

# Molecular cloning of an atypical voltage-gated sodium channel expressed in human heart and uterus: Evidence for a distinct gene family

(complementary DNA/heart muscle/skeletal muscle/cardiovascular system)

ALFRED L. GEORGE, JR.\*<sup>‡</sup>, TIMOTHY J. KNITTLE<sup>†</sup>, AND MICHAEL M. TAMKUN<sup>†‡</sup>

\*Department of Medicine, University of Pennsylvania School of Medicine, Philadelphia, PA 19104; and <sup>†</sup>Departments of Molecular Physiology & Biophysics and Pharmacology, Vanderbilt University School of Medicine, Nashville, TN 37232

Communicated by William A. Catterall, February 27, 1992

**ABSTRACT** Previously cloned voltage-dependent sodium channels exhibit a high degree of homology to one another and appear to comprise a single multigene family. We have now isolated and characterized cDNAs from both human adult heart and fetal skeletal muscle that encode a sodium channel  $\alpha$  subunit that exhibits only moderate primary structure identity with other sodium channels and is prominently expressed in both heart and uterus. The  $\approx 7.2$ -kilobase cDNA sequence, designated hNav2.1, predicts a 1682-amino acid protein that bears 52%, 49%, and 46% overall identity with sodium channels cloned from rat brain, skeletal muscle, and heart, respectively. Positively charged S4 segments are present in hNav2.1, but there are fewer basic residues in repeat domains 1, 3, and 4 than in other cloned sodium channels. The cloning of hNav2.1 provides evidence for greater evolutionary divergence among voltage-dependent sodium channels and suggests that other sodium channel gene subfamilies may exist. The unique amino acid sequences in regions known to be involved in voltage-dependent activation and inactivation suggest that hNav2.1 will have novel gating properties.

Voltage-dependent sodium channels (NaChs) are responsible for the depolarizing phase of the action potential in most electrically excitable cells and are essential for many diverse physiological processes (1, 2). All previously reported rat NaCh cDNA sequences (3–8) predict proteins that exhibit striking similarity to one another ( $>60\%$  overall amino acid sequence identity) and appear to comprise a multigene family. There is evidence that this NaCh gene family has related members in such diverse species as eel (9), *Drosophila* (10, 11), and man (12, 13). NaChs may have a common ancestry with other voltage-dependent ion channels and appear to have evolved from primitive single-transmembrane-domain channels by gene duplication (14). In view of the extensive structural diversity that exists among members of the voltage-dependent potassium channel family, it is surprising that a greater variety of mammalian NaCh isoforms have not been identified.

The predominant isoform in adult heart is insensitive to tetrodotoxin and is responsible for most of the sodium current observed in mammalian cardiac myocytes (15). In addition, there is evidence that both the rat brain I isoform (8) and another partially characterized NaCh cDNA sequence (6) are present in heart, but their exact cellular location is unknown. The presence of other functional NaChs in heart is supported by both pharmacological and electrophysiological studies. In rat heart homogenates,  $\approx 20\%$  of the binding sites for saxitoxin exhibit high affinity, and these data may indicate the presence of a NaCh isoform other than the predominant

tetrodotoxin-insensitive channel (16). In addition, voltage-clamp recordings of heart cells from a variety of species have demonstrated the existence of a class of low-amplitude, sodium- and voltage-dependent, slow inward currents that may contribute to the plateau phase of the cardiac action potential (17–20). These observations suggest the existence of a class of voltage-dependent, sodium-conducting channels that exhibit functional properties very distinct from those of the predominant cardiac NaCh.

We report here the cloning of a cDNA from adult human heart and fetal human skeletal muscle that is likely to encode a unique mammalian NaCh isoform.<sup>§</sup> The differences in overall amino acid sequence between this channel and previously described mammalian NaChs suggest that this protein is a member of a distinct subfamily of voltage-dependent NaChs. We have designated this channel hNav2.1, adopting a nomenclature developed for the potassium channels (21).

## MATERIALS AND METHODS

**RNA Isolation.** Freshly explanted human myocardium (left ventricle) was obtained from the Cardiac Transplantation Program of Vanderbilt University Hospital. All other human tissues were obtained as fresh frozen surgical pathology specimens or procured from cadaveric organ transplant donors by the National Disease Research Interchange (Philadelphia). Total RNA was isolated by the guanidinium thiocyanate method (22), and poly(A)<sup>+</sup> RNA was purified by oligo(dT)-cellulose chromatography (23).

**cDNA Library Construction and Screening.** An oligo(dT)-primed, size-selected [ $>2.0$  kilobases (kb)] human heart cDNA library was constructed in  $\lambda$ gt10 by using a cDNA synthesis system (Pharmacia) according to the manufacturer's instructions. Approximately  $1.5 \times 10^6$  primary plaques were screened (24) with a human cardiac cDNA fragment (25) homologous to nucleotides (nt) 4407–5121 of the rat brain II NaCh (3). Hybridizations were performed as described (26). A second  $\lambda$ gt10 cDNA library, made from fetal human skeletal muscle and generously provided by Eric Hoffman (27), was screened with a cDNA probe corresponding to nt 367–4402 of the rat skeletal muscle NaCh isoform SkM1 (5). Approximately  $6 \times 10^5$  amplified recombinants were screened with hybridization and washing conditions as described (13).

**Nucleotide Sequence Analysis.** Nucleotide sequence was determined on both strands of cDNAs hNa9-1 and hNa5 and on most of one strand of clones 1B-1 and 2A by the dide-

Abbreviations: NaCh, sodium channel; nt, nucleotide(s).

<sup>‡</sup>To whom reprint requests should be addressed at: Department of Molecular Physiology & Biophysics, Room 724, Medical Research Building, Vanderbilt University Medical Center, 21st and Garland, Nashville, TN 37232.

<sup>§</sup>The sequence reported in this paper has been deposited in the GenBank data base (accession no. M91556).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

oxynucleotide chain termination method using Sequenase 2.0 (United States Biochemical) and appropriate oligonucleotide primers. Portions of the 3' untranslated regions were not sequenced. Double-stranded templates were generated either by the exonuclease III deletion method (28) or by subcloning of restriction fragments.

**Northern Blot Analysis.** Samples (10  $\mu$ g) of total RNA were size-fractionated in formaldehyde/agarose gels, electrophoreted to nylon membranes (Nytran; Schleicher & Schull), and probed under high stringency with a cDNA probe corresponding to the carboxyl terminus and 3' untranslated region of hNav2.1. Exact hybridization and washing conditions were as described (26).

## RESULTS

**Isolation and Characterization of NaCh cDNAs.** Five positive clones that could be identified by sequence analysis as members of the NaCh gene family were isolated from the human heart cDNA library. The deduced peptide sequence of three of these clones was nearly identical to either the rat heart I (RH-I; ref. 8) or the rat brain I (RB-I; ref. 3) NaCh isoform and most likely represent the human homologs of these channels. Two cDNAs designated hNa5 (5.0 kb) and hNa7 (2.8 kb) were homologous to other NaCh nucleotide sequences but exhibited less sequence identity to previously cloned rat channels than did the other clones. A restriction fragment ( $\approx$ 0.6 kb) of hNa5 representing the 5' end of this clone was used to rescreen the library under high stringency conditions. Multiple overlapping cDNAs were isolated and the largest of these, hNa9-1 (2.4 kb), was also further characterized.

The initial screening of the fetal skeletal muscle cDNA library resulted in the identification of two overlapping cDNAs designated 1B-1 (2.1 kb) and 2A (4.9 kb). Sequence analysis revealed significant homology with previously cloned NaChs, but not of a sufficient degree to indicate that these clones were human homologs of any known channel sequence. When compared with cDNAs hNa9-1 and hNa5, clones 1B-1 and 2A were virtually identical in nucleotide sequence except that hNa9-1 extended 351 nt further in the 5' direction than 1B-1. All the cDNAs are represented in Fig. 1.

The nucleotide sequence of hNa9-1 and hNa5 was determined except for parts of the 3' untranslated region. The complete sequence, designated hNav2.1, consists of 139 nt of 5' untranslated sequence, an open reading frame of 5046 nt, and an  $\approx$ 2.0-kb 3' untranslated region that includes a polyadenylation signal (AATAAA) and a poly(dA) tail. The nucleotide sequence immediately surrounding the assigned initiation codon (TGGAAATGT, nt 135–143) resembles the

consensus eukaryotic initiation sequence (29) only in the presence of a purine nucleotide at position –3 (relative to the start codon). Interestingly, 5' of the assigned start codon are two ATG triplets (nt 29–31 and 37–39) within a region that is more similar to the initiating sequences of previously cloned mammalian NaChs (ATGCCAACATGG); however, both precede short open reading frames. The possibility of a cloning artifact causing a frameshift between the assigned start codon and upstream ATG triplets was excluded by amplifying this region from human heart and uterus via the PCR and sequencing the amplified cDNAs. The sequence of the PCR-generated fragments was identical to the sequence obtained from the hNa9-1 clone. An in-frame termination codon (TAA) is present at nt 5186–5188.

**Primary Structure of hNav2.1.** The deduced primary structure of hNav2.1 consists of 1682 amino acids and has a calculated molecular weight of 193,472. Prediction of transmembrane topology by hydropathy analysis reveals a profile similar to other NaChs, with four large (229–280 residues) hydrophobic domains each composed of at least six potential membrane-spanning  $\alpha$ -helical segments including a positively charged amphipathic segment (S4). Comparison of hNav2.1 to each of the five complete rat NaCh sequences reveals a uniform pattern of overall primary structure homology (overall amino acid identity is 52% for rat brain II, 51% for both rat brain I and III, 49% for rat skeletal muscle SkM1, and 46% for rat heart I) indicating a high degree of relatedness to this family of ion channel proteins. Significant homology is also evident from comparisons of hNav2.1 to NaChs from the eel electric organ (44% amino acid identity) and the *Drosophila* para locus (22%). Fig. 2 shows the primary sequence of hNav2.1 aligned with the rat brain II and rat heart I NaChs. A search of available protein and nucleic acid data bases (GenBank, Swiss-Prot, EMBL; December 1991) revealed significant homology between hNav2.1 and the  $\alpha_1$  subunit of L-type voltage-dependent calcium channels, although the extent of primary sequence identity (9–14%) was less than that found for NaChs. hNav2.1 shares no significant overall amino acid identity with voltage-dependent potassium channels or other known protein sequences. These findings support the contention that hNav2.1 is a voltage-dependent NaCh.

Regional comparisons of hNav2.1 to other NaChs reveal a high degree of homology within the repeat domains, where up to 60% amino acid identity is observed (Table 1). In contrast, there is poor conservation of primary sequence within the interdomain regions ID1–2 and ID2–3. The highest degree of amino acid sequence identity with other NaChs is found within two short segments (SS1 and SS2) of the S5–S6 interhelical region (Fig. 3) that are believed to form membrane-penetrating hairpin structures that contribute to the formation of the ion pore (31) and various neurotoxin binding sites (32, 33). The primary sequence in this region is much more homologous to the NaCh gene family than to the  $\alpha_1$  subunit of voltage-dependent calcium channels (Fig. 3).

The primary sequence of hNav2.1 has distinct features in regions known to be important in voltage-dependent activation and inactivation. The S4 segments, which are believed to function as voltage sensors, collectively exhibit fewer positive charges than is typical of other NaChs (Fig. 4). The greatest differences occur in domain 4, where there are only four arginine or lysine residues as compared to eight for other NaChs. Histidines (residues 1355 and 1367) replace arginines at two positions in the S4 segment of domain 4. Other S4 segment variations include the substitution of glutamines (residues 207 and 1346) for arginines in domains 1 and 4 and of aliphatic residues (Ile-1028 and Leu-1349) for arginines in domains 3 and 4. Only the S4 segment in domain 2 has the same number of positively charged residues as do other NaChs. The ID3–4 region, which has an essential role in

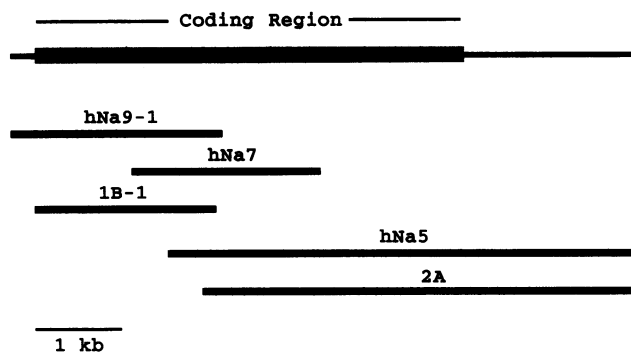


FIG. 1. hNav2.1 cDNAs. The bar represents the longest open reading frame encoded by these overlapping cDNAs, and untranslated regions are shown as flanking black lines. Clones hNa9-1, hNa7, and hNa5 were isolated from the human cardiac cDNA library; clones 1B-1 and 2A were isolated from the human fetal skeletal muscle library.

hNav2.1 MLASPEPKGLVPFTKESFELIKQHIA KTHNEDHEED LKPNPDLEVGGKLPFIYGNLSQGMVSEPLEDVPYKKNTFIVLNKNTIRFRNAASILCTLSPFNC  
 RH-I MA NL-LPRGTSSFR--R--LAAIEKRM-EKQARGSA-SQ-SREGLQEEAPR-QL--QAS---DL---PPRELIG---L--F-STQ- ---GKT---S-TNA-YV-S--HP  
 RB-II MARSV-VPPGDSFRF--R--LAAIE-R--E EKAKRPKQ- RKD--DENG- -S--A--S---DIPPE---L---IN---GKA-S-S-T-A-YI-T---P

**I 81** **I 82** **I 83** **I 84**

107 IRRITIKVLVHPFFQFILISVLIDCVFMSLTNLKWRPVLNTLLGIYTFEILVLKFAFGVWAGSFSFLGDPWNWLDVSVTVFEVLIIRYSPLEDFIPTLTQTARTLRILKIIPLNQGLKSLVG  
 121 V-RAAV-I---SL-SML-MCTI-TN---AQHDP-P-TKYV-Y-FTA---IL---FCLHA-T-R---I-MAYTTEFVD-GNVSA-R-F-V--A--T-SVIS---TI--  
 118 --KLA--I---SL-NVL-MCTI-TN---TMS-P-D-TKNV-Y-FT---S-I-IL---FCLD-T-RN---T-IT-AVYTFEVN-GNVSA-R-F-V--A--T-SVIP---TI--

**I 85** **I 86**

229 VLIHCLKQLIGVILITLFFLSIFSLIGMGLFMGNLKHKC FRWQENE NETLHNRT GNPYYIRETENFYLEGERYALCGNRTDAGQCEGYVCVKAGIN  
 243 A--QSV-K-AD-MV--V-C--V-A--LQ-----R--VRNFTELNGTNGSVEADGLVWNSLD V-LNDPANVLLKN-TTDV---SS---T---R-L--E-  
 240 A--QSV-K-SD-M--V-C--V-A--LQ-----RN-- LQ--PDNSTFEINITSFFNSLDW-G-AF---VNMFNWDE--EDKSH--F--QND-----SS-----I-----R-

**I 87** **I 88** **I 89** **I 90**

329 PDQGFNTNDSFGWALFALFRMAQDYPEVLYHLILYASGVYMIFFVVSFLSFYMASFLGILAMAYEEERQVGEISKKEIPKFPQ  
 349 --H-Y-S---A--FL---T--CW-R--QQT-RSA--I---ML-I--G---LVN-I-AVV---QN-ATIAETEEK-KR--EAMEMLKK EHEA LTIRGVDVTSR SS  
 360 -NY-Y-S--T-S--FLS---T--FW-N-Q-T-R-A--T---L-I--G---LVN-I-AVV---QN-ATL-EABQK-AE---MLEQLKKQEEAQAASASERDFSGAGGIG

**I 91** **I 92** **I 93** **I 94**

418 TGKEL QEGNETDEAKTIQIEMKK RSP IST  
 462 LEMSLAPVTNHERKS KRRK RLSSGTEDGDDRLPKSDSDGPRALNQLSLTHGLSRTSMRPRS RGSIFTFR RDQGEADFADDE--AGE  
 482 VFSESSSVASKLSSKSE---KNRRKKKKQKEQAGEEKEKDAVRKSASEDSIR--GFQFSLEGSRLTYEKFRS--HQSLLSIRGSLFSPRRNSRSLFNFKGVKRDIGSENFADDEH--FED

**I 95** **I 96** **I 97** **I 98**

447 SESHRSTLLVP WPLRHPS AQCGPGPGASAGPYVLNGKRNSTVDCNGVSLLGAGDAEA--PGSYLLRPMVLDLDRPDTTTPSEEPGPGQMLTPQAPCADGFE--PGARQALSAVSVLTS  
 555 NDSRRDLFVPHRHGERRPSNSQASRASRGITPLMNGKMSAVDCNGVSL VGGPSAL--PVQQL PEGTTT ETEIRKRSSSYHVSML--PSRQAMSMASILNT  
 604

**II 81** **II 82** **II 83**

461 KEELEKSKKICPLYWYFAKTLIWNCSPCWKLKEFVHRITIMAPTDLFLIICILNVCFLTEHYPMKQNTLLNIGNLVIGIFTAEMIPIKIAMHPYGFQVGNWIFDSMIVFHLI  
 774 L---E-HRK--PC-NR--QHY--E-CPL-MSI-QK-KFVV--D-A--TITM--V-TL-MAL--N-TAEFEEM-QV---T---LD--Y--Q---I---ILSLM  
 715 M---E-RQK--PC---NMC---D-CKP---V-HV-NLVV--D-V--AIT--V--TL-MAM---TE-FSSV-SV---T---FL---D--Y--E---GF--SLSLM

**II 84** **II 85** **II 86**

583 ELCLANVAGMALLRLFRMLRIFKLGKYWPTFQILMWSLSNSVAKDLVLLFTFFSAAFGMKLFQGNKYEEFVCHIDKQLPRWHMHDFHFSFLNVFRILCGEWETLWDCMEVAGQSW  
 796 --G-SRMGNLSV--S--L--V--A--S--LNT-IKIIG-VG--GN-T-V-AIIV-IF-VV--Q---S-LRHR-SDSGL---MD---A--II---I--M---S--SL  
 837 --G---E-LSV--S--L--V--A--S--LNM-IKIIG-VG--GN-T-V-AIIV-IF-VV--Q---S-K-C--K-SN-E---H---I--V---I--M---TM

**II 87** **II 88** **II 89** **II 90**

705 CIPFYLMLVILIGNLLVLYLFAL VSSFSSCKDVTAEENEAQNLQAVARIKGINV LKILCKTQ NVPKDMDHVNVEYVKED ISDHTLSLSNTQDFLKD K  
 918 -LLVFL-MV---V--N---LL---ADNLTAPDEDG-MN---L--QR-LRFVKRTWDFCCGILRRRPKPKAALATHSQLPSCITAPSPPPPEVKEVPPARKETFEEDKRPQ  
 959 -LTVFM-MV---V--N---LL---DNLAATDD-MN---I--G-MQ--DFVKR --REFI-KAFV RKQKAL D-IKPL--LNNKDKSC--N--TI-IGKDLNY--GN

**II 91** **II 92** **II 93** **II 94**

809 EKSSG TEKNATENES QSLIPSPS VSETVPIASGESDIENLDNKEIQS KSGDGSKEKIKQSSSSECT VDI AISEEEM  
 1040 GT P-DSEPVCPPIAVA--DTEDQEEDEEN-LGTEESSKQESQVSGHEPYQ--PRAWVSQSETTS-EAGASTQADWQEQKTEPQAPGCGETPED-Y--G--ADMTNADLLEQIPDLG  
 1073 GTT--IGSSV--YVVD--D YM-F-NN--L TV---L---F--NTE-FS-E -DMEE---LNAT---G--GAPAEQGQ-P-A-P

**II 95** **II 96** **II 97** **II 98**

888 FYGERSKHLKNGCRGSSQIGSGASKKGIWQIRKTCCKIVENNWFKCFGLVTLTSTGTAFEDIYDQKTKILLEYADMIFTYIFILEMLLKMAYGFKAYFSNGWYRDFVVI  
 1161 EDVKDPEDCFTE--V-RCPCCMVDTQSP--V-WRL---YR---HS--ET--IFMI--S-A---LEE---VL---KM---V---V---K--T-A-CW---LI-D  
 1162 EESL-PEACFTED-V-KFKCC--IEEG--L-W-L---Y---H--ET--VFMI--S-A---E---TM---KV---V---V---Q---M--T-A-CW---LI-D

**II 99** **II 100** **II 101** **II 102**

1010 VFCLSL IGKTREEL KPLISMFLRPLRVLSQFERMKVVRALIKTTLPLNVLVCLMIWLFISIMGVDLFAGRFYECIDPTSGE RFPSEVMNKSRCESLLFNESMLWENAKMF  
 1283 -SLV--VANTLGFA-MGPI-S-RTLRA---A--R--G-R--N--VGAIPSIM--L--IF-----N---K-GR--NQ-E-DLPLNYTIVN---E---FNVGTGELY-TKV-V--  
 1284 -SLV--TANALGYS--GAI-S-RTLRA---A--R--G-R--N--LGAIPSIM--L--IF-----N---K--H--NY-T-MFDV-VVNNYSECQALIES-QTAR-K-V-V--

**II 103** **II 104** **II 105** **II 106**

1126 DNVNGFSLSLQVATFNWITMNSAIDSVAVNIQPHFEVNIYMYCYFINIFIGVFLPLSLMILTVIINDFNKHKIKLGGSNIFTVQKQKQYRRLKKLMYEDSQRPVPRPLNKLQGFIFV  
 1405 ---A-Y-A--K--MD--YA-V--RGYEE--QW-D-L--I--VV---S-FT-NLF-G---QQ-K---QD--M-EE-K-K-I---GSKKP-K-I---I---I---I  
 1406 ---L-Y---K--MD--YA-V--RN-EL--KY-D-L--L--VI---S-FT-NLF-G---QQ-K-F--QD--M-EE-K-Y-NAM--GSKKP-K-I---A--F--MV--F

**III 81** **III 82** **III 83** **III 84**

1248 VTSQAFNVIVMVLICFOAIAMIDTDVQSLQMSIALYWINIFVLMYIMECILKLIARCFYFTIWNIFDFMVVIFISITGLCLPMTVGSYLVPPSLVQLILLSRIHILKLGKPKVPHNL  
 1527 --K--D-TI-F--LNMVT--VE--D--PEKVNI-AK--LL--AIF-G--V-MA-L-HY--NS-----V--L--V-TV-SDIIQK-FFS-T-FRV-R-A--GRI---IR-A-GIRT-  
 1528 --K-V-DISI-I--LNMVT--VE--D--QE-TNI---LV-IV-F-G--V--SL-HY---G-----V--L--V-MF-AELIEK-F-S-T-FRV-R-A--GRI---I-A-GIRT-

**III 85** **III 86** **III 87** **III 88**

1370 MLPLMLSPALLNIIILLFVLMFYAVFGMYNFAVYKKEAGINDVSNFETFGNSMLCLFQVAIFAGWGLDAIFNSKWSDCDP DKINPGTVRGDCGNPSGVIFYFVSVYLLISLIVNM  
 1649 LFA--M---F--G--L-----SI---A-----W---D-MF--Q--A-----ITTS---L-SP-L-TGPPY---NLPNSN-S R-N--S-A--LF-TT--I--F--V--  
 1650 LFA--M---F--G--L-----I---S-----R-V--D-MF-----I---ITTS---L-AP-L--GPP---E-DH--SS-K-----F---I--F--V--

**III 89** **III 90** **III 91** **III 92**

1491 YIVVMEFLNIASKKKNKLTSEDDFRKFFQWKRFPDPTQYIDSSKLSDFAAALDPLPMAKPNKQGLIALDLPMAGVDRIHCLDILLAFTRKVMQGVDRMEKVSEIESGFLLANPFKIT  
 1769 --AII-L-NFSV-TEESTEP---DM-YEI-EK--EA--F-EYLA---D--SE--RI---QIS--NM---VS-----M--F-----L-ESGE-DALKIQM-EK-MAA--S--S  
 1771 --A-IL-NFSV-TEESAEP---EM-YE-EK--A--F-EFC-----LI---V---M---VS-----F-----L-ESGE-DALRIQM-ER-MAS--S--VS

**III 93** **III 94** **III 95** **III 96**

1613 CEPITTLTKRKQEAVSATIIQRAY KNYRLRRNDK NTSDI HMDIGDRDV HAT K EGAYFDK AKEKSPIQSQI  
 1891 Y-----R--H-E-----FRHLLQRSVKHASFLFRQAGGSLSDAPEREGLIAYMMNGNFRSRASPLSSSISSTSPPSYDSV-RATSDNLGVRAS-YRSR-DLADFPSPD  
 1893 Y-----E--IV-----RRYLLKQKVK -VSSIIYK--GKEDE GTPIKE--ITDKLNENSTPEKT--TPSTTSPSYDSV--PEK-KFEK--SE--D KKG

**III 97** **III 98** **III 99** **III 100**

1683  
 2013 RDRESIV  
 1999 DRESKK

FIG. 2. Amino acid sequence of hNav2.1, deduced from its cloned cDNA, is compared with the sequences of rat brain II (RB-II; ref. 3) and rat heart I (RH-I; ref. 8). Translation starts from the first in-frame methionine found in an open reading frame of 5046 nucleotides. Identical amino acid residues are represented as dashes in RB-II and RH-I, and gaps were inserted to maximize the alignment. The suggested locations of the 24 putative membrane-spanning regions are indicated with bold lines above the sequence. Potential extracellular consensus N-linked glycosylation sites are indicated by filled circles. Potential sites for phosphorylation by cyclic nucleotide-dependent protein kinase are indicated by asterisks.

inactivation (34), is not as well conserved in hNav2.1 as in other NaChs, and there are differences in the number and distribution of charged residues.

There are five potential sites for N-linked glycosylation in regions of hNav2.1 predicted to be extracellular. Like other NaChs, these potential sites are clustered in the S5-S6 interhelical regions of domains 1 and 3. Potential sites for cyclic nucleotide-dependent phosphorylation are present in

the ID1-2 (Ser-442) and ID2-3 (Thr-777, Ser-869, and Ser-905) regions of hNav2.1. The density of these sites in the ID1-2 region is much less than in rat brain NaChs (1). A potential site for phosphorylation by protein kinase C that has been shown to affect inactivation kinetics in the rat brain IIA NaCh isoform (34) is absent.

**Tissue Distribution of hNav2.1.** Steady-state levels of hNav2.1 transcripts were examined in various tissues by

Table 1. Amino acid sequence homology between hNav2.1 and other NaCh isoforms

| NaCh isoform* | % amino acid identity with hNav2.1† |    |       |    |       |    |       |    |    |
|---------------|-------------------------------------|----|-------|----|-------|----|-------|----|----|
|               | N                                   | D1 | ID1-2 | D2 | ID2-3 | D3 | ID3-4 | D4 | C  |
| RB-I          | 43                                  | 53 | 20    | 59 | 16    | 58 | 50    | 57 | 43 |
| RB-II         | 43                                  | 52 | 18    | 60 | 33    | 57 | 50    | 59 | 46 |
| RB-III        | 33                                  | 52 | 18    | 59 | 28    | 56 | 50    | 59 | 46 |
| SkM1          | 40                                  | 51 | 14    | 60 | 4     | 58 | 50    | 54 | 41 |
| RH-I          | 37                                  | 48 | 13    | 56 | 7     | 56 | 54    | 54 | 41 |
| CSC-1         | —                                   | —  | —     | —  | —     | —  | —     | —  | 41 |
| Eel           | 32                                  | 45 | 16    | 58 | 9     | 54 | 41    | 49 | 26 |
| Para          | 21                                  | 38 | 5     | 39 | 4     | 43 | 17    | 40 | —  |

\*RB, rat brain (I and II, ref. 3; III, ref. 4); SkM1, adult rat skeletal muscle (5); RH-I, rat heart (8); CSC-1, rat cardiac sodium channel (partial cDNA sequence, ref. 6); Eel, *Electrophorus electricus* electric organ (9); Para, *Drosophila para* locus (11).

†Only exact matches were considered. Gaps were considered as a single nonidentity independent of length. N, amino terminus; D, domain; ID, interdomain region; C, carboxyl terminus.

Northern blot analysis (Fig. 5). A 7.8-kb transcript was detected in human heart (both cardiomyopathic and normal) and human uterus. Signals were also present in RNA from adult skeletal muscle and cultured rhabdomyosarcoma (RD) cells, but no specific hybridization was observed in RNA from human liver, cultured vascular smooth muscle cells, or retinoic acid-treated cultured neuroblastoma (LAN) cells. Faint signals were detected in brain, kidney, and spleen RNA.

## DISCUSSION

Voltage-dependent NaChs comprise a multigene family with at least six structurally distinct isoforms known to exist in mammalian brain, skeletal muscle, and heart (3–8). We have deduced the primary structure of a protein from human heart and fetal skeletal muscle that is highly related to the NaCh gene family but does not appear to represent the human homolog of any previously cloned channel. This assertion is based on the recent observations that cloned human NaChs typically exhibit >90% overall amino acid identity with the corresponding sequences in rat (12, 13). The protein we designate hNav2.1 bears ≈50% overall sequence identity with rat NaChs and appears to represent a distinct NaCh isoform. The extent of homology between hNav2.1 and other NaChs is reminiscent of that among different gene subfamilies of voltage-dependent potassium channels (35). By analogy, hNav2.1 may belong to a distinct subfamily of voltage-dependent NaCh genes.

Although the functional diversity of NaChs in excitable membranes is not as extensive as that for voltage-dependent potassium channels, there is evidence for the existence of functional phenotypes not clearly represented by previously identified NaCh isoforms. Specifically, a class of NaChs

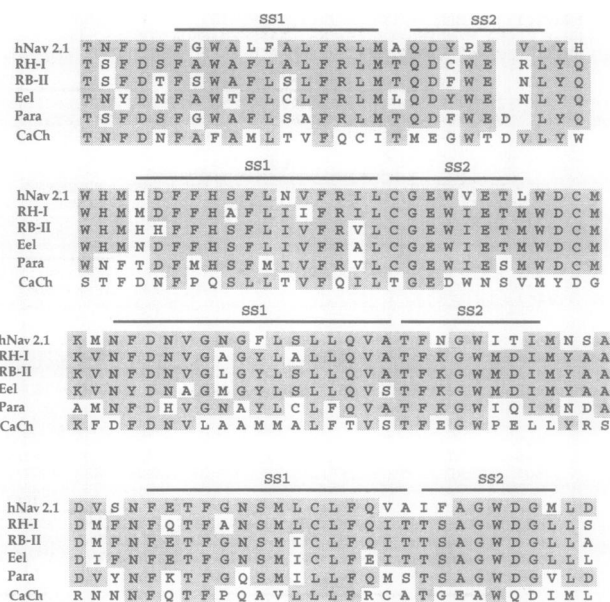


FIG. 3. Comparison of S5–S6 interhelical regions of hNav2.1 to other voltage-dependent ion channels: RH-I, rat heart I NaCh (8); RB-II, rat brain II NaCh (3); Eel, *Electrophorus electricus* electric organ NaCh (9); Para, *Drosophila para* locus NaCh (11); CaCh, rabbit heart dihydropyridine receptor  $\alpha_1$  subunit (30). Identical amino acid residues are shaded. Suggested locations of the putative ion pore-forming segments SS1 and SS2 are indicated by horizontal lines. Domains 1–4 are listed top to bottom.

exhibiting activation and inactivation kinetics significantly slower than that responsible for the fast inward sodium current of the action potential has been described in heart (17–20), glia (36), some neuronal preparations (37), and pathologic conditions of cardiac (38) and skeletal (39) muscle. The functional attributes of these NaChs are quite distinct from channels that have been expressed from cDNA and indicate the existence of a distinct class of NaChs. The relationship between this functional class of NaChs and the hNav2.1 channel is not clear, but certain structural features of this protein are consistent with novel functional properties.

There are several features of the hNav2.1 primary sequence that distinguish it from other NaChs. The length of hNav2.1 (1682 residues) is significantly less than that of other NaChs (1819–2018 residues), and the basis for this difference is not limited to an isolated region of the sequence. The most striking difference between hNav2.1 and other NaChs exists in the S4 segments. Although this motif is clearly present in hNav2.1, the absolute number of positively charged residues is less than in other NaChs. Interestingly, histidine residues replace two of the highly conserved S4 segment arginines in

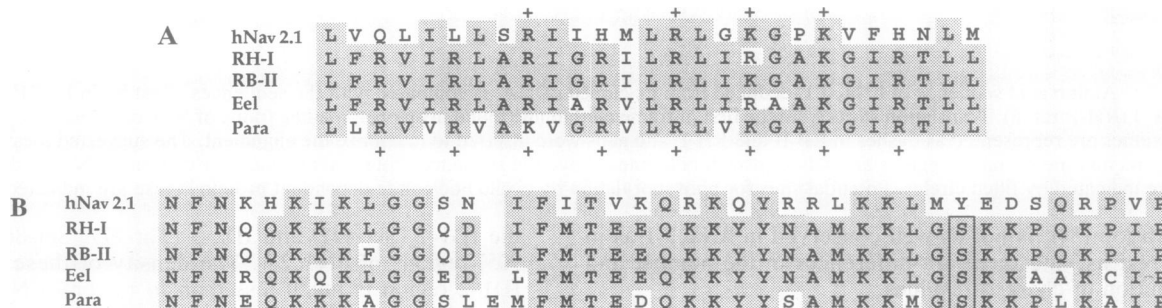


FIG. 4. Comparison of primary sequence of structures involved in activation and inactivation. (A) Alignment of S4 segments of domain 4 for hNav2.1 and other NaChs (see Fig. 3 legend). Identical amino acid residues are shaded. Positively charged residues are indicated (+). (B) Alignment of the interdomain 3–4 region for hNav2.1 and other NaChs. Identical amino acid residues are shaded. The rectangle indicates the position of a highly conserved site for phosphorylation by protein kinase C that is absent from hNav2.1.

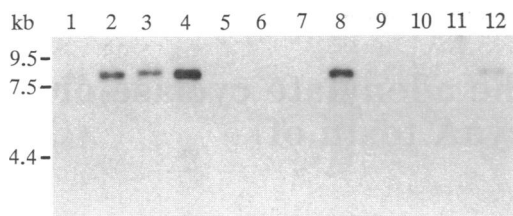


FIG. 5. Northern blot analysis of the distribution of hNav2.1 transcripts in human tissues. Lanes: 1, adult skeletal muscle; 2, cardiomyopathic right atrium; 3, cardiomyopathic left ventricle; 4, normal left ventricle; 5, brain (temporal lobe cortex); 6, kidney; 7, liver; 8, uterus; 9, spleen; 10, vascular smooth muscle cells; 11, retinoic acid-treated neuroblastoma (LAN) cells; 12, rhabdomyosarcoma (RD) cells. Locations of mRNA size standards electrophoresed in the same gel are indicated at left. The size of the hNav2.1 transcript is 7.8 kb.

domain 4. The electrostatic charge present on histidine at physiological pH ( $pK_a \approx 6.0$ ) is expected to be highly dependent upon the environment and it is difficult to predict precisely the effect this amino acid substitution will have on voltage sensing. The functional consequences of replacing arginine residues by glutamine in the S4 segment of domain 1 has been examined by Stühmer *et al.* (40) by site-directed mutagenesis of the rat brain II NaCh. Neutralization of S4-segment positive charge in domain 1 results in decreased steepness in the voltage dependence of activation and causes a shift in the current-voltage relationship toward more positive potentials. Similarly, certain voltage-dependent potassium channels have differences in both the number of S4 segment basic residues and in the slope of the voltage dependence of activation (35). These findings suggest that hNav2.1 may encode a NaCh with altered activation kinetics.

The tissue distribution of hNav2.1 mRNA transcripts suggests that this NaCh is prominently expressed in adult human heart and uterus. We infer that hNav2.1 is also expressed in fetal skeletal muscle, since it was cloned from this tissue, and low levels of the transcript are present in adult skeletal muscle. The finding of a NaCh mRNA in uterus was unexpected. The existence of functional NaChs in uterine smooth muscle has been demonstrated electrophysiologically (41), but the structural nature of these channels has not been determined. Additional studies are needed to define the role of hNav2.1 in these events. Interestingly, cardiac muscle and myometrium also share an unusual, voltage-dependent potassium channel (42).

We thank Deborah Gilmore, Melissa House, Begonia Tome, and Berndt Fakler for technical assistance and Drs. Steven Roberds, David Lovinger, Todd Verdoorn, Roland Kallen, and Robert Barchi for advice and critical review of the manuscript. This research was supported by National Institutes of Health Grant GM41325 (M.M.T.) and by grants-in-aid from the American Heart Association (National Chapter and Southeastern Pennsylvania Affiliate). M.M.T. is an Established Investigator of the American Heart Association. A.L.G. is a Lucille P. Markey Scholar, and this work was supported in part by a grant from the Lucille P. Markey Charitable Trust.

- Catterall, W. A. (1988) *Science* **242**, 50–61.
- Trimmer, J. S. & Agnew, W. S. (1989) *Annu. Rev. Physiol.* **51**, 401–418.
- Noda, M., Ikeda, T., Kayano, T., Suzuki, H., Takeshima, H., Kurasaki, M., Takahashi, H. & Numa, S. (1986) *Nature (London)* **320**, 188–192.
- Kayano, T., Noda, M., Flockerzi, V., Takahashi, H. & Numa, S. (1988) *FEBS Lett.* **228**, 187–194.
- Trimmer, J. S., Cooperman, S. S., Tomiko, S. A., Zhou, J., Crean, S. M., Boyle, M. B., Kallen, R. G., Sheng, Z., Barchi, R. L., Sigworth, F. J., Goodman, R. H., Agnew, W. S. & Mandel, G. (1989) *Neuron* **3**, 33–49.
- Sills, M. N., Xu, Y. C., Baracchini, E., Goodman, R. H.,

- Cooperman, S. S., Mandel, G. & Chien, K. R. (1989) *J. Clin. Invest.* **84**, 331–336.
- Kallen, R. G., Sheng, Z., Yang, J., Chen, L., Rogart, R. B. & Barchi, R. L. (1990) *Neuron* **4**, 233–242.
- Rogart, R. B., Cribbs, L. L., Muglia, L. K., Kephart, D. D. & Kaiser, M. W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8170–8174.
- Noda, M., Shimizu, S., Tanabe, T., Takai, T., Kayano, T., Ikeda, T., Takahashi, H., Nakayama, H., Kaknaoka, Y., Minamino, N., Kikawa, K., Matsuo, H., Raftery, M. A., Hirose, T., Inayama, S., Hayashida, H., Miyata, T. & Numa, S. (1984) *Nature (London)* **312**, 121–127.
- Salkoff, L., Butler, A., Wei, A., Scavarda, N., Giffen, K., Ifune, C., Goodman, R. & Mandel, G. (1987) *Science* **237**, 744–749.
- Loughney, K., Kreber, R. & Ganetzky, B. (1989) *Cell* **58**, 1143–1154.
- Gellens, M. E., George, A. L., Chen, L., Chahine, M., Horn, R., Barchi, R. L. & Kallen, R. G. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 554–558.
- George, A. L., Komisarof, J., Kallen, R. G. & Barchi, R. L. (1992) *Ann. Neurol.* **31**, 131–137.
- Hille, B. (1992) *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA), pp. 525–544.
- Brown, A. M., Lee, K. S. & Powell, T. (1981) *J. Physiol. (London)* **318**, 455–477.
- Rogart, R. B. (1986) *Ann. N.Y. Acad. Sci.* **479**, 402–430.
- Bkaily, G., Jacques, D., Yamamoto, T., Scultoreanu, A. & Payet, M. D. (1988) *Can. J. Physiol. Pharmacol.* **66**, 1017–1022.
- Coraboeuf, E. (1980) in *The Slow Inward Current and Cardiac Arrhythmias*, eds. Zipes, D. P., Bailey, J. C. & Elharrar, V. (Nijhoff, The Hague, The Netherlands), pp. 25–95.
- Bkaily, G., Jacques, D., Scultoreanu, A., Yamamoto, T., Carrier, D., Vigneault, D. & Sperelakis, N. (1991) *J. Mol. Cell. Cardiol.* **23**, 25–39.
- Lee, K. S. (1990) *J. Mol. Cell. Cardiol.* **22** (Suppl. I), S.15 (abstr.).
- Chandy, K. G. (1991) *Nature (London)* **352**, 26.
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
- Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 1408–1412.
- Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
- Knoth, K., Roberds, S., Potet, C. & Tamkun, M. (1988) *Nucleic Acids Res.* **16**, 10932.
- Tamkun, M. M., Knoth, K. M., Walbridge, J. A., Kroemer, H., Roden, D. M. & Glover, D. M. (1991) *FASEB J.* **5**, 331–337.
- Koenig, M., Hoffman, E. P., Bertelson, C. J., Monaco, A. P., Feener, C. & Kunkel, L. M. (1987) *Cell* **50**, 509–517.
- Henikoff, S. (1984) *Gene* **28**, 351–359.
- Kozak, M. (1984) *Nucleic Acids Res.* **12**, 857–872.
- Mikami, A., Imoto, K., Tanabe, T., Niidome, T., Mori, Y., Takeshima, H., Narumiya, S. & Numa, S. (1989) *Nature (London)* **340**, 230–233.
- Guy, H. R. & Conti, F. (1990) *Trends Neurosci.* **13**, 201–206.
- Terlau, H., Heinemann, S. H., Stühmer, W., Pusch, M., Conti, F., Imoto, K. & Numa, S. (1991) *FEBS Lett.* **293**, 93–96.
- Thomsen, W. J. & Catterall, W. A. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 10161–10165.
- West, J. W., Numann, R., Murphy, B. J., Scheuer, T. & Catterall, W. A. (1991) *Science* **254**, 866–868.
- Wei, A., Covarrubias, M., Butler, A., Baker, K., Pak, M. & Salkoff, L. (1990) *Science* **248**, 599–603.
- Barres, B. A., Chun, L. L. Y. & Corey, D. P. (1989) *Neuron* **2**, 1375–1388.
- Llinas, R. R. (1988) *Science* **242**, 1654–1664.
- Jacques, D. & Bkaily, G. (1991) *Biophys. J.* **59**, 259a (abstr.).
- Bkaily, G., Jasmin, G., Tautu, C., Prochek, L., Yamamoto, T., Scultoreanu, A., Peyrow, M. & Jacques, D. (1990) *Muscle Nerve* **13**, 939–948.
- Stühmer, W., Conti, F., Suzuki, H., Wang, X., Noda, M., Yahagi, N., Kubo, H. & Numa, S. (1989) *Nature (London)* **339**, 597–603.
- Ohya, Y. & Sperelakis, N. (1989) *Am. J. Physiol.* **257**, C408–C412.
- Folander, K., Smith, J. S., Antanavage, J., Bennett, C., Stein, R. B. & Swanson, R. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 2975–2979.