THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 16, Issue of April 21, pp. 11880–11890, 2000

Printed in U.S.A.

The Mouse Na^+-Sulfate Cotransporter Gene *Nas1*

CLONING, TISSUE DISTRIBUTION, GENE STRUCTURE, CHROMOSOMAL ASSIGNMENT, AND TRANSCRIPTIONAL REGULATION BY VITAMIN D*

(Received for publication, January 3, 2000)

Laurent Beck‡ and Daniel Markovich§

From the Department of Physiology and Pharmacology, the University of Queensland, Brisbane, Queensland 4072, Australia

NaSi-1 is a Na^+-sulfate cotransporter expressed on the apical membrane of the renal proximal tubule and plays an important role in sulfate reabsorption. To understand the molecular mechanisms that mediate the regulation of NaSi-1, we have isolated and characterized the mouse NaSi-1 cDNA (mNaSi-1), gene (*Nas1*), and promoter region and determined *Nas1* chromosomal localization. The mNaSi-1 cDNA encodes a protein of 594 amino acids with 13 putative transmembrane segments, inducing high affinity Na^+-dependent transport of sulfate in *Xenopus* oocytes. Three different mNaSi-1 transcripts derived from alternative polyadenylation and splicing were identified in kidney and intestine. The *Nas1* gene is a single copy gene comprising 15 exons spread over 75 kilobase pairs that maps to mouse chromosome 6. Transcription initiation occurs from a single site, 29 base pairs downstream to a TATA box-like sequence. The promoter is AT-rich (61%), contains a number of well characterized *cis*-acting elements, and can drive basal transcriptional activity in opossum kidney cells but not in COS-1 or NIH3T3 cells. We demonstrated that 1,25-dihydroxyvitamin D_3 stimulated the transcriptional activity of the *Nas1* promoter in transiently transfected opossum kidney cells. This study represents the first characterization of the genomic organization of a Na^+-sulfate cotransporter gene. It also provides the basis for a detailed analysis of *Nas1* gene regulation and the tools required for assessing *Nas1* role in sulfate homeostasis using targeted gene manipulation in mice.

Sulfate is the fourth most abundant anion in mammalian plasma, is present in nearly all cell types, and is essential for a variety of metabolic and cellular processes (1). The largest group of sulfoconjugates in mammals is sulfated proteoglycans, which are required for normal structure and function of bone and cartilage. Accordingly, three human congenital chondrodysplasias were recently shown to be caused by mutations in a sulfate transport protein gene (DTDST),[1] leading to undersulfation of proteoglycans in the extracellular matrix of bone and cartilage, and associated developmental abnormalities (2–4). Considering the importance of sulfate at a cellular and biochemical level, it is likely that mechanisms regulating the serum sulfate levels are essential for the maintenance of normal physiology. However, little is established about the molecular factors that regulate sulfate homeostasis, and the physiological consequence of a disturbance in sulfate homeostasis is mostly unknown.

In mammals, the regulation of sulfate homeostasis is largely determined by the kidney with the majority of the filtered sulfate load reabsorbed in the proximal segment of the nephron. Transepithelial transport of sulfate from the renal lumen to the blood compartment involves entry through the brush-border membrane (BBM) by a Na^+-dependent transport system, translocation across the cell and efflux across the basolateral membrane by an anion exchange system (5). Early transport studies in BBM vesicles suggested that Na^+-sulfate cotransport across the BBM is the rate-limiting step in the overall sulfate reabsorptive process (6, 7). By expression cloning, we isolated a cDNA (NaSi-1) from rat kidney encoding a high affinity Na^+-dependent sulfate transporter (8). NaSi-1 mRNA is expressed in kidney and small intestine and encodes a glycosylated protein (8) that has been localized by immunohistochemistry to the BBM of proximal tubular cells (9).

Recently, factors known to regulate renal Na^+-sulfate reabsorption were found to regulate NaSi-1 expression in the kidney. Vitamin D was shown to modulate concomitantly serum sulfate concentration, renal sulfate handling, and the expression (mRNA and protein levels) and activity of the NaSi-1 cotransporter (10). High sulfate intake in rats led to a reduction in both NaSi-1 mRNA and protein (11), whereas low sulfate intake (reduced methionine diet) led to an increase in both NaSi-1 mRNA and protein (12). Thyroid hormone, growth hormone, heavy metals, potassium intake, and anti-inflammatory agents were also found to regulate NaSi-1 expression (13–17). It is suggested that these modulators could alter serum sulfate levels via the regulation of NaSi-1 expression *in vivo*, suggesting that sulfate homeostasis is controlled, at least in part, by NaSi-1. However, the underlying mechanisms involved in the regulation of NaSi-1 expression by these factors, as well as NaSi-1 contribution to body sulfate homeostasis, have yet to be defined.

* This work was supported in part by the National Health and Medical Research Council of Australia (to D. M.). 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) AF199365, AF199366, AF199380, and AF200305–AF200319.

‡ Recipient of a University of Queensland postdoctoral research fellowship.

§ To whom correspondence should be addressed: Dept. of Physiology and Pharmacology, the University of Queensland, Brisbane, Queensland 4072, Australia. Tel.: 61 7 3365 1400; Fax: 61 7 3365 1766; E-mail: danielm@plpk.uq.edu.au.

[1] The abbreviations used are: DTDST, diastrophic dysplasia sulfate transporter; BBM, brush-border membrane; RACE, rapid amplification of cDNA ends; UTR, untranslated region; OK, opossum kidney; RT-PCR, reverse transcriptase-polymerase chain reaction; LA-PCR; long and accurate PCR; 1,25-$(OH)_2D_3$, 1α,25-dihydroxyvitamin D_3; DR, direct repeat; VDR, vitamin D receptor; VDRE, vitamin D-responsive element; hRXRα, human retinoid X receptor α; TRE, thyroid hormone-responsive elements; GRE, glucocorticoid-responsive element; EST, expressed sequence tag; bp, base pair; kb, kilobase pair; PIPES, 1,4-piperazinediethanesulfonic acid.

TABLE I
Oligonucleotides

Name	Sequence[a]	Position[b]	Strand
Sense *Nas*1-specific primers			
FN-24	TGTTGAAGGCACCTGCTCAGG	−24	Sense
FN17	ATGCTTTGGTCTATCGCCGCTTTC	17	Sense
FN110	GTGCCTACATCCTCTTTGTTATTG	110	Sense
FN252	CTTTCACCTTCTGCTAATTGGA	252	Sense
FN390	CACTGCCTTCTTATCTATGTGG	390	Sense
FN665	CAGTCACAGGAGCAAAATATCGG	665	Sense
FN823	GAATGTCGCTGCCTCCACTTTGG	823	Sense
FN1386	TCCTCTAGGTTCATTACCAGTTTG	1386	Sense
FN1547	TGCCTTCCACTCTCTGTACCTCA	1547	Sense
Antisense *Nas*1-specific primers			
RN44	AGGAGAAAGCGGCGATAGAC	44	Antisense
RN88	TGATGAGAGGGAGTGGCAAGA	88	Antisense
RN137	ATGGCAATAACAAAGAGGATGT	137	Antisense
RN226	GTGAAGAACGCATGATCCCAAA	226	Antisense
RN308	TTCCATTTCTCTATTGATGTTGCT	308	Antisense
RN535	GGGCGGCAGATTCATTGAAATA	535	Antisense
RN687	CCGATATTTTGCTCCTGTGACTG	687	Antisense
RN909	GAGCCAAATCCAAGACAAAAGTAG	909	Antisense
RN1020	CCCAAGTTTTTCATATTCTT	1020	Antisense
RN1220	GTCATTTTTGTCAGTTTCTTGGC	1220	Antisense
RN1647	CATGTCAATGACTTTCAGGTGG	1647	Antisense
RN1989	AGGCGGGTAGATGCTCTTTGATTG	1989	Antisense

[a] Primers were designed either from the rat or the mouse NaSi-1 cDNA sequences and are written from the 5′ to 3′ direction.
[b] The number indicated refers to the nucleotide position within the mNaSi-1 cDNA sequence (A of ATG initiation codon defined as +1) where the 5′ end of the primer will anneal.

In order to provide insights into the molecular mechanisms underlying tissue-specific and hormonal regulation of NaSi-1 and its role in sulfate homeostasis, we have cloned and characterized the mouse NaSi-1 cDNA and its corresponding gene. This study represents the first characterization of the genomic structure of a Na$^+$-coupled sulfate transporter gene. We have also determined the pattern of NaSi-1 expression in mouse adult tissues, identified the existence of alternative transcripts, determined its chromosomal localization, and demonstrated that the transcriptional activity of the promoter region is elevated in response to 1,25-$(OH)_2D_3$ stimulation in a transiently transfected renal cell line.

EXPERIMENTAL PROCEDURES

PCR Amplifications and Sequencing—Oligonucleotides used during this study are listed in Table I. The mouse NaSi-1 (designated mNaSi-1) cDNA coding sequence was cloned using RT-PCR. Total RNA (5 μg) isolated from mouse kidney cortex was reverse-transcribed and PCR-amplified using primers derived from the rat NaSi-1 cDNA (designated rNaSi-1) coding sequence (8). The PCR products were subcloned into pCR 2.1 vector (Invitrogen) and sequenced in both directions. Semi-quantitative PCR amplification of mNaSi-1 was performed by comparing its abundance to β-actin. Preliminary optimization of the conditions showed that coamplification of mNaSi-1 and β-actin transcripts was occurring linearly through cycles 23–30. Total RNA treated with RNase-free DNase I enzyme was reverse-transcribed, and PCR amplification (30 cycles) was performed using 0.4 μM mNaSi-1 primers (FN252 and RN1220) and 0.1 μM β-actin primers. The identity of the mNaSi-1 PCR-amplified products in each tissue was confirmed by restriction enzyme digestion. Fluorescence from ethidium bromide staining of each mNaSi-1 signal was compared with that of β-actin, by calculating the ratio (fluorescence units of mNaSi-1/fluorescence units of β-actin). Dye-termination sequencing was performed using the Big Dye™ Termination kit (Perkin-Elmer) following the manufacturer's protocol, and gel separation was performed at the Australian Genome Research Facility, the University of Queensland.

5′- and 3′-RACE—The 5′- and 3′ end of mNaSi-1 cDNA were isolated using 5′- and 3′-RACE techniques, respectively, essentially as described by Chen (18). For 5′-RACE, primer RN137 was used to reverse-transcribe mouse kidney total RNA and for a first round of PCR amplification. After nested amplification using primer RN88, PCR products were obtained and subcloned into pCR 2.1 vector. For 3′-RACE, kidney total RNA was reverse-transcribed using SuperScript II (Life Technologies, Inc.) and an oligo(dT)/adapter primer. PCR amplification using *Taq* DNA polymerase (Biotech International) was carried out using primer FN1547. The 3′-RACE technique was also used for identifying mNaSi-1 variants. In this case, total RNA from mouse kidney was reverse-transcribed using display THERMO-RT™ (Display Systems Biotech) reverse transcriptase, and PCR amplifications were performed using a 16:1 blend of *Taq* and *Pfu* (Promega) DNA polymerases with primers FN-24, FN110, FN823, or FN1547. The 5′-RACE technique was also used to confirm the position of the transcription start site (see below). In this case, primer RN535 was used for reverse transcription and first round of PCR, and primer RN226 was used for nested amplification. In addition, the display THERMO-RT™ reverse transcriptase was used to permit the utilization of high temperatures (42 °C for 40 min and 65 °C for 15 °C) avoiding an artificial termination due to the secondary structure of mRNA.

Xenopus laevis Oocytes and Transport Assays—Methods for handling of oocytes, *in vitro* transcription, and transport assay have been described previously (8, 19). Stages V and VI oocytes were injected with either 50 nl of water (control) or 5 ng of mNaSi-1 cRNA using a Nanoject automatic oocyte injector (Drummond Scientific Co.).

Northern and Southern Blot Analyses—Northern analysis of total RNA (25 μg) from mouse tissues (Fig. 5) was performed as described previously (20). Full-length mNaSi-1 cDNA was ^{32}P-labeled by random priming and used as a probe. After stripping, the membranes were rehybridized with a ^{32}P-labeled 1.3-kb mouse β-actin cDNA probe. Mouse genomic DNA (10 μg) prepared from mouse liver was digested with restriction enzymes (Fig. 3), separated on 0.7% agarose gels, and capillary transferred to positively charged nylon membranes (Hybond XL, Amersham Pharmacia Biotech). After UV-cross linking, the membranes were hybridized (16–18 h at 65 °C) with a full-length ^{32}P-labeled mNaSi-1 cDNA probe in Church's buffer (0.5 M Na_2HPO_4/NaH_2PO_4, pH 7.2, 7% SDS, 10 mM EDTA). The membranes were then washed to high stringency and exposed to Kodak X-Omat AR5 film at −80 °C for 48 h.

Isolation and Characterization of Mouse Genomic Nas1 Clones—The genomic clones were isolated from a λ FIX II mouse (129sv strain) genomic DNA library (Stratagene) using the method described by Lardelli and Lendahl (21). Five positive λ clones were purified and further analyzed. Some large introns were isolated from mouse genomic DNA using LA-PCR, as described elsewhere (22). Introns 1, 2, and 6–8 were amplified using primer pairs FN17/RN226, FN110/RN308, FN390/RN687, FN665/RN909, and FN823/RN1020, respectively. Identity and further characterization of the λ clones and PCR products were confirmed by Southern analysis and/or direct sequencing of the coding regions. Exon sizes were determined by nucleotide sequencing, and intron sizes were determined by either nucleotide sequencing or estimated from the size of corresponding PCR-generated DNA fragments using exon-specific primers. Location of and sequences at intron/exon boundaries of the *Nas1* gene were determined by direct sequencing using *Nas1*-specific oligonucleotides.

Primer Extension Analysis—Primer extension analysis was performed using protocols and reagents provided by Promega (primer extension system). Briefly, two *Nas1*-specific primers located in exon 1 (RN44 and RN88) were end-labeled using T4 polynucleotide kinase and [γ-^{32}P]dATP. Total RNA (10 μg) was mixed with 0.1 pmol of the labeled primer, denatured 5 min at 90 °C, and incubated at 55 °C for 16 h in hybridization buffer (0.4 M NaCl, 1 mM EDTA, 40 mM PIPES, pH 6.4, 80% formamide). After reverse transcription using avian myeloblastosis virus-reverse transcriptase, the labeled cDNAs were separated through a 6% polyacrylamide gel. *Hin*fI-ϕx174-digested DNA was end-labeled and used as a molecular weight marker. The gel was dried and exposed to Kodak X-Omat AR5 film for 48 h at −80 °C.

Chromosomal Localization by Radiation Hybrid—Fine mapping of *Nas1* was undertaken using the T-31 Radiation Hybrid Panel of the mouse genome (Research Genetics). Primers used for screening were FN1386 and RN1647, generating an intense 1.5-kb band from mouse genomic DNA and a faint 950-bp band from Chinese hamster DNA. The panel was screened twice using these primers and a third time using primers FN1547 and RN1989. Data analysis was performed by the Jackson Laboratory Mouse Radiation Hybrid Data Base.

Plasmid Construction, Cell Culture, and Transient Transfections—Three fragments containing 3229, 1203, and 457 bp of *Nas1* 5′-flanking sequence, respectively, were PCR-amplified from the λP2 clone (Fig. 4*A*), subcloned into pCR2.1 vector, and sequenced. These fragments were then inserted upstream of a luciferase reporter gene in a promoterless luciferase expression vector (pGL3-Basic, Promega) by restriction enzyme digestion and ligation. Plasmids were designated p*Nas1*-3229, p*Nas1*-1203, and p*Nas1*-457, respectively. The 1203-bp promoter fragment was also cloned in reverse orientation and designated p*Nas1*-1203R. Correct insertion and sequence were verified by enzyme restriction digestion and sequencing. COS-1 and NIH3T3 cells were cultured in Dulbecco's modified Eagle's medium (Life Technologies, Inc.) with 10% (v/v) fetal bovine serum (Life Technologies, Inc.). OK cells were maintained in Ham's F12/Dulbecco's modified Eagle's medium (1:1) containing 10% fetal bovine serum. At 80% confluence, cells were cotransfected using LipofectAMINE™ 2000 reagent (Life Technologies, Inc.), with 0.8 μg of the *Nas1* gene promoter-luciferase reporter plasmid and 0.8 μg of pRSVβGal plasmid (gift of Dr. M. Waters, University of Queensland) as an internal control for transfection efficiency. Incubation with plasmids and LipofectAMINE was carried out for 24 h in normal growth medium, as recommended by the manufacturer. Controls were performed by transfection with pGL3-Basic (promoter-less plasmid) and pGL3-Control (containing the SV40 promoter). In experiments involving vitamin D, cells were cotransfected with 0.2 μg of a VDR expression vector alone or together with 0.2 μg of a human RXRα expression vector (VDR/pSG5 and RXR/pSG5 plasmids, respectively; generous gift from Dr. John White, McGill University). The VDR and hRXRα expression plasmids were cotransfected to ensure that a sufficient concentration of receptor was available for binding to the overexpressed *Nas1* gene. Incubation with plasmids and LipofectAMINE was carried out for 24 h, after which the medium was replaced by fresh Dulbecco's modified Eagle's medium containing 10% fetal bovine serum with varying concentrations of 1,25-$(OH)_2D_3$ for an additional 24 h. Cells were then harvested in cell lysis buffer, and the lysate was assayed for luciferase and β-galactosidase activities using protocols and reagents provided by Roche Molecular Biochemicals. Luciferase activity was measured using a Trilux 1450 Microbeta (Wallac) luminometer.

Data Presentation and Statistics—Data are shown as means ± S.D. Statistical significance was determined by unpaired Student's t test, with $p < 0.05$ considered significant. For the transport kinetic studies in oocytes, the Michaelis-Menten and generalized Hill equations were used to calculate K_m and V_{max} values using non-linear regression.

RESULTS

Mouse NaSi-1 cDNA Cloning and Expression—The mouse Na$^+$-sulfate cotransporter cDNA, mNaSi-1, was cloned using a combination of RT-PCR and 5′- and 3′-RACE techniques. The mNaSi-1 cDNA is 2246 bp long, with 28 bases of 5′-UTR, an open reading frame of 1782 bases, and 436 bases of 3′-UTR. The 3′-UTR contains a polyadenylation signal (AATAAA) at position 2180. The open reading frame encodes a protein of 594 amino acids (Fig. 1*A*) with a calculated molecular mass of 66.1 kDa, containing 13 putative transmembrane domains, predicted by the TopPred2 program (23). The mNaSi-1 protein contains one potential protein kinase A (Thr^{404}) and five potential protein kinase C (Ser^{213}, Thr^{218}, Ser^{230}, Thr^{322}, and Thr^{422}) phosphorylation sites (Fig. 1*A*). Consensus sequences for *N*-glycosylation were found at Asn positions 140, 174, and 590 (Fig. 1*A*). Alignment of the mouse and rat NaSi-1 amino acid sequences shows 93.6% identity and 96% similarity. Nucleotide sequence identity is 91% between mouse and rat NaSi-1. When this work was initiated, no ESTs with homology to mNaSi-1 were identified. At the submission of this manuscript, a search in the EST data base identified approximately 50 murine ESTs from kidney, all identical to mNaSi-1. Homology searches using BLAST (24) and PSI-BLAST (25) revealed significant homology to 22 other proteins (Fig. 1*B*), although the closest relatives are the recently reported human Na$^+$-sulfate cotransporter SUT-1 (49% identity (26)) and the Na$^+$-dicarboxylate transporters sharing ~32–43% protein sequence identity with mNaSi-1. Of particular interest is a consensus pattern previously established for Na$^+$-coupled symporters (PROSITE PS01271) known as the Na$^+$-sulfate signature, present at amino acids 522–538 in the mNaSi-1 protein containing a very high degree of homology with other related proteins (Fig. 1*B*).

To determine the functionality of the isolated clone, we injected mNaSi-1 cRNA into *Xenopus* oocytes followed by [^{35}S]sulfate radiotracer uptake assay. Sulfate uptake in mNaSi-1-cRNA-injected oocytes was Na$^+$-dependent and showed typical Michaelis-Menten saturation, with a calculated K_m value for sulfate of 0.20 ± 0.06 mM and V_{max} of 49.2 ± 4.1 pmol/h (data not shown), in agreement with the BBM Na$^+$-sulfate cotransporter (5). mNaSi-1 mRNA expression was screened by RT-PCR in 23 murine tissues (Fig. 2). An amplified mNaSi-1 fragment was obtained in RNA from kidney, duodenum/jejunum, ileum, and colon. Lower levels of mNaSi-1 mRNA expression were observed in cecum, testis, adrenal, and adipose tissue. By normalizing the mNaSi-1 mRNA signal to β-actin, the relative abundance of mNaSi-1 in kidney and ileum was found to be similar and approximately twice as high as those found in duodenum/jejunum and colon (Fig. 2*B*; $n = 3$). The physiological importance of the low level of expression of mNaSi-1 found in testis, adrenal, and adipose tissue remains to be determined.

Genomic Organization of the Mouse Na$^+$-Sulfate Cotransporter Gene, Nas1—Southern blotting was used to estimate the size and complexity of the gene encoding mNaSi-1, designated *Nas1* (Fig. 3). Results show that the estimated size of the *Nas1* gene was approximately 45 kb, which was lower than the actual size of the *Nas1* gene determined from genomic cloning (see below). This was due both to the presence of comigrating bands and large introns that did not hybridize with the cDNA probe. Blots washed at both high and low stringency gave similar results, suggesting that *Nas1* is a single copy gene.

Screening of a genomic λ phage library led to the isolation of five λ *Nas1* clones containing the 5′-flanking region and most of the *Nas1*-coding region (Fig. 4*A*). Introns 1, 2, and 6–8 not present in the λ clones were obtained using LA-PCR (Fig. 4*A*). The λ clones and PCR-amplified introns overlapped, covering a 80-kb region comprising the entire *Nas1* gene (Fig. 4*A*). Southern analysis of the *Nas1* genomic clones was consistent with the data obtained from Southern blotting of mouse genomic DNA and confirmed that *Nas1* is a single copy gene. The resulting exon-intron organization of the mouse *Nas1* gene is shown in Fig. 4*B*. The *Nas1* spans ~75 kb and contains 15 exons. The translation initiation site is present in exon 1. Exon sizes range from 49 to 188 bp, except for exon 15, which is 555 bp and contains the TGA stop codon (Fig. 4*C* and Table II). Intron sizes range from 70 bp to 15 kb (Table II). All exon-intron boundaries conform to canonical splice donor and acceptor consensus sequences, and the codon phase usage is mainly

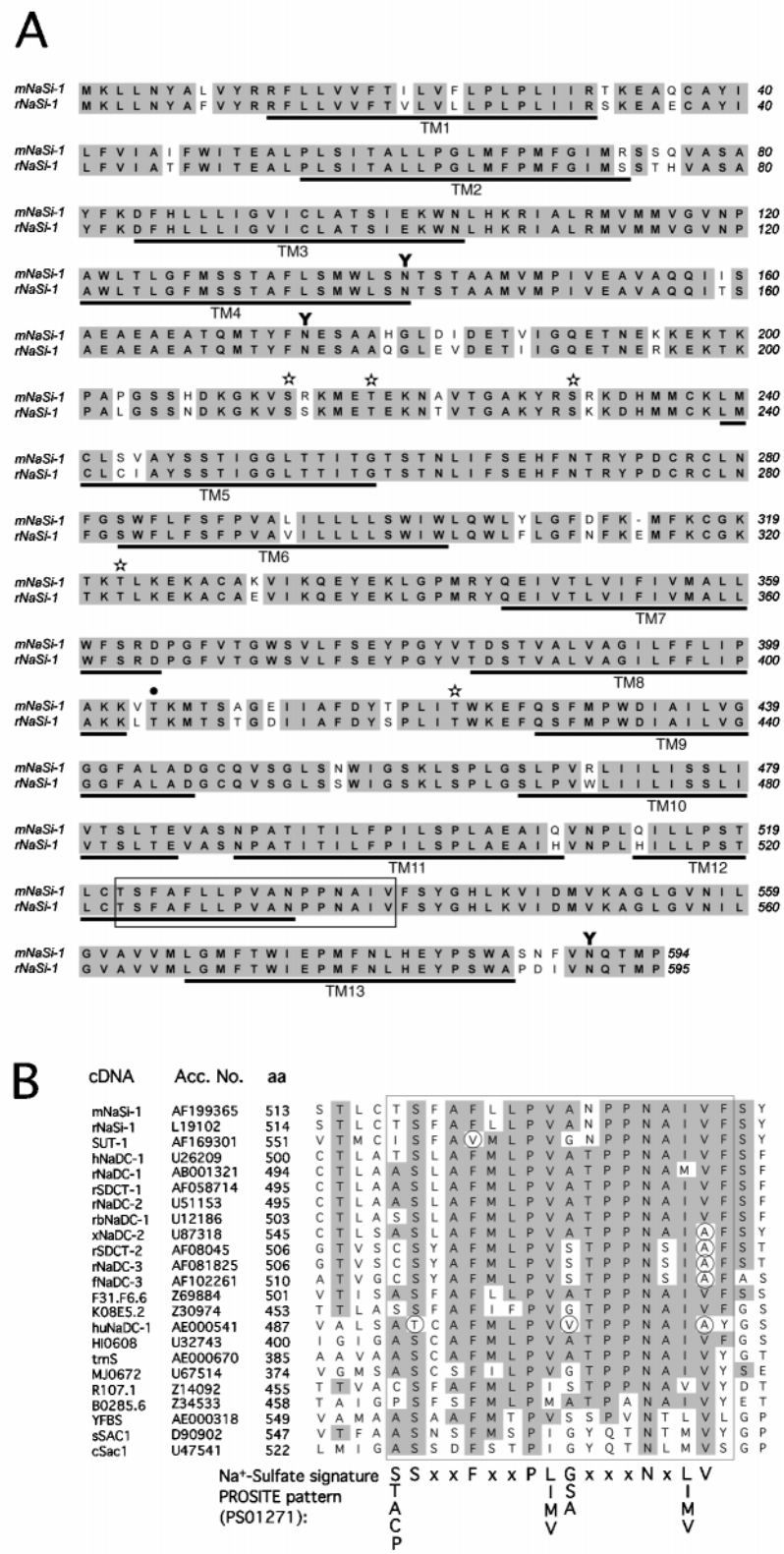

FIG. 1. **Sequence alignments of mNaSi-1.** *A*, comparison of the deduced amino acid sequences of mNaSi-1 and rNaSi-1 (8). Amino acids that are identical are depicted with *gray shading*. Putative transmembrane domains (*TM1–TM13*) are *underlined*. Potential phosphorylation (☆, protein kinase C; ●, protein kinase A) and *N*-glycosylation (***Y***) sites are labeled. The Na$^+$-sulfate signature is *boxed* (see *B*). *B*, multiple sequence alignment of the Na$^+$-sulfate signature PROSITE pattern among mNaSi-1 homologues, with GenBank^{TM} accession numbers indicated. The PROSITE consensus pattern is indicated at the *bottom* of the alignment; amino acids that differ from this consensus sequence are *circled*. Sequence alignments were made using the ClustalW and MacVector programs.

0 or II (Table II). Comparison of predicted protein transmembrane domains to exon border structure showed that each predicted transmembrane segment is encoded by a separate exon, with the exception of transmembrane domains 10 and 11, which are encoded by the same exon (exon 13). In addition, splicing mostly occurred near membrane/aqueous transitions (Fig. 4*D*).

Mapping of Nas1 to Mouse Chromosome 6—Nas1 was mapped by analysis of the data from the T-31 Radiation Hybrid panel in The Jackson Laboratory Mouse Radiation Hybrid data base. The data placed *Nas1* on mouse chromosome 6 in the most likely position between marker D6Mit170 (LOD score 20.8) and D6Mit380 (LOD score 11.2). *Nas1* is 2.3 centi-rays distal to D6Mit170, which has been assigned map positions of 4.4 centimorgans (MIT) and 4.0 centimorgans (MGD and Chromosome Committee). The *Nas1* gene maps very close to the calcitonin receptor gene, which has been assigned a position of 4.5 centimorgans (MGD and Chromosome Committee) in

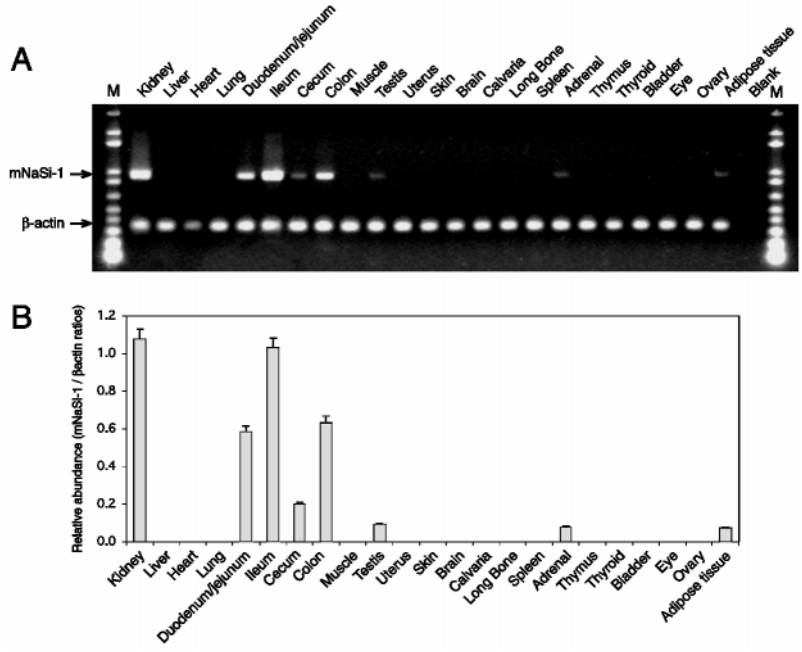

FIG. 2. **Tissue distribution of mouse Na$^+$-sulfate cotransporter.** *A*, total RNA from various mouse tissues (indicated at the *top*) was reverse-transcribed using an oligo(dT) primer and PCR-amplified using mNaSi-1 (FN252 and RN1220, *upper bands*) and β-actin (*lower bands*) primers. An aliquot of each PCR reaction was electrophoresed on 1.2% agarose gels and visualized with ethidium bromide. A water blank is shown. *B*, densitometric analysis of mNaSi-1 RT-PCR-amplified mRNA derived from three separate experiments. Data are shown as mean ± S.D.

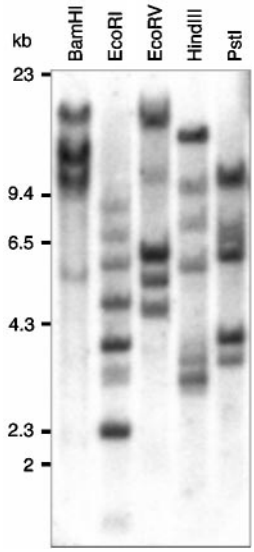

FIG. 3. **Southern blot analysis of mouse genomic DNA.** Mouse liver DNA (10 μg) was digested with *Bam*HI, *Eco*RI, *Eco*RV, *Hin*dIII, or *Pst*I, as indicated, electrophoresed on a 0.7% agarose gel, transferred to a nylon membrane, and hybridized with a full-length ^{32}P-labeled mNaSi-1 cDNA probe.

mouse and 7q21.3-q21.3 in human.

Alternative Polyadenylation and Splicing of the Nas1 Gene—By using Northern blot analysis, two major transcripts of equal intensity (2.2 and 2.5 kb) were detected in kidney, ileum, duodenum/jejunum, and colon but not in liver (Fig. 5*A*), confirming the RT-PCR data (Fig. 2). Normalizing the mNaSi-1 mRNA signal with β-actin showed that renal and ileal mNaSi-1 transcripts were of similar abundance, in agreement with RT-PCR data (Fig. 2). Whereas the 2.2-kb transcript most probably corresponds to the 2246-bp cDNA fragment characterized above, the larger transcript is most likely derived from alternative polyadenylation, as previously shown for the rNaSi-1 transcripts (8, 27). To test this, we performed 3′-RACE on total RNA from mouse kidney (Fig. 5*B*). Two bands were obtained using primers FN110, FN823, and FN1547 (Fig. 5*B, lanes 2–4*). Sequencing analysis showed that the smaller band corresponded to the 2246-bp mNaSi-1 cDNA, whereas the larger band contained an additional 254 nucleotides at the 3′ end, generating a 2500-bp cDNA fragment containing a polyadenylation signal at position 2437. Sequence comparison between the 2.5-kb clone and the genomic λ3 clone showed absence of introns in the 3′-UTR. In addition to the 2.2- and 2.5-kb clones, we could also amplify an additional faint band at 2.1 kb using a sense primer further upstream (Fig. 5*B, lane 1*). This fragment was 100% identical to the 2.2-kb cDNA, with the exception of a missing 129-bp region, corresponding to exon 2 sequence (Fig. 5*C*). To determine the functionality of these clones, we injected the corresponding cRNAs into *Xenopus* oocytes. Both the 2.2- and 2.5-kb clones induced significant Na$^+$-sulfate cotransport (at a similar rate), whereas the 2.1-kb clone showed an activity comparable to water-injected oocytes (Fig. 5*D*).

Transcription Initiation Site and Nucleotide Sequence of the Nas1 5′-Flanking Region—To identify the transcription start site, primer extension assays were performed. A single major product of 70 and 114-nucleotides was identified using primers RN44 and RN88, respectively (Fig. 6*A*). This located the transcription start site, designated +1, at 26 bp upstream from the translation initiation ATG codon (Fig. 6*C*). To confirm these data, we performed 5′-RACE, which gave rise to a 277-bp band (Fig. 6*B*). Sequence analysis of this fragment confirmed the primer extension findings, locating the transcription start site 28-bp upstream from the ATG codon (Fig. 6*C*).

A 3229-bp region of the *Nas1* promoter has been isolated and sequenced (Fig. 7). This region was found to be A + T-rich (61%). The center of an atypical TATA sequence, TATTTAA, is located 29 bp upstream of the transcription start site. A classical CAAT box consensus sequence is present at position −91 on the negative strand. However, a canonical TATA box and CAAT box are positioned further upstream at −201 and −424, respectively, but are considered too far from the transcription start site to have promoter function. No GC box motif or Sp1-binding sites were detected. A repeated GA-rich region of unknown function was found from position −218 to −410. The *Nas1* promoter contains a number of potential *cis*-acting elements recognized by well characterized transcription factors that may play a role in the basal or chronic regulation of the *Nas1* gene. These include two AP-1 sites, one AP-2 site, one AP-4 site, two CAAT/enhancer-binding protein (C/EBP) binding sites, three Oct-1 sites, three NF-Y sites, and two NFAT

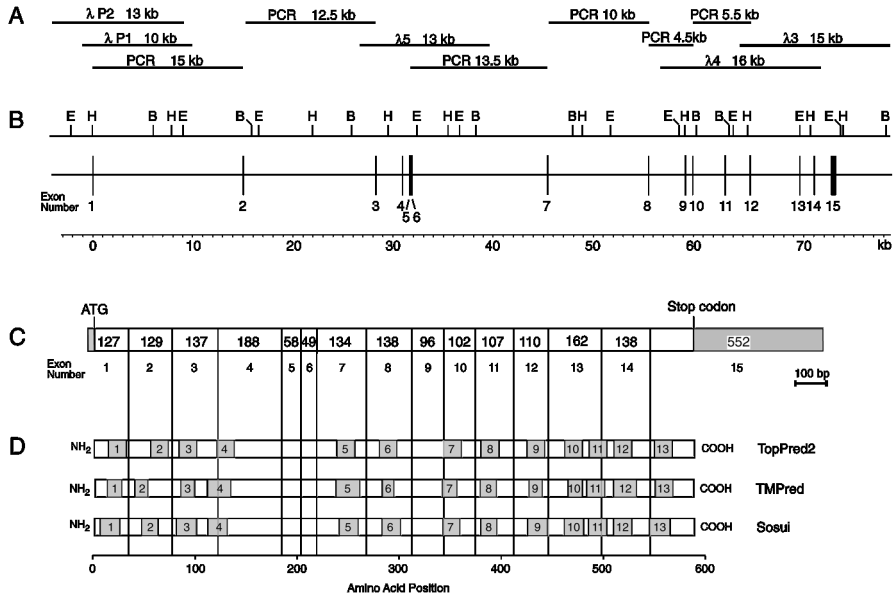

FIG. 4. **Organization of the mouse Na$^+$-sulfate cotransporter *Nas1* gene.** *A*, genomic *Nas1* clones and LA-PCR products obtained are indicated. *B*, genomic structure of the mouse *Nas1* gene. *Top*, partial restriction map of *Nas1* gene. *E, Eco*RV; *H*, *Hin*dIII; *B*, *Bam*HI. *Middle, Nas1* gene organization. The *horizontal line* indicates gene introns, and *vertical lines* represent exons (*numbered 1–15*). The position of introns was determined by sequencing across exon/intron boundaries of genomic fragments and LA-PCR products. *Bottom*, scale in kb. *C*, exons 1–15 are represented by *boxes.* The *white portions* of the exons identify the protein-coding sequences, and the size of each exon is indicated in bp. The translation initiation site is present in exon 1. The *gray portions* of the exons represent the untranslated regions. *D*, a comparison of the predicted 13 transmembrane spanning domain models and the exon border structure. *Gray boxes* indicate predicted transmembrane domains of the protein. The programs used for predicting the transmembrane spanning domains were TopPred2 (23), TMPred (55), and Sosui (56).

TABLE II
Exon/intron organization of the mouse Nas1 gene

Intron number	Location	5′ splice donor[a]	Intron size[b]	3′ splice acceptor[a]	Amino acid[c]	Codon phase[d]	Exon number	Exon size
			bp					*bp*
1	99	ACCAAG/gtaagcaagc. . .	15,000	. . .ccctttgcag/GAAGCA	Lys	0	1	127
2	228	TCACAG/gtaacataat. . .	12,500	. . .gcaatttcag/GTGGCT	Gln	0	2	129
3	365	AGCCTG/gtgagtatta. . .	2,500	. . .ctcgtttcag/GCTGAC	Trp	II	3	137
4	553	TTGATG/gtatcatgta. . .	600	. . .tatcttacag/AAACTG	Glu	I	4	188
5	611	TCCAGG/gtaaagacta. . .	70	. . .tatgtttcag/AAGCAG	Gly	II	5	58
6	660	GAAAAG/gtacattaca. . .	13,500	. . .ttgattacag/AATGCA	Lys	0	6	49
7	794	GATCTT/gtataaagct. . .	10,000	. . .cttattgcag/CTCTGA	Phe	II	7	134
8	932	ATTCGA/gtaagtagac. . .	3,500	. . .tctctttcag/CTTTAA	Asp	II	8	138
9	1028	AATGAG/gttaaaattg. . .	560	. . .tcttccacag/GTATCA	Arg	II	9	96
10	1130	TTCAGA/gttgagtatc. . .	3,100	. . .tatttttcag/GTACCC	Glu	II	10	102
11	1237	AAATTA/gttgagtatc. . .	2,400	. . .tattttgcag/TTGCTT	Ile	I	11	107
12	1347	TGTCAG/gtaatgctca. . .	4,800	. . .tgtattgtag/GTATCA	Gln	0	12	110
13	1509	CCTTTG/gtgagtatga. . .	1,250	. . .tccatgccag/GCTGAA	Leu	0	13	162
14	1647	GACATG/gtaagtcagc. . .	1,500	. . .tgttactaag/GTTAAA	Met	0	14	138
							15	555

[a] Exon sequences are indicated by uppercase letters and intron sequences by lowercase letters.
[b] Intron size was determined by restriction analysis, direct sequencing for small introns, and by estimating the size of PCR amplification products for large introns.
[c] Amino acids encoded at the splice sites are indicated.
[d] Introns that do not split codon triplets are indicated by phase 0, interruption after the first nucleotide by phase I, and interruption after the second nucleotide by phase II.

sites. There were many GATA-1-binding sites, located at nucleotides −153, −400, −605, −682, −928, −948, −1222, −1348, −2088, and −3030. Consensus sequences for the binding of other transcription factors activated by mitogenic or differentiation signals (c-Ets-1, Sox-5, hepatic nuclear factor 4, upstream stimulating factor, FREAC4, and Pit-1) are also present (Fig. 7). However, the evaluation of many of these sites in relation to known *Nas1* functions will require additional studies.

Structure of Putative Steroid-Thyroid Hormone-responsive Elements in the Nas1 Promoter—Three regions, named A, B and C, containing direct repeat-like sequences similar to steroid-thyroid hormone-responsive elements were found (Fig. 7). Within the region A (−2549 to −2515), the sequence 5′-AGTTCAgaaTGTCCT-3′ bears strong resemblance to the consensus inverted repeat sequence (5′-AGGTCA—TGACCT-3′) of TREs. This same sequence also had similarities to the consensus core binding motif (5′-(A/G)G(G/T)TCA-3′) of VDREs as well as the mouse osteopontin VDRE (28). Within the same region, the sequences 5′-AGCTCActctgtAGTTCA-3′ and 5′-AGTTCAgaatgtCCTTGA-3′ show strong similarities with DR6-type and palindromic type consensus VDREs (29), respectively. Region B (−525 to −508) contains a consensus DR6-type structure (5′-GGTTCAtcaaaaGGGGCA-3′). Slightly downstream, region C (−496 to −468) contains a DR3-type structure (5′-GTGTGAacaAAGTCA-3′) with similarities to rat osteocalcin (30) and rat calbindin (31) VDREs, as well as a second DR3-type structure (5′-AGTTAAtttCATTCA-3′) with similarities to mouse os-

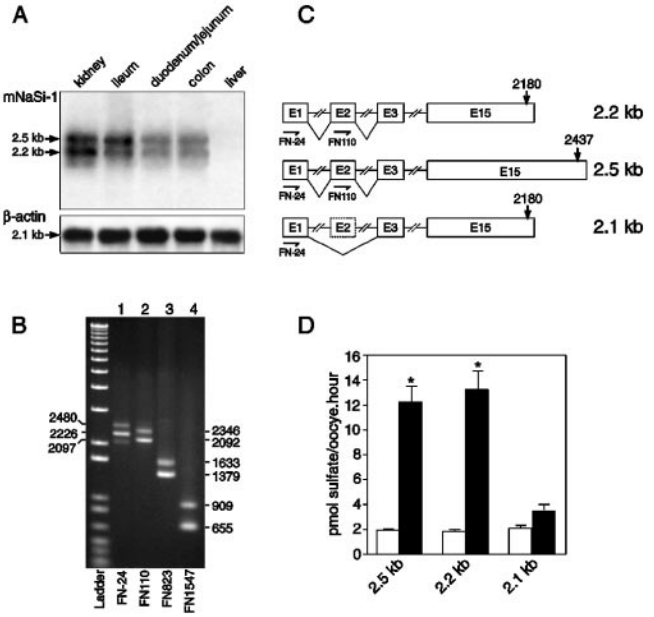

FIG. 5. **Nas1 transcription products.** *A*, Northern analysis of total kidney mRNA. *Top panel*, total RNA (25 μg) from mouse kidney, ileum, duodenum/jejunum, colon, and liver was hybridized to a full-length 2.2-kb mNaSi-1 cDNA probe. *Lower panel,* after stripping, the same membrane was hybridized to a 1.3-kb β-actin cDNA probe, as a control for RNA quality and loading. Exposure times were 24 h for mNaSi-1 probe and 1 h for β-actin probe. *B*, 3′-RACE. Total RNA was reverse-transcribed with an oligo(dT)/adapter primer and PCR-amplified using an antisense adapter primer and sense mNaSi-1-specific primer as indicated. *C*, structure of the *Nas1* transcription products. Two forms of mNaSi-1 mRNA were derived from alternative polyadenylation (*2.2* and *2.5 kb*), the third form was derived from alternative splicing and lacked exon 2 (*2.1 kb*). Exons (*E*) are indicated by *boxes. D*, mNaSi-1 cDNA variants were subcloned, *in vitro* transcribed, and either water (*open bars*) or cRNA (*filled bars*) corresponding to each transcript were injected into *Xenopus* oocytes (8–10 oocytes per condition). $^{35}SO_4^{2-}$ uptake was performed day 3 post-injection at room temperature for 30 min, in the presence of sodium. Data are shown as mean ± S.D. *, $p < 0.01$ when compared with water-injected oocytes.

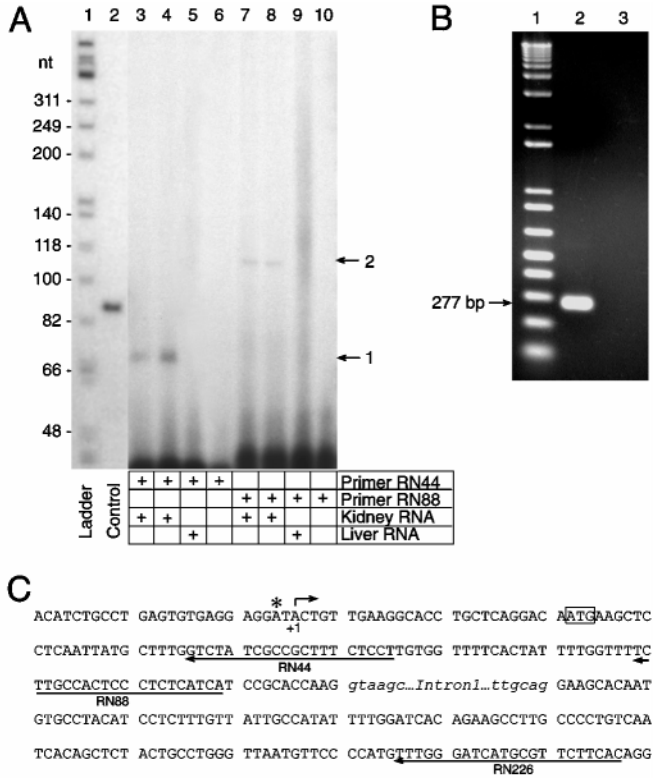

FIG. 6. **Mapping the transcription start site of *Nas1* using primer extension and 5′-RACE.** *A*, primers RN44 and RN88 (see *C*) were mixed with either no RNA (*lanes 6* and *10*) or with 10 μg of total RNA prepared from mouse kidney or liver, as indicated. A control primer extension reaction was also included. The primers were extended with avian myeloblastosis virus-reverse transcriptase at 42 °C for 30 min. The reaction products were size-fractionated on a 6% denaturing polyacrylamide gel, followed by exposure to film for 1.5 (*lanes 1* and *2*) or 48 h (*lanes 3–10*). The sizes of the primer extension products were determined by their migration relative to a molecular weight marker (*Hin*fI-digested ϕx174, labeled with [γ-^{32}P]dATP). *B*, total RNA (5 μg) isolated from mouse kidney was reverse-transcribed using primer RN535. After two rounds of PCR amplification with primer RN226 (see *C*), a 277-bp fragment was obtained (*lane 2*). DNA size markers (*lane 1*) and a PCR blank (*lane 3*) are shown. *C*, 5′-flanking sequence of the *Nas1* gene. Primers used in primer extension and 5′-RACE experiments are indicated. The ATG translation initiation site is *boxed*. The transcription start sites as mapped by primer extension (defined as position +1) and 5′-RACE are indicated by the *arrow* and *asterisk*, respectively. Identical results were obtained in two additional experiments.

teopontin VDRE (28). Overlapping the two DR3, is a DR4-type structure (5′-AAGTCAgttaATTTCA-3′) with similarities to mouse Pit-1 VDRE (32). In addition to these three regions, five GREs are detected at position +621, +377, +248, −580, and −792. These motifs are similar to the consensus inverted palindromic GRE sequence 5′-AGAACAnnnTGTTCT-3′ that is distinct from the TRE/VDRE core binding motif (33). Within this consensus GRE, the core binding motif TGTTCT is generally well conserved, whereas the left-most hexanucleotide sequence can be quite variable (33).

Transcriptional Activity of the Nas1 Promoter—To determine whether the 5′-flanking region of the *Nas1* gene can initiate basal transcription, reporter constructs were made with the 5′ upstream region of *Nas1* fused to a luciferase reporter gene and transfected into OK, COS-1, and NIH3T3 cells (Fig. 8). In OK cells, the construct containing the *Nas1* sequence −457 to +70 (p*Nas1*-457) was able to induce the highest luciferase expression when compared with constructs containing more downstream sequences of the *Nas1* promoter (p*Nas1*-3229 and p*Nas1*-1203), suggesting the possible presence of downstream negative regulatory elements (Fig. 8). The luciferase activity driven from the sequence −1203 to +70 inserted in the reverse orientation (p*Nas1*-1203R) was not significantly different from the promoterless pGL3 vector. In contrast to OK cells, none of the *Nas1* promoter fragments were able to drive expression of the luciferase gene in COS-1 or NIH3T3 cells (Fig. 8).

Transactivation by VDR, RXR, and 1,25-(OH)$_2$D$_3$—The potential transcriptional activity of the *Nas1* promoter in response to 1,25-(OH)$_2$D$_3$ was initially tested using the p*Nas1*-1203 construct (Fig. 9*A*). Cotransfection of p*Nas1*-1203 and the VDR expression vector into OK cells did not result in increased luciferase activity, when compared with transfection with p*Nas1*-1203 alone. In contrast, in OK cells expressing both the VDR and hRXRα, the increase in luciferase activity was 8.9-fold higher than with activation of p*Nas1*-1203 alone. Under these conditions, in the presence of 0.5 and 50 nM 1,25-(OH)$_2$D$_3$, the promoter activity was further increased by 1.8- and 4.1-fold, respectively (Fig. 9*A*). Similar experiments were performed for the p*Nas1*-3229, p*Nas1*-1203, p*Nas1*-457, and p*Nas1*-1203R constructs, and the effect of 1,25-(OH)$_2$D$_3$ in OK cells coexpressing the VDR and hRXRα is summarized in Fig. 9*B*. When transfected with p*Nas1*-3229 or p*Nas1*-1203 constructs, the luciferase activity increased markedly upon exposure of cells to 0.5 or 50 nM 1,25-(OH)$_2$D$_3$. In contrast, the luciferase activity in OK cells transfected with p*Nas1*-457 or p*Nas1*-1203R was not affected by 1,25-(OH)$_2$D$_3$ (Fig. 9*B*).

DISCUSSION

NaSi-1 is a high affinity Na$^+$-sulfate cotransporter present on the BBM of the renal proximal tubule (9) and ileum (27). The

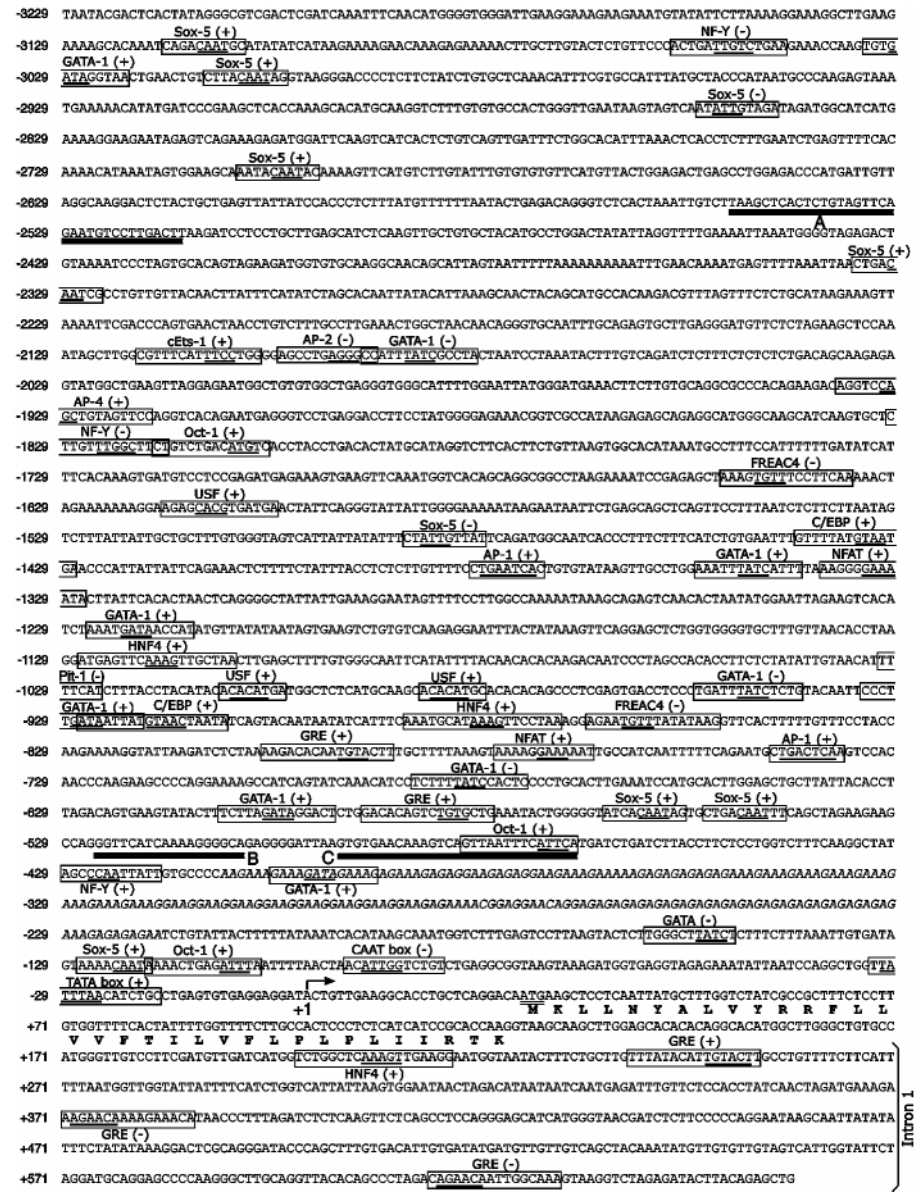

FIG. 7. **Nucleotide sequence of the 5′ region of the *Nas1* gene.** Nucleotide +1 denotes the transcription start site and is marked by an *arrow*. Amino acid sequence corresponding to exon 1 is in *boldface type* and indicated *below* the coding strand. The translation initiation ATG codon is *double underlined*. The TATA and CAAT boxes, as well as putative transcription factor binding motifs, with core sequence *underlined*, are *boxed*. Only the best potential binding motif for each of the transcription factors shown is boxed (core similarity >0.9, matrix similarity >0.85). Regions containing potential steroid/thyroid-responsive elements are *underlined* and labeled *A–C*, respectively. The GA-rich region of unknown function is in *italic uppercase. AP-1*, associated protein 1; *AP-2*, associated protein 2; AP-4, associated protein 4; *C/EBP*, CAAT/enhancer-binding protein; *c-Ets-1*, c-Ets-1 proto-oncogene product; *GATA-1*, GATA binding factor 1; *GRE*, glucocorticoid-responsive element; *HNF4*, hepatic nuclear factor 4; *FREAC4,* forkhead related activator 4; *NFAT*, nuclear factor of activated T cells; *NF-Y*, nuclear factor Y; *Oct-1*, octamer factor 1; *Sox-5*, SRY-related HMG box; *USF*, upstream stimulating factor. The analysis of the transcription factor-binding sites was performed using the MatInspector program (57).

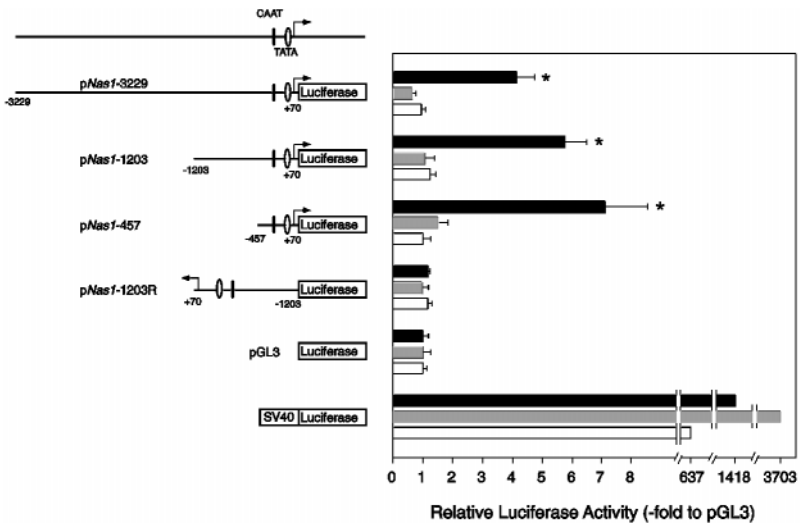

FIG. 8. **Transcriptional activity of the *Nas1* promoter region.** OK (*filled bars*), COS-1 (*gray bars*), and NIH3T3 (*open bars*) cells were transiently transfected using LipofectAMINE 2000 with 0.8 μg of reporter vector containing the luciferase gene under the control of the *Nas1* promoter and 0.8 μg of the pRSVβGal control plasmid. Transfected cells were harvested after 24 h and assayed for luciferase and β-galactosidase activity. Data are shown relative to activity observed with the pGL-3 vector (promoterless vector). The pGL-3 control vector (containing the SV40 promoter) was used as a positive control. Data are means ± S.D. from triplicate determinations, and the results are representative of three separate experiments.

tissue distribution, hormonal regulation, and characteristics of expressed activity of NaSi-1 suggest that this transporter may play a crucial role in the maintenance of sulfate homeostasis. Our report describes the first characterization of the genomic structure and fine chromosomal localization of a mammalian Na$^+$-coupled sulfate transporter gene. We also demonstrate a

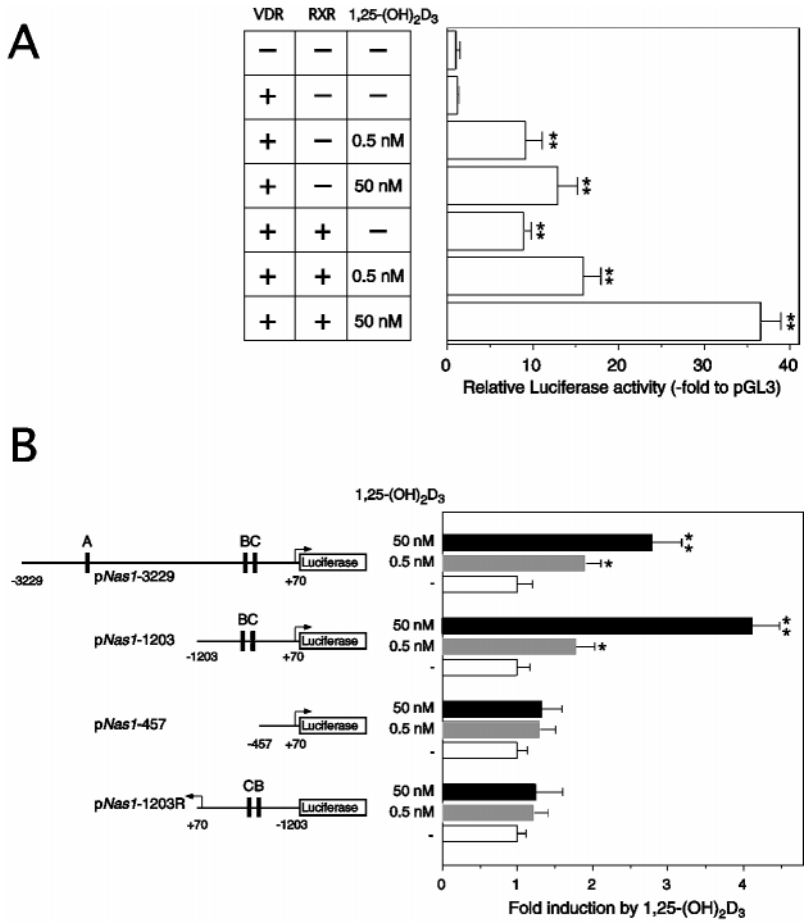

FIG. 9. **Regulation of the mouse *Nas1* promoter by VDR/RXR and 1,25-$(OH)_2D_3$.** *A*, a vector containing 1.2 kb of *Nas1* promoter sequence (p*Nas1*-1203, 0.8 μg) was transfected into OK cells with or without the VDR (0.1 μg) or hRXRα (0.1 μg) expression vector, as indicated. After 24 h, cells were incubated with 1,25-$(OH)_2D_3$ (0.5 or 50 nM) or ethanol as control for an additional 24 h before being harvested and assayed for luciferase and β-galactosidase activity. Results are represented as fold induction in OK cells as compared with p*Nas1*-1203 alone. **, $p < 0.01$ when compared with transfection with p*Nas1*-1203 alone. *B*, OK cells were transfected with 0.8 μg of a *Nas1* reporter plasmid (p*Nas1*-3229, p*Nas1*-1203, p*Nas1*-457, or p*Nas1*-1203R, as indicated), 0.1 μg of VDR expression vector, and 0.1 μg of hRXRα expression vector. After 24 h, cells were incubated with 1,25-$(OH)_2D_3$ (0.5 nM or 50 nM) or ethanol as control for an additional 24 h before being harvested. Results are expressed as fold induction by 1,25-$(OH)_2D_3$, as compared with no 1,25-$(OH)_2D_3$ treatment. Data are means ± S.D. from triplicate determinations and are representative of two separate experiments. *, $p < 0.05$ and **, $p < 0.01$ when compared with no 1,25-$(OH)_2D_3$ treatment.

cell-specific transcriptional activation of *Nas1* expression as well as a transactivation of the *Nas1* gene expression by 1,25-$(OH)_2D_3$.

The characteristics of expressed activity of mNaSi-1 protein in *Xenopus* oocytes, as well as the overall pattern of mNaSi-1 mRNA tissue distribution, support the finding that we have cloned the mouse proximal tubular BBM Na$^+$-sulfate cotransporter (34). Both kinetic data and tissue distribution were very similar to the rat NaSi-1 (8). Three mNaSi-1 cDNA variants were identified in mouse kidney mRNA. The two main transcripts (2.2 and 2.5 kb) could be detected by Northern analysis, and comparison with the genomic sequence showed that they were derived from alternative polyadenylation, as shown for the two rNaSi-1 transcripts (2.3 and 2.9 kb) detected in rat kidney mRNA (8, 27). The third mNaSi-1 variant (2.1 kb) was considered to be the result of alternative splicing of exon 2. This alternative spliced version of mNaSi-1 could not be detected on Northern blots and did not show significant sulfate transport when injected in *Xenopus* oocytes, despite the fact that no frameshift was introduced by exon 2 splicing. A possible explanation for the loss of function is that deletion of exon 2, encoding transmembrane segment 2, would lead to an inverted membrane topology of the protein due to the lack of one transmembrane segment. Alternatively, the removal of transmembrane segment 2 could disrupt the native signal anchor sequence and perturb the sequential mechanism of membrane insertion and folding (35). In view of our data, we conclude that the alternatively spliced mRNA, detected by RT-PCR only, represents a rare transcript that is unlikely to have significant biological relevance.

The mNaSi-1 protein secondary structure model of 13 transmembrane segments contrasts with the secondary structure prediction of rNaSi-1 protein, which was initially predicted to contain 8 transmembrane segments (8). The difference is most probably due to the difference in the prediction method, which was previously based on hydropathy analysis. The new prediction was performed using the TopPred2 program (23), featuring a more precise analysis of the hydropathy data and taking into account the inside positive rule (36). Consensus sequences for *N*-glycosylation were found at positions 140, 174, and 590; however, according to the 13 transmembrane helices model, only Asn^{590} is suggested to be extracellular and thus possibly glycosylated. This is consistent with the study of Pajor and Sun (37) showing glycosylation of rabbit NaDC-1 to only occur on Asn^{578}. However, although the 13 transmembrane domain model is likely to represent a better prediction than the 8 transmembrane domain model, additional work is needed to validate this model.

The *Nas1* gene contains 15 exons distributed among 75 kb without obvious pattern of exon organization. Although exons encoding transmembrane domains were similar in size, no sequence identity was detected at the nucleotide or amino acid level between them (data not shown), suggesting that these exons did not arise through duplication events. The recently identified SUT-1 transporter shares the highest sequence identity with mNaSi-1; however, its genomic structure is presently unknown and thus no comparison was possible. The genomic structure of human *NaDC-1* gene was recently reported and appeared to be significantly different from the *Nas1* gene, containing 12 exons distributed over 23.8 kb of genomic DNA (38). Similarly, the genomic structure of sulfate/anion exchangers differ considerably from the *Nas1* gene. The human *CLD* gene comprises of 21 exons spanning 39 kb (39), whereas the rat *Dtdst* gene contains only 5 exons spanning approximately

20 kb (40). In contrast, despite sharing no homology with mNaSi-1 cDNA, the human Na^+-glucose *SGLT1* transporter shares a comparable gene structure with *Nas1* consisting of 15 exons distributed among 72 kb (41). Particularly, the exon sizes and their distribution are very similar to that of *Nas1*. Moreover, most of the 35 members of the *SGLT1* gene family share a common core structure of 13 transmembrane helices (42), as is the case for *Nas1*. Superimposing the *Nas1* exon boundaries on the mNaSi-1 protein secondary structure model shows that splicing frequently occurs near membrane/aqueous transitions, as is also observed with the *SGLT1* gene (41) and other genes encoding membrane proteins, such as the murine band 3 (43), human skeletal muscle sodium channel (44), and *GLUT1-*, *GLUT2-,* and *GLUT4*-facilitated glucose transporters (45). It remains to be elucidated whether the similarities between the two Na^+-cotransporters, *SGLT1* and *Nas1,* are the result of a possible common evolutionary origin.

The *Nas1* gene was mapped to mouse chromosome 6, close to marker D6Mit170, which has been assigned a map position of 4.0 centimorgans. The mouse chromosome 6 region from centromere to map position 28 centimorgans is a region of conserved synteny with human chromosome 7. It contains 46 identified genes whose human homologues map to chromosome 7, between regions q14 and q35. Within this region, only one gene (centromere autoantigen E gene) maps to another chromosome (4q24-q25). Altogether these linkage data suggest that the human homologue of the *Nas1* gene most likely resides on human chromosome 7q. Interestingly, the human *SUT-1* gene, which displays 49% amino acid identity to mNaSi-1, was mapped to 7q33, close to 7q32 (26). Due to their high protein identity and possible chromosomal colocalization, studies in humans are warranted to determine whether the *SUT-1* and *NAS1* genes could have derived from a gene duplication event. A similar situation was described previously for the *DRA* and *PDS* genes, encoding for sulfate and chloride anion exchangers sharing 45% homology and both residing on human chromosome 7q21–31.1 (46).

Transcription initiation of the *Nas1* gene occurs at a single site and is under the control of an atypical TATA box located 29 bp upstream of +1, yielding mRNA with a short 26-bp 5′-UTR. The core and flanking residues of the atypical TATA box, T$\underline{\text{TAT}_0\text{TTAA}}$C, differ from the extended canonical sequence (G/C) $\underline{\text{TAT}_0\text{A(A/T)AA}}$(G/A) by having a T in the −3 position, a T in the +1 position, and a C in the +5 position. Nonetheless, a recent study showed that these bases can be present in these positions but at low frequencies (8, 8, and 11%, respectively) (47). Interestingly, the promoter region is AT-rich (61%), rather than GC-rich, a feature that, together with the unique transcription start site, has been observed for genes that are regulated during development and differentiation (48). The AT-rich feature is also consistent with the restricted pattern of *Nas1* gene expression since GC-rich promoters are commonly associated with widely expressed "housekeeping" genes. The role of *Nas1* during development and differentiation is yet unknown. However, it is interesting to note the presence of multiple binding sites for Sox-5, a transcription factor that has critical roles in the regulation of numerous developmental processes; upstream stimulating factor (USF), a ubiquitous factor involved in development; and hepatic nuclear factor 4 (HNF-4), a thyroid hormone receptor-like factor expressed in kidney from day 10.5 post-coitum and involved in development. In addition, two potential binding sites for FREAC-4, a transcription factor predominantly expressed in kidney (49) and having important roles in embryonic development, as well as regulation of tissue-specific gene expression (50, 51), were found. FREAC-4 has recently been shown to be regulated by the c-*Ets-1* proto-oncogene (51). Interestingly, a c-*Ets-1* putative binding site was also found in the *Nas1* promoter region.

When placed upstream of a reporter gene, the *Nas1* promoter could initiate basal gene transcription in a cell-specific manner, since the promoter was only active in OK cells and not in COS-1 or NIH3T3 cells, suggesting tissue specificity of promoter activity. Only FREAC-4-binding sites were identified as potential *cis*-acting elements associated with kidney-specific gene expression in this region, but their actual role in the cell-specific expression of *Nas1* remains unknown.

In a recent study, vitamin D was shown to modulate renal Na^+-sulfate cotransport (10). Vitamin D-deficient rats showed lower plasma sulfate levels and an increased fractional excretion of sulfate, which correlated with decreases in BBM Na^+-sulfate cotransport activity and rNaSi-1 protein and mRNA abundance. Moreover, this modulation was shown to be the result of a direct effect of vitamin D, with no independent action of parathyroid hormone or calcium levels (10). The data presented here extend these observations by demonstrating that vitamin D and VDR/RXR transactivated the *Nas1* promoter in OK cells. A comparable transactivation by VDR and 1,25-$(OH)_2D_3$ was also observed for the renal Na^+-dependent phosphate transporter gene, *NPT2* (52). Our data suggest that the previously reported effect of vitamin D on sulfate homeostasis (10) may, at least in part, be mediated by a transcriptional activation of the *Nas1* gene. Although the first 1.2 kb of *Nas1* promoter is sufficient for vitamin D transactivation, we cannot rule out that a further upstream region may play a role in this phenomenon, and more detailed studies are required to identify the specific VDRE loci in the *Nas1* promoter.

Finally, the presence of five putative GREs and a typical TRE in the *Nas1* promoter may also be of significant importance for *Nas1* gene regulation and hormonal control of sulfate homeostasis. Glucocorticoids have been shown to regulate renal Na^+-sulfate cotransport at the BBM level (53), and experimentally induced hypothyroidism in rats led to a decrease of NaSi-1 mRNA and protein levels with no change in membrane fluidity, suggesting a possible down-regulation of the *Nas1* gene (54).

In summary, we have isolated and characterized the murine Na^+-sulfate cotransporter cDNA, gene, and promoter region. In addition, we have investigated its expression in murine tissues, determined its chromosomal localization, identified cDNA variants, and demonstrated that this gene can be transcriptionally activated by 1,25-$(OH)_2D_3$ in a renal cell line. This study provides the framework for a more detailed analysis of *Nas1* gene expression through the characterization of *Nas1* promoter function and the tools required for assessing the role of *Nas1* in the maintenance of sulfate homeostasis through the generation and analysis of *Nas1*-deficient mice.

Acknowledgments—We are grateful to Dr. John White (McGill University, Montreal, Canada) for the generous gift of the VDR and hRXRα expression vectors and Dr. Michael Waters (University of Queensland, Brisbane, Australia) for providing the pRSVβGal plasmid.

REFERENCES

1. Tallgren, L. (1980) *Acta Med. Scand.* **640,** (suppl.) 1–100
2. Hastbacka, J., de la Chapelle, A., Mahtani, M. M., Clines, G., Reeve-Daly, M. P., Daly, M., Hamilton, B. A., Kusumi, K., Trivedi, B., Weaver, A., and *et al.* (1994) *Cell* **78,** 1073–1087
3. Hastbacka, J., Superti-Furga, A., Wilcox, W. R., Rimoin, D. L., Cohn, D. H., and Lander, E. S. (1996) *Am. J. Hum. Genet.* **58,** 255–262
4. Superti-Furga, A., Hastbacka, J., Wilcox, W. R., Cohn, D. H., van der Harten, H. J., Rossi, A., Blau, N., Rimoin, D. L., Steinmann, B., Lander, E. S., and Gitzelmann, R. (1996) *Nat. Genet.* **12,** 100–102
5. Murer, H., Manganel, M., and Roch-Ramel, F. (1992) in *Handbook of Physiology* (Winhager, E., ed) Vol. 2, pp. 2165–2188, Oxford University Press, Oxford
6. Besseghir, K., and Roch-Ramel, F. (1987) *Renal Physiol.* **10,** 221–241
7. Frick, A., and Durasin, I. (1986) *Pfluegers Arch.* **407,** 541–546
8. Markovich, D., Forgo, J., Stange, G., Biber, J., and Murer, H. (1993) *Proc. Natl.*

Acad. Sci. U. S. A. **90,** 8073–8077
9. Lotscher, M., Custer, M., Quabius, E. S., Kaissling, B., Murer, H., and Biber, J. (1996) *Pfluegers Arch. Eur. J. Physiol.* **432,** 373–378
10. Fernandes, I., Hampson, G., Cahours, X., Morin, P., Coureau, C., Couette, S., Prie, D., Biber, J., Murer, H., Friedlander, G., and Silve, C. (1997) *J. Clin. Invest.* **100,** 2196–2203
11. Markovich, D., Murer, H., Biber, J., Sakhaee, K., Pak, C., and Levi, M. (1998) *J. Am. Soc. Nephrol.* **9,** 1568–1573
12. Sagawa, K., DuBois, D. C., Almon, R. R., Murer, H., and Morris, M. E. (1998) *J. Pharmacol. Exp. Ther.* **287,** 1056–1062
13. Puttaparthi, K., Markovich, D., Halaihel, N., Wilson, P., Zajicek, H. K., Wang, H., Biber, J., Murer, H., Rogers, T., and Levi, M. (1999) *Am. J. Physiol.* **276,** C1398–C1404
14. Sagawa, K., Han, B., DuBois, D. C., Murer, H., Almon, R. R., and Morris, M. E. (1999) *J. Pharmacol. Exp. Ther.* **290,** 1182–1187
15. Markovich, D., and Knight, D. (1998) *Am. J. Physiol.* **274,** F283—F289
16. Markovich, D., Wang, H., Puttaparthi, K., Zajicek, H., Rogers, T., Murer, H., Biber, J., and Levi, M. (1999) *Kidney Int.* **55,** 244–251
17. Sagawa, K., Benincosa, L. J., Murer, H., and Morris, M. E. (1998) *J. Pharmacol. Exp. Ther.* **287,** 1092–1097
18. Chen, Z. (1996) *Trends Genet.* **12,** 87–88
19. Markovich, D., Bissig, M., Sorribas, V., Hagenbuch, B., Meier, P. J., and Murer, H. (1994) *J. Biol. Chem.* **269,** 3022–3026
20. Beck, L., Soumounou, Y., Martel, J., Krishnamurthy, G., Gauthier, C., Goodyer, C. G., and Tenenhouse, H. S. (1997) *J. Clin. Invest.* **99,** 1200–1209
21. Lardelli, M., and Lendahl, U. (1994) *BioTechniques* **16,** 420–422
22. Cheng, S., Fockler, C., Barnes, W. M., and Higuchi, R. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 5695–5699
23. von Heijne, G. (1992) *J. Mol. Biol.* **225,** 487–494
24. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
25. 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
26. Girard, J. P., Baekkevold, E. S., Feliu, J., Brandtzaeg, P., and Amalric, F. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 12772–12777
27. Norbis, F., Perego, C., Markovich, D., Stange, G., Verri, T., and Murer, H. (1994) *Pfluegers Arch. Eur. J. Physiol.* **428,** 217–223
28. Noda, M., Vogel, R. L., Craig, A. M., Prahl, J., DeLuca, H. F., and Denhardt, D. T. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 9995–9999
29. Carlberg, C., and Polly, P. (1998) *Crit. Rev. Eukaryotic Gene Expr.* **8,** 19–42
30. Demay, M. B., Gerardi, J. M., DeLuca, H. F., and Kronenberg, H. M. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87,** 369–373
31. Darwish, H. M., and DeLuca, H. F. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 603–607
32. Rhodes, S. J., Chen, R., DiMattia, G. E., Scully, K. M., Kalla, K. A., Lin, S. C., Yu, V. C., and Rosenfeld, M. G. (1993) *Genes Dev.* **7,** 913–932
33. Lucas, P. C., and Granner, D. K. (1992) *Annu. Rev. Biochem.* **61,** 1131–1173
34. Lücke, H., Stange, G., and Murer, H. (1979) *Biochem. J.* **182,** 223–229
35. von Heijne, G. (1996) *Prog. Biophys. Mol. Biol.* **66,** 113–139
36. Hartmann, E., Rapoport, T. A., and Lodish, H. F. (1989) *Proc. Natl. Acad. Sci. U. S. A.* **86,** 5786–5790
37. Pajor, A. M., and Sun, N. (1996) *Am. J. Physiol.* **271,** C1808—C1816
38. Aruga, S., Moe, O. W., Preisig, P. A., Pajor, A., and Alpern, R. J. (1999) *J. Am. Soc. Nephrol.* **10,** 50 (abstr.)
39. Haila, S., Hoglund, P., Scherer, S. W., Lee, J. R., Kristo, P., Coyle, B., Trembath, R., Holmberg, C., de la Chapelle, A., and Kere, J. (1998) *Gene (Amst.)* **214,** 87–93
40. Satoh, H., Susaki, M., Shukunami, C., Iyama, K., Negoro, T., and Hiraki, Y. (1998) *J. Biol. Chem.* **273,** 12307–12315
41. Turk, E., Martin, M. G., and Wright, E. M. (1994) *J. Biol. Chem.* **269,** 15204–15209
42. Wright, E. M., Loo, D. D., Panayotova-Heiermann, M., Hirayama, B. A., Turk, E., Eskandari, S., and Lam, J. T. (1998) *Acta. Physiol. Scand. Suppl.* **643,** 257–264
43. Kopito, R. R., Andersson, M., and Lodish, H. F. (1987) *J. Biol. Chem.* **262,** 8035–8040
44. McClatchey, A. I., Lin, C. S., Wang, J., Hoffman, E. P., Rojas, C., and Gusella, J. F. (1992) *Hum. Mol. Genet.* **1,** 521–527
45. Bell, G. I., Kayano, T., Buse, J. B., Burant, C. F., Takeda, J., Lin, D., Fukumoto, H., and Seino, S. (1990) *Diabetes Care* **13,** 198–208
46. Everett, L. A., Glaser, B., Beck, J. C., Idol, J. R., Buchs, A., Heyman, M., Adawi, F., Hazani, E., Nassir, E., Baxevanis, A. D., Sheffield, V. C., and Green, E. D. (1997) *Nat. Genet.* **17,** 411–422
47. Bucher, P. (1990) *J. Mol. Biol.* **212,** 563–578
48. Smale, S. T., and Baltimore, D. (1989) *Cell* **57,** 103–113
49. Pierrou, S., Hellqvist, M., Samuelsson, L., Enerback, S., and Carlsson, P. (1994) *EMBO J.* **13,** 5002–5012
50. Ernstsson, S., Pierrou, S., Hulander, M., Cederberg, A., Hellqvist, M., Carlsson, P., and Enerback, S. (1996) *J. Biol. Chem.* **271,** 21094–21099
51. Cederberg, A., Hulander, M., Carlsson, P., and Enerback, S. (1999) *J. Biol. Chem.* **274,** 165–169
52. Taketani, Y., Miyamoto, K., Tanaka, K., Katai, K., Chikamori, M., Tatsumi, S., Segawa, H., Yamamoto, H., Morita, K., and Takeda, E. (1997) *Biochem. J.* **324,** 927–934
53. Renfro, J., Clark, N., Metts, R., and Lynch, M. (1989) *Am. J. Physiol.* **256,** 1176–1183
54. Sagawa, K., Murer, H., and Morris, M. E. (1999) *Am. J. Physiol.* **276,** F164—F171
55. Hofmann, K., and Stoffel, W. (1993) *Biol. Chem. Hoppe-Seyler* **347,** 166
56. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998) *Bioinformatics* **14,** 378–379
57. Quandt, K., Frech, K., Karas, H., Wingender, E., and Werner, T. (1995) *Nucleic Acids Res.* **23,** 4878–4884

The Mouse Na^+-Sulfate Cotransporter Gene*Nas1* : CLONING, TISSUE DISTRIBUTION, GENE STRUCTURE, CHROMOSOMAL ASSIGNMENT, AND TRANSCRIPTIONAL REGULATION BY VITAMIN D

Laurent Beck and Daniel Markovich

J. Biol. Chem. 2000, 275:11880-11890.
doi: 10.1074/jbc.275.16.11880