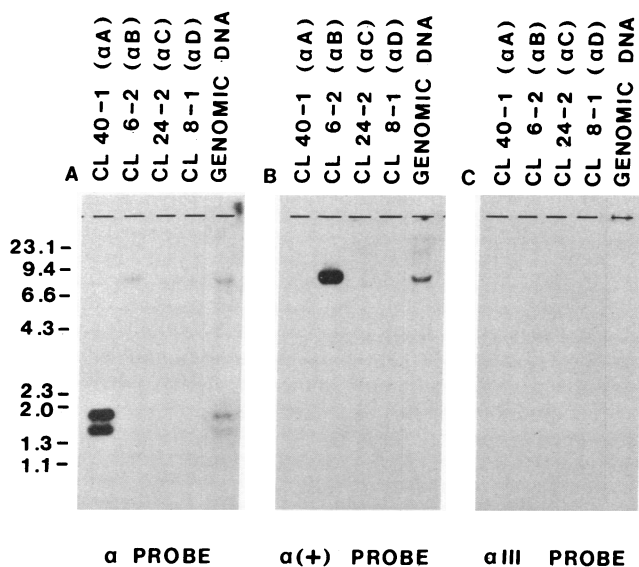


FIG. 2. Southern blot hybridization of representative cloned DNAs and genomic DNA to isoform-specific probes. DNA (1.5  $\mu$ g) from clones 40-1 ( $\alpha$ A), 6-2 ( $\alpha$ B), 24-2 ( $\alpha$ C), and 8-1 ( $\alpha$ D) was digested to completion with *Eco*RI, fractionated by electrophoresis in 0.8% agarose, and transferred to nylon membranes. Human genomic DNA (7.5  $\mu$ g) was analyzed in the same manner. DNA bound to the filters was hybridized to either an  $\alpha$ -specific probe (A), an  $\alpha(+)$ -specific probe (B), or an  $\alpha$ III-specific probe (C). The 8.3-kb band in the genomic DNA lane A is due to cross-hybridization of the  $\alpha$  probe with the  $\alpha(+)$  gene under the low-stringency wash conditions employed for analysis of the genomic DNA. The 16.5-kb band in the genomic DNA lane in B is due to cross-hybridization of the  $\alpha(+)$  probe with the  $\alpha$ C gene; the faint upper band in this same lane appears to be a partial digestion product.

8.3-kb genomic fragment hybridizing with the  $\alpha(+)$ -specific probe is the same size as that observed when the  $\alpha$ B clone is hybridized with this probe. In addition, a weaker, 16.5-kb band is seen; this band represents cross-hybridization of  $\alpha$ C sequences with the  $\alpha(+)$  probe. The faint uppermost band appears to represent a partial digestion product. The  $\alpha$ III-specific probe fails to hybridize to any band in the genomic digest under the hybridization and wash conditions used (Fig. 2C).

**Nucleotide and Amino Acid Sequence Analysis.** To verify the identity of the  $\alpha$ A and  $\alpha$ B genes and to determine whether all four genes encode  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase catalytic subunit sequences, limited DNA sequence analysis was performed. Approximately 6.7 kb of sequence was determined, most of which represented intron sequences. To identify protein-coding regions, the sequences were translated in all three reading frames and the deduced amino acid sequences were compared to the amino acid sequences of the rat  $\alpha$ -subunit isoforms. Intron/exon boundaries were readily apparent by an abrupt end to amino acid homology, by the presence of sequences homologous to splice site (22) and lariat acceptor site (23) consensus sequences in the expected positions, and by comparison to corresponding regions of the sheep  $\alpha$ -subunit gene (M.M.S., J.B.L., J. Greeb, and G. E. Shull, unpublished data). Only those regions representing coding sequences are shown in Figs. 3 and 4. The nucleotide sequence of fragments from clone 40-1 ( $\alpha$ A), clone 23-1 ( $\alpha$ B), clone 41-1 ( $\alpha$ C) and clone 78-1 ( $\alpha$ D) are shown in Fig. 3. Deduced amino acid sequences of these same regions are shown in Fig. 4.

The sequence of clone 40-1 ( $\alpha$ A) shown in Fig. 3 exhibits 100% nucleotide identity to the published sequence of the corresponding region of the HeLa (human)  $\alpha$ -subunit cDNA (12) and an overall identity of 89% to the same region of the rat  $\alpha$  isoform cDNA (7). At the amino acid level, identity to the HeLa  $\alpha$  sequence is 100% and identity to the rat  $\alpha$  sequence is 98% (Fig. 4). (As the amino acid identity of HeLa  $\alpha$  to rat  $\alpha$  is 97%, the same as that observed in pairwise comparisons between rat, sheep, and pig  $\alpha$  isoforms, it seems clear that the HeLa sequence represents the  $\alpha$  isoform.) These results confirm that  $\alpha$ A represents the gene encoding the  $\alpha$  isoform. When the nucleotide sequence of clone 23-1 ( $\alpha$ B) is compared to sequences of the corresponding regions of the HeLa  $\alpha$  and rat brain  $\alpha$ ,  $\alpha(+)$ , and  $\alpha$ III cDNAs, the highest similarity (92%) is seen with the rat  $\alpha(+)$  sequence. This degree of nucleotide identity is similar to that observed when comparing the  $\alpha$  isoforms of different mammalian species. The amino acid identity of  $\alpha$ B with rat  $\alpha(+)$  is 100%, whereas identity with HeLa or rat  $\alpha$  is 92% and with rat  $\alpha$ III is 90%. These homology comparisons provide confirmation that  $\alpha$ B encodes the  $\alpha(+)$  isoform. The sequence from clone 41-1 ( $\alpha$ C) exhibits high nucleotide (76–81%) and amino acid (80–89%) similarity to the HeLa  $\alpha$  and rat  $\alpha$ ,  $\alpha(+)$ , and  $\alpha$ III isoform sequences. The sequence from clone 78-1 ( $\alpha$ D) has a nucleotide identity of 68–76% and an amino acid identity of 66–75% with the HeLa  $\alpha$  and rat  $\alpha$ ,  $\alpha(+)$ , and  $\alpha$ III isoform sequences.

## DISCUSSION

Using sheep  $\alpha$ -subunit and rat  $\alpha(+)$ -subunit cDNA as probes, we isolated four classes of human genomic clones exhibiting high sequence similarity to  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase catalytic-subunit sequences. Southern hybridization to isoform-specific probes and DNA sequence analysis showed that two of these classes,  $\alpha$ A and  $\alpha$ B, represent separate genes encoding the  $\alpha$  and  $\alpha(+)$  isoforms, respectively. A third hybridizing sequence,  $\alpha$ C, is physically linked to the  $\alpha$ B gene by 10–20 kb and appears to represent  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase sequences, as it hybridizes strongly with each fragment from the sheep kidney  $\alpha$ -subunit cDNA used as a probe and exhibits high nucleotide

	739		
Hu $\alpha$ A	AG/CAAGCCCTGTGATTTCGAAATGGTGAGAAAAATGACATAAATGCGGAGGAAGTTGGTTGGGGAT		
HeLa $\alpha$	-----A-----G-----C-----A-----T-----C-----T-----		
Rat $\alpha$	-----C-----G-----A-----G-----C-----A-----G-----G-----C-----G-----A-----C-----		
Rat $\alpha$ +	-----T-----G-----C-----C-----G-----G-----A-----G-----C-----A-----G-----G-----C-----G-----A-----C-----		
Rat $\alpha$ III	-----C-----GG-----A-----A-----G-----CAGG-----G-----G-----G-----C-----C-----		
	874		
Hu $\alpha$ A	AG/GTGGATnACTCTCGCTCAGTGGTGAATCAGAACCAGACTAGGCTCCAGATTCACAATGAAAACCCCTGGAGACGAGGAACATTCCTTC		
HeLa $\alpha$	-----A-----		100% *
Rat $\alpha$	-----A-----A-----C-----G-----A-----G-----G-----G-----C-----C-----G-----C-----G-----T-----A-----		89% *
Rat $\alpha$ +	-----A-----A-----C-----G-----A-----G-----G-----G-----C-----C-----T-----G-----C-----G-----T-----T-----CC-----C-----T-----CTGT-----		72% *
Rat $\alpha$ III	-----CA-----T-----C-----G-----C-----T-----G-----T-----CC-----C-----G-----C-----G-----C-----C-----A-----CC-----CA-----		76% *
	874		
Hu $\alpha$ B	AG/GTGGATAACTCATCCTTAACAGGAGAGTCGGAGCCCCAGACCCGCTCCCCGAGTTCACCCATGAGAACCCTGGAGACCCGCAATATCTGTTTCTCCACCAACTGTGTTGAAG/GT		
HeLa $\alpha$	-----C-----GC-----C-----T-----A-----A-----A-----TA-----G-----T-----A-----T-----AA-----A-----GA-----G-----C-----TGCC-----T-----A-----A-----T-----		74%
Rat $\alpha$	-----C-----AC-----C-----T-----T-----A-----A-----A-----T-----G-----G-----T-----AA-----C-----T-----AA-----G-----C-----TGCC-----A-----		78%
Rat $\alpha$ +	-----C-----G-----		92%
Rat $\alpha$ III	-----C-----T-----C-----G-----T-----C-----A-----T-----T-----G-----C-----G-----A-----C-----C-----A-----G-----C-----ACC-----T-----G-----		80%
	1939		
Hu $\alpha$ C	GGATCCCATTTAATACAGATGAAATAAATTTCCCATGGACAACCTTTGTTTGGGGCCTCATATCCATGATTGACCCCTCCCGAGCTGCAGTGCCTGATGCTGTGAGCAAGTGTCCGAGTGCAGGAATTAAG		
HeLa $\alpha$	-----AG-----G-----C-----T-----TG-----G-----T-----C-----T-----T-----G-----C-----T-----G-----C-----T-----G-----A-----G-----C-----T-----C-----G-----A-----A-----T-----		78%
Rat $\alpha$	-----C-----AG-----G-----C-----T-----G-----C-----G-----T-----C-----C-----C-----T-----T-----C-----C-----G-----A-----C-----C-----T-----G-----		81%
Rat $\alpha$ +	-----C-----AA-----G-----C-----G-----GC-----G-----C-----T-----CA-----G-----G-----C-----C-----G-----T-----T-----C-----A-----A-----T-----A-----G-----CA-----A-----C-----C-----C-----CAA-----		76%
Rat $\alpha$ III	-----C-----TG-----C-----G-----CTGT-----CG-----G-----C-----A-----CA-----C-----C-----T-----G-----C-----C-----G-----A-----T-----C-----G-----A-----C-----C-----C-----		78%
	739		
Hu $\alpha$ D	AG/CAAGCTCTCGTCATCCGAGATTCCGAGAAGAAGACCCTCCCTCAGACAGCTGGTGGTGGGGGACATGTGGAGGTCAAAGGAGGAGACAGATCCCTGCAGACATCAGGGTGTCTTCTCAGGGGTGTCGG/GT		
HeLa $\alpha$	-----C-----T-----G-----T-----A-----GGT-----A-----T-----G-----AAA-----G-----G-----G-----AG-----T-----T-----TC-----G-----A-----A-----GA-----T-----T-----C-----AA-----CA-----A-----G-----AA-----T-----C-----CAA-----		
Rat $\alpha$	-----C-----G-----T-----A-----GGG-----T-----G-----AAC-----G-----TG-----C-----C-----T-----TC-----G-----G-----C-----GA-----T-----TC-----AA-----CA-----A-----G-----AA-----T-----C-----CAA-----		
Rat $\alpha$ +	-----G-----G-----G-----GGGA-----T-----CAG-----AA-----G-----G-----G-----C-----A-----C-----G-----A-----G-----G-----T-----GTG-----C-----T-----C-----C-----A-----CA-----C-----C-----C-----T-----CAA-----		
Rat $\alpha$ III	-----C-----T-----G-----G-----AGGT-----A-----T-----CAGG-----GAA-----G-----G-----G-----C-----C-----G-----A-----G-----T-----C-----G-----G-----C-----T-----CG-----CA-----CA-----C-----GG-----C-----T-----CAA-----		
	874		
Hu $\alpha$ D	AG/GTGGATAACTCATCTCACGGGGAGTCTGAGCCCCAGCCCCGCTCCCTCTGAGTTTACCCATGAAAACCCCTGGAA		
HeLa $\alpha$	-----C-----G-----G-----T-----T-----A-----A-----A-----A-----TA-----G-----TC-----A-----T-----C-----AA-----		70%
Rat $\alpha$	-----C-----A-----T-----T-----A-----A-----A-----A-----T-----G-----C-----G-----T-----C-----AA-----C-----G-----T-----G-----		70%
Rat $\alpha$ +	-----C-----G-----A-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----G-----		76%
Rat $\alpha$ III	-----C-----T-----C-----G-----T-----C-----A-----A-----T-----A-----C-----G-----C-----GC-----A-----C-----C-----A-----G-----		68%

FIG. 3. Nucleotide sequence analysis. The nucleotide sequence of fragments from clones 40-1 ( $\alpha$ A), 23-1 ( $\alpha$ B), 41-1 ( $\alpha$ C), and 78-1 ( $\alpha$ D) was determined. Only sequences corresponding to coding regions are shown. Below each of these sequences are the corresponding sequences from a HeLa cell  $\alpha$ -subunit cDNA (data from ref. 12) and from rat brain  $\alpha$ ,  $\alpha$ (+), and  $\alpha$ III cDNAs (data from ref. 7). Identical bases are indicated by dashes; n represents an undetermined nucleotide. Intron/exon boundaries are indicated by slash marks. The superscript numerals refer to the numbers of the corresponding nucleotides of the rat  $\alpha$ -isoform cDNA sequence. Percent identity of genomic clone sequences to the HeLa  $\alpha$  and rat brain  $\alpha$ ,  $\alpha$ (+), and  $\alpha$ III sequences are shown on the right. Stars indicate that the undetermined nucleotide was not included when calculating the percent identities.

(76–81%) and amino acid (80–89%) identity with the three previously identified isoforms. The fourth sequence,  $\alpha$ D, also hybridizes with all of the sheep kidney  $\alpha$ -subunit cDNA fragment probes, suggesting that it represents yet another  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase catalytic-subunit sequence. However, when the sequence of this gene is compared to the sequence of corresponding regions of three rat isoforms, the nucleotide (68–76%) and amino acid (66–75%) identities obtained are less than those observed in pairwise comparisons of the rat isoforms to each other; for the rat isoforms, nucleotide identities in this region range from 76% to 80%, while amino acid identities are 87–91%.

In Southern hybridization analysis, none of the phage clones hybridized with the  $\alpha$ III-specific probe. Furthermore, this probe did not yield a signal on genomic Southern blots. It is possible that the human does not have an isoform corresponding to the rat  $\alpha$ III isoform or that the corresponding isoform in humans has diverged from that of the rat in the region used as a specific probe, so that it does not hybridize to the probe under the conditions used. We cannot rule out the possibility that additional  $\alpha$ -subunit isoforms exist in the human genome. The hybridization stringencies used in screening the library allowed us to identify only two clones encoding the  $\alpha$ D gene but, in addition, allowed identification of a single clone encoding part of the  $\text{H}^+$ ,  $\text{K}^+$ -ATPase (M.M.S., G. E. Shull, and J.B.L., unpublished data). The rat  $\text{H}^+$ ,  $\text{K}^+$ -ATPase exhibits an overall nucleotide identity of 64% to the  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase  $\alpha$  isoform cDNA used as probe; however, certain regions of the  $\text{H}^+$ ,  $\text{K}^+$ -ATPase and  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase cDNAs exhibit greater identities, and these may have been responsible for the hybridization observed. Thus,

the existence of additional  $\alpha$ -subunit isoforms in the human genome cannot be precluded on the basis of the present data.

The  $\alpha$ C and  $\alpha$ D sequences do not correspond to any previously-identified isoforms; therefore, it is not known whether these genes encode functional isoforms of the  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase. The sequence obtained so far gives no indication that  $\alpha$ C or  $\alpha$ D represent pseudogenes. There are no in-frame stop codons in the coding regions of either gene, and in the case of  $\alpha$ D, the three intron/exon boundaries examined exhibit homology to splice-site consensus sequences (22) and correspond to the same boundaries seen in the sheep  $\alpha$ -subunit gene (data not shown). In addition, sequences homologous to a lariat acceptor site consensus sequence (23) are found in the correct position 5' to each of the two exons examined (data not shown).

Nucleotide and amino acid homology comparisons of  $\alpha$ D to published  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase sequences suggest that this gene could encode either an additional  $\alpha$  isoform or a closely related cation-transport ATPase. It is known that the  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase exhibits homology to other mammalian cation-transport ATPases such as the  $\text{Ca}^{2+}$ -ATPase and the  $\text{H}^+$ ,  $\text{K}^+$ -ATPase. However,  $\alpha$ D clearly does not encode either of these enzymes, as it exhibits only 72% amino acid identity with the corresponding region of the  $\text{H}^+$ ,  $\text{K}^+$ -ATPase (24) and only 31–34% identity with the  $\text{Ca}^{2+}$ -ATPase (25, 26). Southern hybridization analysis using the  $\text{H}^+$ ,  $\text{K}^+$ -ATPase cDNA as a probe provided further confirmation that the  $\alpha$ D gene does not encode the  $\text{H}^+$ ,  $\text{K}^+$ -ATPase (data not shown). However, it is conceivable that there are additional cation-transport ATPases with high sequence similarity to the  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase. For example, recent biochemical evidence

	163		
Hu $\alpha$ A	QALVIRNGEKMSINAEVVVDG		
HeLa $\alpha$	-----		
Rat $\alpha$	-----D-----		
Rat $\alpha(+)$	-----E-Q-----		
Rat $\alpha$ III	-----E-QV-----		
	208		
Hu $\alpha$ A	VDxSSLTGESEPTQTRSPDFNTNENPLETRNIAF		
HeLa $\alpha$	--N-----	100% *	
Rat $\alpha$	--N-----	98% *	
Rat $\alpha(+)$	--N-----E--H-----C-	91% *	
Rat $\alpha$ III	--N-----C-HD-----T-	87% *	
	208		
Hu $\alpha$ B	VDNSSLTGESEPTQTRSPDFNTNENPLETRNICFFSTNCVCE		
HeLa $\alpha$	-----D--N-----A-----	92%	
Rat $\alpha$	-----D--N-----A-----	92%	
Rat $\alpha(+)$	-----	100%	
Rat $\alpha$ III	-----DC--D-----T-----	90%	
	563		
Hu $\alpha$ C	GFPFNTDEINFPMDNLCFVGLISMIDPPRAAVPDAVSKCRSAGIK		
HeLa $\alpha$	--Q-D--DV--I-----G-----	87%	
Rat $\alpha$	--Q-D--V--V-----G-----	89%	
Rat $\alpha(+)$	--K-D--L--TEK--M-----G-----	82%	
Rat $\alpha$ III	--A-DC-DV--TT-----M-----G-----	80%	
	163		
Hu $\alpha$ D	QALVIRDSEKTIPTSEQLVVGDIIVEVKGDDQIPADIRVLSQGCRC		
HeLa $\alpha$	-----NG--MS-NA-EV-----L-----R-----L-II-AN--K		
Rat $\alpha$	-----NG--MS-NA-DV-----L-----R-----L-II-AN--K		
Rat $\alpha(+)$	-----EG--MQ-NA-EV-----L-----RV-----L-II--H--K		
Rat $\alpha$ III	-----EG--MQVNA-EV-----L--I-----RV--L-II-AH--K		
	208		
Hu $\alpha$ D	VDNSSLTGESEPTQTRSPDFNTNENPLETRNIAF		
HeLa $\alpha$	-----T--PD--N-----	72%	
Rat $\alpha$	-----T--PD--N-----	72%	
Rat $\alpha(+)$	-----T--P-----	75%	
Rat $\alpha$ III	-----T--PDC--D-----	66%	

FIG. 4. Amino acid sequence analysis. Amino acid sequences were deduced from the nucleotide sequences shown in Fig. 3. Identical residues are indicated by dashes; x refers to an unidentified amino acid (other residues are represented by standard one-letter abbreviations). Percent identity of clone sequences to the HeLa  $\alpha$ -subunit sequences (data from ref. 12) and rat brain  $\alpha$ ,  $\alpha(+)$ , and  $\alpha$ III sequences (data from ref. 7) are shown on the right. Stars indicate that the unidentified amino acid was not considered when calculating the percent identities.

suggests that there may be a Na<sup>+</sup>-dependent, K<sup>+</sup>-independent ATPase expressed in mammalian kidney (27).

Although it is known that the Na<sup>+</sup>,K<sup>+</sup>-ATPase is regulated developmentally (5, 28) and by hormones (29–34) and changes in the ionic environment of the cell (35, 36), little is known about the genetic basis of this regulation. Characterization of the genes encoding multiple isoforms of the catalytic subunit will facilitate studies of the molecular mechanisms involved in tissue-specific, developmental, and hormonal regulation of expression of this important enzyme.

We thank Gary Shull for valuable advice and discussions, Michael Hughes and Diana Pugh for technical assistance, Karen Wohlfeil and Becky Howland for typing the manuscript, and Greg Wernke and Gwen Kraft for preparing the figures. This work was supported in part by National Institutes of Health Grant HL28573 and National Science Foundation Instrument Grant DMB8414251. M.M.S. is a predoctoral fellow of the Albert J. Ryan Foundation.

1. Stahl, W. L. (1986) *Neurochem. Int.* **8**, 449–476.

2. Sweadner, K. J. (1979) *J. Biol. Chem.* **254**, 6060–6067.  
 3. Lytton, J., Lin, J. C. & Guidotti, G. (1985) *J. Biol. Chem.* **260**, 1177–1184.  
 4. Matsuda, T., Iwata, H. & Cooper, J. R. (1984) *J. Biol. Chem.* **259**, 3858–3863.  
 5. Schmitt, C. A. & McDonough, A. A. (1986) *J. Biol. Chem.* **261**, 10439–10444.  
 6. Lytton, J. (1985) *J. Biol. Chem.* **260**, 10075–10080.  
 7. Shull, G. E., Greb, J. & Lingrel, J. B (1986) *Biochemistry* **25**, 8125–8132.  
 8. Shull, G. E., Schwartz, A. & Lingrel, J. B (1985) *Nature (London)* **316**, 691–695.  
 9. Kawakami, K., Noguchi, S., Noda, M., Takahashi, H., Ohta T., Kawamura, M., Nojima, H., Nagano, K., Hirose, T., Inayama, S., Hayashida, H., Miyata, T. & Numa, S. (1985) *Nature (London)* **316**, 733–736.  
 10. Ovchinnikov, Y. A., Modyanov, N. N., Broude, N. E., Petrukhin, K. E., Grishin, A. V., Arzamazova, N. M., Aldanova, N. A., Monastyrskaya, G. S. & Sverdlov, E. D. (1986) *FEBS Lett.* **201**, 237–245.  
 11. Schneider, J. W., Mercer, R. W., Caplan, M., Emanuel, J. R., Sweadner, K. J., Benz, E. J. & Levenson, R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 6357–6361.  
 12. Kawakami, K., Ohta, T., Nojima, H. & Nagano, K. (1986) *J. Biochem.* **100**, 389–397.  
 13. Bell, G. I., Karam, J. H. & Rutter, W. J. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 5759–5763.  
 14. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).  
 15. Frischauf, A. M., Lehrach, H., Poustka, A. & Murray, N. (1983) *J. Mol. Biol.* **170**, 827–842.  
 16. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.  
 17. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.  
 18. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.  
 19. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.  
 20. Norrander, J., Kempe, T. & Messing, J. (1983) *Gene* **26**, 101–106.  
 21. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.  
 22. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.  
 23. Keller, E. B. & Noon, W. A. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7417–7420.  
 24. Shull, G. E. & Lingrel, J. B (1986) *J. Biol. Chem.* **261**, 16788–16791.  
 25. MacLennan, D. H., Brandl, C. J., Korczak, B. & Green, N. M. (1985) *Nature (London)* **316**, 696–700.  
 26. Brandl, C. J., Green, N. M., Korczak, B. & MacLennan, D. H. (1986) *Cell* **44**, 597–607.  
 27. Proverbio, F., Proverbio, T. & Marin, R. (1986) *Biochim. Biophys. Acta* **858**, 202–205.  
 28. Vigne, P., Frelin, C. & Lazdunski, M. (1982) *J. Biol. Chem.* **257**, 5380–5384.  
 29. Lo, C., August, T. R., Liberman, U. A. & Edelman, I. S. (1976) *J. Biol. Chem.* **251**, 7826–7833.  
 30. Lin, M. H. & Akera, T. (1978) *J. Biol. Chem.* **253**, 723–726.  
 31. Chaudhury, S., Ismail-Beigi, F., Gick, G. G., Levenson, R. & Edelman, I. S. (1987) *Mol. Endocrinol.* **1**, 83–89.  
 32. Geering, K., Girardet, M., Bron, C., Kraehenbühl, J. & Rossier, B. C. (1982) *J. Biol. Chem.* **257**, 10338–10343.  
 33. Garg, L. C., Knepper, M. A. & Burg, M. B. (1981) *Am. J. Physiol.* **240**, F536–F544.  
 34. Harik, S. I. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4067–4070.  
 35. Kim, D., Marsh, J. D., Barry, W. H. & Smith, T. W. (1984) *Circ. Res.* **55**, 39–48.  
 36. Wolitzky, B. A. & Fambrough, D. M. (1986) *J. Biol. Chem.* **261**, 9990–9999.