

Primary structure of rat brain sodium channel III deduced from the cDNA sequence

Toshiaki Kayano, Masaharu Noda, Veit Flockerzi*, Hideo Takahashi and Shosaku Numa

Departments of Medical Chemistry and Molecular Genetics, Kyoto University Faculty of Medicine, Kyoto 606, Japan

Received 12 December 1987

The complete amino acid sequence of a third sodium channel (designated sodium channel III) from rat brain has been deduced by cloning and sequence analysis of the cDNA. This protein is homologous in amino acid sequence and shares characteristic structural features with other sodium channels.

Na⁺ channel; cDNA cloning; Nucleotide sequence; Voltage-gated ionic channel; (Rat brain)

1. INTRODUCTION

The sodium channel is a transmembrane protein responsible for the voltage-dependent modulation of the sodium ion permeability of excitable membranes and thus plays an essential role in generating action potentials [1]. We have previously elucidated the complete amino acid sequences of the sodium channel from the electric organ of the eel *Electrophorus electricus* [2] and two distinct sodium channels (designated sodium channels I and II) from rat brain [3] by cloning and sequence analysis of the cDNAs. We have also cloned part of a third type of rat brain sodium channel cDNA encoding a sequence of 573 amino acids that is highly homologous with the carboxy-terminal sequences of sodium channels I and II [3]. The present investigation deals with the isolation of the

whole protein-coding sequence of the third type of cDNA and with the complete amino acid sequence of this novel sodium channel (designated sodium channel III) deduced from the cDNA sequence. The accompanying paper [4] concerns the functional expression of the cloned sodium channel III cDNA and the tissue distribution of the mRNAs encoding rat sodium channels I, II and III.

2. MATERIALS AND METHODS

Total RNA was extracted from the whole brain of male Wistar rats (~200 g body wt) by the guanidinium thiocyanate method [5], and poly(A)⁺ RNA was isolated by oligo(dT)-cellulose chromatography [6]. The procedures used to clone transcripts formed by extension of a synthetic oligodeoxyribonucleotide primer into the *Pst*I site of the plasmid pBR322 have been described [7]. Five primers (I–V) were prepared using an automatic DNA synthesizer (Applied Biosystems): I, 5'-CAGGGTCATGTATT-3' (complementary to nucleotides 4507–4521); II, 5'-GTTGCCTTCTTGGAT-3' (2968–2982); III, 5'-TGGGAAACCTGTCTC-3' (1559–1573); IV, 5'-TAGTTGGGGTTTCGT-3' (1068–1082); and V, 5'-CTAACA-GGGTTAGC-3' (339–353); 1 nmol of primer I, II or III or 3 nmol of primer IV was extended on 200 µg of rat brain poly(A)⁺ RNA, or 4 nmol of primer V on 300 µg of poly(A)⁺ RNA. The probes (A–H) used for screening the cDNA libraries constructed by primer extension were as follows: A, *Pst*II(4158)/*Bst*EII(4474) fragment from clone prSCH203; B, *Hpa*II(4027)/*Pst*I(4158) fragment from prSCH306; C, *Fnu*DII(2873)/*Hpa*II(3081) fragment from prSCH321; D, *Fnu*4HI(2618)/*Fnu*4HI(2832) fragment from prSCH331; E,

Correspondence address: S. Numa, Departments of Medical Chemistry and Molecular Genetics, Kyoto University Faculty of Medicine, Yoshida, Sakyo-ku, Kyoto 606, Japan

* *Present address:* Physiologische Chemie, Medizinische Fakultät der Universität des Saarlandes, D-6650 Homburg/Saar, FRG

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00766

*Hae*III(1298)/*Hpa*II(1504) fragment from prSCH417; F, *Sau*3AI(997)/*Rsa*I(1049) fragment from prSCH616; G, *Dde*I(644)/*Dde*I(773) fragment from prSCH705; and H, *Ava*II(235)/*Hin*II(303) fragment from prSCH722; probes A, C, D and E were labelled by nick-translation [8] with [α - 32 P]dCTP, and probes B, F, G and H were labelled at the 5'-end with [γ - 32 P]ATP. *Escherichia coli* strain HB101 or MC1061 was used for transformation [9], and transformants were screened at 60°C as in [10]. DNA sequencing was carried out according to [11].

3. RESULTS AND DISCUSSION

Clone prSCH203, which was isolated [3] from the Okayama-Berg library [12] derived from rat brain poly(A)⁺ RNA, carries a cDNA sequence (nucleotides 4133–6409; for nucleotide numbers, see fig.2) encoding the carboxy-terminal 573 amino acid residues of sodium channel III and the 3'-noncoding region (for all cDNA clones, see fig.1). For cloning cDNA sequences further upstream, a synthetic oligodeoxyribonucleotide primer complementary to a sequence in the 5'-portion of the prSCH203 cDNA (primer I; for all primers, see section 2 and fig.1) was extended by reverse transcriptase using rat brain poly(A)⁺ RNA as template. The single-stranded cDNA obtained was converted to double-stranded cDNA, which was cloned in the plasmid pBR322. Screening ($\sim 1.5 \times 10^5$ transformants) of the resulting cDNA library with a probe derived from the 5'-terminal region of prSCH203 (probe A; for all probes, see section 2 and fig.1) yielded 3 hybridization-positive clones including prSCH306 (carrying nucleotides 3964–4521). The same filters were rescreened with probe B derived from prSCH306 to give prSCH321 (nucleotides 2866–4092) and subsequently with probe C derived from prSCH321 to give prSCH331 (nucleotides 2586–3478). A second cDNA library constructed by extension of primer II was next screened ($\sim 9 \times 10^4$ transformants) with probe D derived from prSCH331 to harvest 18 positive clones including prSCH417 (nucleotides 1292–2981). Screening ($\sim 6 \times 10^5$ transformants) of a third cDNA library prepared by extension of primer III with probe E derived from prSCH417 yielded 11 positive clones including prSCH616 (nucleotides 927–1547) and prSCH628 (nucleotides 933–1569). A fourth cDNA library resulting from extension of primer IV was then screened ($\sim 2.4 \times 10^5$ transformants)

with probe F derived from prSCH616 to give 5 positive clones including prSCH705 (nucleotides 598–1081). A second screening ($\sim 1.1 \times 10^5$ transformants) of the same library with probe G derived from prSCH705 yielded 6 positive clones including prSCH722 (nucleotides 224–900). Finally, a fifth cDNA library constructed by extension of primer V was screened ($\sim 2.4 \times 10^5$ transformants) with probe H derived from prSCH722 to give 4 positive clones including prSCH801 (nucleotides –413 to 349).

Fig.2 shows the 6822-nucleotide cDNA sequence (excluding the poly(dA) tract) for rat brain sodium channel III, determined with the clones given in fig.1. The primary structure of rat sodium channel III was deduced by using the reading frame corresponding to the amino acid sequences of rat sodium channels I and II [3]. The translational initiation site was assigned to the methionine codon composed of nucleotide residues 1–3 because this is the first ATG triplet that appears downstream of a nonsense codon, TAA (nucleotides –24 to –22), found in frame. A translational termination codon (TAA) occurs in frame after codon 1951 specifying lysine. Thus, rat sodium channel III consists of 1951 amino acid residues (including the initiating methionine) and has a calculated M_r of 221375; at positions where amino acid differences are predicted by the nucleotide differences found among the individual clones (see fig.2), the amino acid residues given beneath have been used to calculate the M_r . RNA blot hybridization analysis shows that rat brain contains a major mRNA species of ~ 9000 nucleotides and a minor mRNA species of ~ 7500 nucleotides that are hybridizable with a sodium channel III-specific probe (see fig.5B, panel c, lane 2 in [4]). Clone prSCH203 has a 3'-noncoding sequence (556 nucleotides) much shorter than those of the cDNAs encoding rat sodium channels I and II (2120 and 2328 nucleotides, respectively [3]) and contains the polyadenylation signal AATAAA [13] (nucleotides 6385–6390) 20 nucleotides upstream of the poly(dA) tract. This, together with the sizes of the major mRNA species encoding sodium channels I and II (~ 9000 and ~ 9500 nucleotides, respectively [3,4]), may suggest that the two mRNA species encoding sodium channel III arise from polyadenylation at different sites and that clone prSCH203 is derived from the shorter mRNA species.

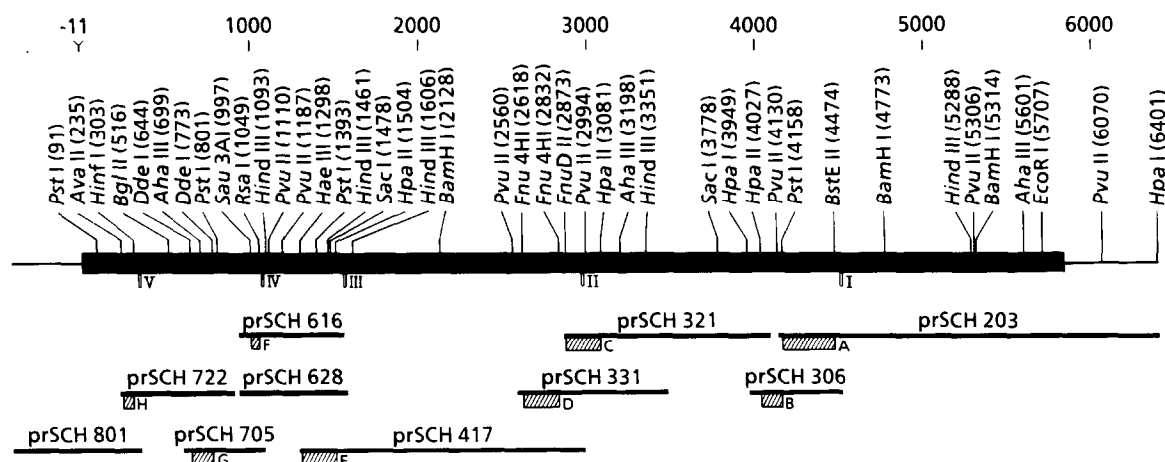


Fig.1. Cloning and restriction map of cDNA encoding rat sodium channel III. All existing sites for the six-base-specific restriction endonucleases shown and the four- and five-base-specific restriction endonuclease sites relevant to the hybridization probes used are presented and identified by numbers indicating the 5'-terminal nucleotide generated by cleavage. The protein-coding region is indicated by a closed box, the synthetic oligonucleotide primers (I-V) used for reverse transcription by small open boxes, and the hybridization probes (A-H) used for screening cDNA libraries by hatched boxes. The extent of cDNA inserts of the individual clones used for nucleotide sequence analysis are shown by thick lines; the poly(dG)-poly(dC) tails and the poly(dA)-poly(dT) tract are not included.

Fig.3 shows the alignment of the amino acid sequences of rat sodium channels I, II and III and the *Electrophorus* sodium channel. The degree of amino acid sequence homology is 87, 85, 87, 61, 62 and 61% for the rat I/rat II, rat I/rat III, rat II/rat III, rat I/*Electrophorus*, rat II/*Electrophorus* and rat III/*Electrophorus* pairs, respectively; for evaluating amino acid sequence homology, a continuous stretch of gaps has been counted as one substitution regardless of its length. Similarity matrix analysis [16] reveals that sodium channel III, like the three other sodium channels, contains four repeated units of homology (positions 111-456, 776-1048, 1240-1554 and 1563-1861 in the aligned sequences). The regions corresponding to the internal repeats are highly conserved among the four sodium channels, whereas the remaining regions, all of which are assigned to the cytoplasmic side of the membrane (see below), are less well conserved, except the short segment between repeats III and IV. Sodium channel III is also similar in hydropathy profile [17] to the three other sodium channels [2,3,14]. Each internal repeat has five hydrophobic segments (S1, S2, S3, S5 and S6) and one positively charged segment (S4), all of which exhibit predicted secondary structure [18]. Thus, the sodium channel III

molecule is assumed to have the same transmembrane topology as proposed for the other sodium channels [2,3] in that the four repeated units of homology, each containing the six presumably α -helical membrane-spanning segments S1-S6, are oriented in a pseudosymmetric fashion across the membrane and in that the amino- and carboxy-termini reside on the cytoplasmic side of the membrane. The proposed transmembrane topology is consistent with five of the six potential *N*-glycosylation sites [19,20] that are conserved in all the four sodium channels (asparagine residues at positions 212, 297, 340, 1418 and 1432; see fig.3) as well as with all the potential cyclic AMP-dependent phosphorylation sites [21] that are conserved in the three rat sodium channels (serine residues at positions 249, 574, 593, 630, 643, 718, 719 and 1556; see fig.3). Sodium channel III, like sodium channels I and II, has a large insertion in the region between repeats I and II, as compared with the *Electrophorus* sodium channel, but the insertion is somewhat shorter in sodium channel III (135 amino acids) than in sodium channels I (194 amino acids) and II (182 amino acids). The inserted segments and their carboxy-terminal neighbouring regions of the three rat sodium channels contain six of the eight conserved potential

[illegible]

Fig.2. Nucleotide sequence of cloned cDNA encoding rat sodium channel III. Nucleotide residues are numbered in the 5'- to 3'-direction, beginning with the first residue of the ATG triplet encoding the initiating methionine, and the nucleotides on the 5'-side of residue 1 are indicated by negative numbers; numbers of the nucleotide residues at the right-hand end of individual lines are given. The deduced amino acid sequence of rat sodium channel III is shown above the nucleotide sequence, and amino acid residues are numbered beginning with the initiating methionine. The 5'-terminal sequence presented does not extend to the 5'-end of the mRNA. The 3'-terminal sequence shown is followed by a poly(dA) tract connected with the vector DNA sequence [12]. The nucleotide differences observed among the individual clones are as follows: G (prSCH705) or A (prSCH722) at nucleotide 810; G (prSCH705) or A (prSCH722) at 819; C (prSCH705) or T (prSCH722) at 833; G (prSCH705) or A (prSCH722) at 855; G (prSCH705 and prSCH628) or A (prSCH616) at 1063; G (prSCH616 and prSCH628) or A (prSCH417) at 1374; G (prSCH616 and prSCH417) or A (prSCH628) at 1537; G (prSCH321) or A (prSCH331) at 3175; G (prSCH321) or A (prSCH331) at 3246; G (prSCH306) or A (prSCH321) at 4092. The resulting amino acid substitutions are also shown.

	1	50	
Rat sodium channel I	MEQTVLPVPPGDSFNFRTRESLAATERIAEAKAKPKPKKKDDENGPKPNSOLEAGKMLPFYIGDIPPEMVSEPLEDPPYINKKTFIVL		93
Rat sodium channel II	MARSVLVPPGDSFNFRTRESLAATERIAEAKAKPKPKKKDDENGPKPNSOLEAGKMLPFYIGDIPPEMVSEPLEDPPYINKKTFIVL		94
Rat sodium channel III	MAQALLVPPGDSFNFRTRESLAATERIAEAKAKPKPKKKDDENGPKPNSOLEAGKMLPFYIGDIPPEMVSEPLEDPPYINKKTFIVL		93
Electrophorus sodium channel	MARKFSSARHEMFRRTFDSLEETEAFTLEKLCSTLEKK---EPESPTPTOLEAGKPLPFYIGDIPPEMLTIPLEDDPPYINKKTFIVL		87
	100	150	200
	NKGKATIRFSATSALYILTPFNPLRKAIAIKLVHSLFSLIMCTILINCVMFTMSPPDWTKNVEYTFGTGYTFESLKIILARGFCLEDTFLRDPNNWLDHITVITFAVYTFEVLGNVS		213
	NKGKATIRFSATSALYILTPFNPLRKAIAIKLVHSLFSLIMCTILINCVMFTMSPPDWTKNVEYTFGTGYTFESLKIILARGFCLEDTFLRDPNNWLDHITVITFAVYTFEVLGNVS		214
	NKGKATIRFSATSALYILTPFNPLRKAIAIKLVHSLFSLIMCTILINCVMFTMSPPDWTKNVEYTFGTGYTFESLKIILARGFCLEDTFLRDPNNWLDHITVITFAVYTFEVLGNVS		213
	SKGNTIMRFAALYIFSPFNPIKRGARIRVVISAFNFIHFTIFSNQIFMTIISNPPASIKITVEYTFGTGYTFESLKIILARGFCLEDTFLRDPNNWLDHITVITFAVYTFEVLGNVS		207
	151	152	153
	250	300	
	ALRTFRVLRAKLTISVIPGLKTIYVGLIQSVKLSQVMILTVFCLSVFALIGLQFMGNLRNKLQWPPDMSFEINITSFFNNSLDMNGTAFNRTNMFNNDEYIEDKSHFYFLEGQND		332
	ALRTFRVLRAKLTISVIPGLKTIYVGLIQSVKLSQVMILTVFCLSVFALIGLQFMGNLRNKLQWPPDMSFEINITSFFNNSLDMNGTAFNRTNMFNNDEYIEDKSHFYFLEGQND		334
	ALRTFRVLRAKLTISVIPGLKTIYVGLIQSVKLSQVMILTVFCLSVFALIGLQFMGNLRNKLQWPPDMSFEINITSFFNNSLDMNGTAFNRTNMFNNDEYIEDKSHFYFLEGQND		333
	ALRTFRVLRAKLTISVIPGLKTIYVGLIQSVKLSQVMILTVFCLSVFALIGLQFMGNLRNKLQWPPDMSFEINITSFFNNSLDMNGTAFNRTNMFNNDEYIEDKSHFYFLEGQND		311
	154	155	
	350	400	450
	ALLCGNSDAGOCPEGYMCVKAGRNPNYGYTSFDTFSWAFSLFRMLTQDFWENLYQLTLRAAGKTYMIFFLVIFLGSFYLINLILAVVAMAYEEQNGATLEAEQKEAEFQOMLEQLK		452
	ALLCGNSDAGOCPEGYMCVKAGRNPNYGYTSFDTFSWAFSLFRMLTQDFWENLYQLTLRAAGKTYMIFFLVIFLGSFYLINLILAVVAMAYEEQNGATLEAEQKEAEFQOMLEQLK		454
	ALLCGNSDAGOCPEGYMCVKAGRNPNYGYTSFDTFSWAFSLFRMLTQDFWENLYQLTLRAAGKTYMIFFLVIFLGSFYLINLILAVVAMAYEEQNGATLEAEQKEAEFQOMLEQLK		453
	ALLCGNSDAGOCPEGYMCVKAGRNPNYGYTSFDTFSWAFSLFRMLTQDFWENLYQLTLRAAGKTYMIFFLVIFLGSFYLINLILAVVAMAYEEQNGATLEAEQKEAEFQOMLEQLK		431
	156		
	500	550	
	KQOEIAQAQAAATASEHSREPSA-----GLSDSSSEASKLSSKSAKERNNRRKKKKKKQKQSGGEEKODDEFHKSSEESIRKKGRFSLIEGNRLTYEKRYSS		551
	KQOEIAQAQAAATASEHSREPSA-----GLSDSSSEASKLSSKSAKERNNRRKKKKKKQKQSGGEEKODDEFHKSSEESIRKKGRFSLIEGNRLTYEKRYSS		554
	KQOEIAQAQAAATASEHSREPSA-----GLSDSSSEASKLSSKSAKERNNRRKKKKKKQKQSGGEEKODDEFHKSSEESIRKKGRFSLIEGNRLTYEKRYSS		554
	KQOEIAQAQAAATASEHSREPSA-----GLSDSSSEASKLSSKSAKERNNRRKKKKKKQKQSGGEEKODDEFHKSSEESIRKKGRFSLIEGNRLTYEKRYSS		487
	157		
	600	650	
	PHOSLLSIRGSLFSPRRNSRSLFPRERAKVSENFADDEHSTFEONERRDLSFVPRHGERPNSLSQTSRSSRMLAGLPANGKMHSTVDCNGVSVLVGGSPVSTSPVGQLLPEV		671
	PHOSLLSIRGSLFSPRRNSRSLFPRERAKVSENFADDEHSTFEONERRDLSFVPRHGERPNSLSQTSRSSRMLAGLPANGKMHSTVDCNGVSVLVGGSPVSTSPVGQLLPEV		673
	PHOSLLSIRGSLFSPRRNSRSLFPRERAKVSENFADDEHSTFEONERRDLSFVPRHGERPNSLSQTSRSSRMLAGLPANGKMHSTVDCNGVSVLVGGSPVSTSPVGQLLPEV		624
	158		
	700	750	800
	IIDKATDODGTTTETEMKRRSSSFHYSHDF---LEDPSRQRAMSAISILNTNTEELLESROKPPCWYKFSNIFLIWDQSPYLKVKHINLVLMVMPFDVLAITICIVLNTLFNAME		788
	IIDKATDODGTTTETEMKRRSSSFHYSHDF---LEDPSRQRAMSAISILNTNTEELLESROKPPCWYKFSNIFLIWDQSPYLKVKHINLVLMVMPFDVLAITICIVLNTLFNAME		779
	IIDKATDODGTTTETEMKRRSSSFHYSHDF---LEDPSRQRAMSAISILNTNTEELLESROKPPCWYKFSNIFLIWDQSPYLKVKHINLVLMVMPFDVLAITICIVLNTLFNAME		731
	IIDKATDODGTTTETEMKRRSSSFHYSHDF---LEDPSRQRAMSAISILNTNTEELLESROKPPCWYKFSNIFLIWDQSPYLKVKHINLVLMVMPFDVLAITICIVLNTLFNAME		586
	159	160	161
	850	900	
	HYPMTEFHHVLTIVGNLVFTGIFTAEMLFKIIAMDPYFFQEGWNIFDGLTVLSLVELGLANVEGLSVLSRFLRVFKLAKSWPTLNLKIKIIGNSVGALGNLTVLVAITVIFFAVG		908
	HYPMTEFHHVLTIVGNLVFTGIFTAEMLFKIIAMDPYFFQEGWNIFDGLTVLSLVELGLANVEGLSVLSRFLRVFKLAKSWPTLNLKIKIIGNSVGALGNLTVLVAITVIFFAVG		899
	HYPMTEFHHVLTIVGNLVFTGIFTAEMLFKIIAMDPYFFQEGWNIFDGLTVLSLVELGLANVEGLSVLSRFLRVFKLAKSWPTLNLKIKIIGNSVGALGNLTVLVAITVIFFAVG		851
	HYPMTEFHHVLTIVGNLVFTGIFTAEMLFKIIAMDPYFFQEGWNIFDGLTVLSLVELGLANVEGLSVLSRFLRVFKLAKSWPTLNLKIKIIGNSVGALGNLTVLVAITVIFFAVG		706
	162	163	164
	950	1000	1050
	MQLFGKSYKCYCKIADQCKLPRWHMDDFFHSFLIVFRVLCGEWETMMDCHVEVAGQMCILVFMVMYVIGNLVVLNLFALLSSFSADNLAATODDNEMNNLQIAYVRMKGAYVVR		1028
	MQLFGKSYKCYCKIADQCKLPRWHMDDFFHSFLIVFRVLCGEWETMMDCHVEVAGQMCILVFMVMYVIGNLVVLNLFALLSSFSADNLAATODDNEMNNLQIAYVRMKGAYVVR		1019
	MQLFGKSYKCYCKIADQCKLPRWHMDDFFHSFLIVFRVLCGEWETMMDCHVEVAGQMCILVFMVMYVIGNLVVLNLFALLSSFSADNLAATODDNEMNNLQIAYVRMKGAYVVR		971
	MQLFGKSYKCYCKIADQCKLPRWHMDDFFHSFLIVFRVLCGEWETMMDCHVEVAGQMCILVFMVMYVIGNLVVLNLFALLSSFSADNLAATODDNEMNNLQIAYVRMKGAYVVR		825
	165	166	167
	1100	1150	
	KIYEFITQSFVRKOKILDEIKPL-DEL-NNRKDNQISNHTT-ETIGRODOLKDVNGTTSQIGTGSSEVKYIIDEISDYMSFINNPSLTVTVPVIAVGESOFENLNTEDFSSSEOLEESKEKL		1145
	KIYEFITQSFVRKOKILDEIKPL-DEL-NNRKDNQISNHTT-ETIGRODOLKDVNGTTSQIGTGSSEVKYIIDEISDYMSFINNPSLTVTVPVIAVGESOFENLNTEDFSSSEOLEESKEKL		1135
	KIYEFITQSFVRKOKILDEIKPL-DEL-NNRKDNQISNHTT-ETIGRODOLKDVNGTTSQIGTGSSEVKYIIDEISDYMSFINNPSLTVTVPVIAVGESOFENLNTEDFSSSEOLEESKEKL		1084
	KIYEFITQSFVRKOKILDEIKPL-DEL-NNRKDNQISNHTT-ETIGRODOLKDVNGTTSQIGTGSSEVKYIIDEISDYMSFINNPSLTVTVPVIAVGESOFENLNTEDFSSSEOLEESKEKL		918
	168	169	170
	1200	1250	
	NESSSSSEGSTVDIGAPA-----E-EQVMEPEE---TLEPEACFTTEGCVORFKCKQISVEEGRGKNNWNLRRITCFRIVEHNNWFETTFIVMILLSSGALAFEDIYIDORKTITKTMLEY		1254
	NESSSSSEGSTVDIGAPA-----E-EQVMEPEE---TLEPEACFTTEGCVORFKCKQISVEEGRGKNNWNLRRITCFRIVEHNNWFETTFIVMILLSSGALAFEDIYIDORKTITKTMLEY		1244
	NESSSSSEGSTVDIGAPA-----E-EQVMEPEE---TLEPEACFTTEGCVORFKCKQISVEEGRGKNNWNLRRITCFRIVEHNNWFETTFIVMILLSSGALAFEDIYIDORKTITKTMLEY		1193
	NESSSSSEGSTVDIGAPA-----E-EQVMEPEE---TLEPEACFTTEGCVORFKCKQISVEEGRGKNNWNLRRITCFRIVEHNNWFETTFIVMILLSSGALAFEDIYIDORKTITKTMLEY		1036
	171	172	173
	1300	1350	1400
	ADKVFTYTFIEMLLKWAYAGQYTFYFNACWLDLFDVDSLSLTANALGYSELGALKSLRTLALRPLRALSRFEGMRVYVYVALLGAIPSMNVLLVCLIFWLFISMGVNLFAGKFY		1374
	ADKVFTYTFIEMLLKWAYAGQYTFYFNACWLDLFDVDSLSLTANALGYSELGALKSLRTLALRPLRALSRFEGMRVYVYVALLGAIPSMNVLLVCLIFWLFISMGVNLFAGKFY		1364
	ADKVFTYTFIEMLLKWAYAGQYTFYFNACWLDLFDVDSLSLTANALGYSELGALKSLRTLALRPLRALSRFEGMRVYVYVALLGAIPSMNVLLVCLIFWLFISMGVNLFAGKFY		1313
	ADKVFTYTFIEMLLKWAYAGQYTFYFNACWLDLFDVDSLSLTANALGYSELGALKSLRTLALRPLRALSRFEGMRVYVYVALLGAIPSMNVLLVCLIFWLFISMGVNLFAGKFY		1156
	174	175	176
	1450	1500	
	HCYNITTTGDTFEITEVNNHSDCLKIERNETARWKNVKNFNDVNGFGLSLLOVATFKGWMIDMAYAVDSRMVLOPKYEECLMYLYFYVIFIFGSSFFTLNLTGVIIDNFNQKKKFG		1494
	HCYNITTTGDTFEITEVNNHSDCLKIERNETARWKNVKNFNDVNGFGLSLLOVATFKGWMIDMAYAVDSRMVLOPKYEECLMYLYFYVIFIFGSSFFTLNLTGVIIDNFNQKKKFG		1484
	HCYNITTTGDTFEITEVNNHSDCLKIERNETARWKNVKNFNDVNGFGLSLLOVATFKGWMIDMAYAVDSRMVLOPKYEECLMYLYFYVIFIFGSSFFTLNLTGVIIDNFNQKKKFG		1430
	HCYNITTTGDTFEITEVNNHSDCLKIERNETARWKNVKNFNDVNGFGLSLLOVATFKGWMIDMAYAVDSRMVLOPKYEECLMYLYFYVIFIFGSSFFTLNLTGVIIDNFNQKKKFG		1275
	177	178	179



Fig.3. Alignment of the amino acid sequences of rat sodium channels I (top), II (second row), III (third row) and the *Electrophorus electricus* sodium channel (bottom). The one-letter amino acid notation is used. The sequence data for rat sodium channels I and II and the *Electrophorus* sodium channel have been taken from [2,3]. At positions where amino acid differences occur (see fig.2 and [3,14]), the residues given beneath have been adopted. Sets of three or four identical residues at one position are enclosed with solid lines, and the fourth residues regarded as conservative substitutions at the same position are enclosed with broken lines. Conservative substitutions are defined as pairs of residues belonging to one of the following groups: S, T, P, A and G; N, D, E and Q; H, R and K; M, I, L and V; F, Y and W [15]. Gaps (—) have been inserted to achieve maximum homology. Amino acid residues are numbered beginning with the initiating methionine, and numbers of the residues at the right-hand end of individual lines are given. Positions in the aligned sequences including gaps are numbered beginning with that of the initiating methionine, and position numbers are given above the sequences. The putative transmembrane segments S1-S6 in each of repeats I-IV are indicated; the termini of these segments have been tentatively assigned.

tains a conserved aspartic acid residue (positions 195, 862, 1328 and 1648) at an equivalent position.

The above-mentioned charged residues in segments S2, S3 and S4 are fully conserved in the putative sodium channel from *Drosophila* as well [24], except that the aspartic acid residue in segment S2 of repeat III (position 1296) is replaced by an asparagine residue and that an additional lysine residue is present in segment S4 of repeat III (position 1350). It seems reasonable to assume that these conserved charged residues in the putative transmembrane segments are essential for the voltage-dependent operation of the sodium channel [2,3]. Furthermore, the cluster of positively charged residues (predominantly lysine) in the region between segment S6 of repeat III and segment S1 of repeat IV as well as the cluster of negatively charged residues in the region following segment S6 of repeat IV [3] is conserved in the vertebrate and insect sodium channels. It is conceivable that these regions, which are assigned to

the cytoplasmic side of the membrane, are involved in the inactivation of the sodium channel [3,25].

The structural features characteristic of the sodium channel are shared by the receptor for calcium channel blockers from rabbit skeletal muscle, which is thought to be a calcium channel [26]. A putative potassium channel component from *Drosophila* has a basic structure corresponding to a single repeat of the sodium channel [27]. These findings suggest that voltage-gated ionic channels represent a family of evolutionarily related gene products [26].

Acknowledgements: We thank Dr Takashi Miyata and Mr Kei-ichi Kuma for computer analysis. This investigation was supported in part by research grants from the Ministry of Education, Science and Culture and the Science and Technology Agency of Japan, the Mitsubishi Foundation and the Japanese Foundation of Metabolism and Diseases and by fellowships from the Japan Society for the Promotion of Science and the Feodor Lynen Program of the Alexander von Humboldt Foundation (V.F.).

REFERENCES

- [1] Hille, B. (1984) *Ionic Channels of Excitable Membranes*, Sinauer, Sunderland, MA.
- [2] Noda, M., Shimizu, S., Tanabe, T., Takai, T., Kayano, T., Ikeda, T., Takahashi, H., Nakayama, H., Kanaoka, Y., Minamino, N., Kangawa, K., Matsuo, H., Raftery, M.A., Hirose, T., Inayama, S., Hayashida, H., Miyata, T. and Numa, S. (1984) *Nature* 312, 121–127.
- [3] Noda, M., Ikeda, T., Kayano, T., Suzuki, H., Takeshima, H., Kurasaki, M., Takahashi, H. and Numa, S. (1986) *Nature* 320, 188–192.
- [4] Suzuki, H., Beckh, S., Kubo, H., Yahagi, N., Ishida, H., Kayano, T., Noda, M. and Numa, S. (1988) *FEBS Lett.* 228, 000–000.
- [5] Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) *Biochemistry* 18, 5294–5299.
- [6] Aviv, H. and Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* 69, 1408–1412.
- [7] Noda, M., Furutani, Y., Takahashi, H., Toyosato, M., Hirose, T., Inayama, S., Nakanishi, S. and Numa, S. (1982) *Nature* 295, 202–206.
- [8] Weinstock, R., Sweet, R., Weiss, M., Cedar, H. and Axel, R. (1978) *Proc. Natl. Acad. Sci. USA* 75, 1299–1303.
- [9] Hanahan, D. (1985) in: *DNA Cloning* (Glover, D.M. ed.) vol.1, pp.109–135, IRL Press, Oxford.
- [10] Hanahan, D. and Meselson, M. (1980) *Gene* 10, 63–67.
- [11] Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.* 65, 499–560.
- [12] Okayama, H. and Berg, P. (1982) *Mol. Cell. Biol.* 2, 161–170.
- [13] Proudfoot, N.J. and Brownlee, G.G. (1976) *Nature* 263, 211–214.
- [14] Numa, S. and Noda, M. (1986) *Ann. NY Acad. Sci.* 479, 338–355.
- [15] Dayhoff, M.O., Schwartz, R.M. and Orcutt, B.C. (1978) in: *Atlas of Protein Sequence and Structure* (Dayhoff, M.O. ed.) vol.5, suppl.3, pp.345–352, National Biomedical Research Foundation, Silver Springs, MD.
- [16] Toh, H., Hayashida, H. and Miyata, T. (1983) *Nature* 305, 827–829.
- [17] Kyte, J. and Doolittle, R.F. (1982) *J. Mol. Biol.* 157, 105–132.
- [18] Chou, P.Y. and Fasman, G.D. (1978) *Annu. Rev. Biochem.* 47, 251–276.
- [19] Hubbard, S.C. and Ivatt, R.J. (1981) *Annu. Rev. Biochem.* 50, 555–583.
- [20] Bause, E. (1983) *Biochem. J.* 209, 331–336.
- [21] Krebs, E.G. and Beavo, J.A. (1979) *Annu. Rev. Biochem.* 48, 923–959.
- [22] Costa, M.R.C., Casnellie, J.E. and Catterall, W.A. (1982) *J. Biol. Chem.* 257, 7918–7921.
- [23] Costa, M.R.C. and Catterall, W.A. (1984) *J. Biol. Chem.* 259, 8210–8218.
- [24] Salkoff, L., Butler, A., Wei, A., Scavarda, N., Giffen, K., Ifune, C., Goodman, R. and Mandel, G. (1987) *Science* 237, 744–749.
- [25] Armstrong, C.M. (1981) *Physiol. Rev.* 61, 644–683.
- [26] Tanabe, T., Takeshima, H., Mikami, A., Flockerzi, V., Takahashi, H., Kangawa, K., Kojima, M., Matsuo, H., Hirose, T. and Numa, S. (1987) *Nature* 328, 313–318.
- [27] Tempel, B.L., Papazian, D.M., Schwarz, T.L., Jan, Y.N. and Jan, L.Y. (1987) *Science* 237, 770–775.