Volume 15 Number 20 1987 Nucleic Acids Research

Nucleotide sequence of the putative sodium channel gene from *Drosophila*: the four homologous domains

Lawrence Salkoff, Alice Butler, Nancy Scavarda and Aguan Wei

Department of Anatomy and Neurobiology, Washington University School of Medicine, St Louis, MO 63110, USA

Submitted July 26, 1987 Accession no. Y00461

The Nucleotide sequence of all four homology domains of the presumptive sodium channel gene from Drosophila (1) has been determined. Additional 5′ and 3′ translated portions are also included. The deduced amino acid sequence of the Drosophila gene is highly conserved with respect to the alpha subunit of the sodium channel from electric eel (2) and rat brain (3). Like the vertebrate protein, each of the four homology domains of the Drosophila channel contains six proposed membrane spanning regions (S1 to S6). Ten introns interrupt the coding regions of the four homology domains.

```
         .         .       120         .         .       150         .         .       180
-------------------------------------------------------CTTGAGACCACATTGGCGGTCTCCAACTGGAGC
                                                       LeuGluThrThrLeuAlaValSerAsnTrpSer
                                                                                        60
         .         .       210         .         .       240         .         .       270
ACAGCAATCCTCTCCCAAACAGCACCCTGTATTAAGCGCGCTTTTTTCAAATTCCAGACCGTCTGTGTAGTATCTAAGCGGTTCCGGAAG
ThrAlaIleLeuSerGlnThrAlaProCysIleLysArgAlaPhePheLysPheGlnThrValCysValValSerLysArgPheArgLys
                                                                                        90
         .         .       300         .         .       330         .         .       360
AACTACATCCATCGCTTCACGGGCACCAAGAGCCTCTTCCTCTTCTACCCATGGAGCCCAGCAAGGCGCGTTTGCGTCTACATCGCCACC
AsnTyrIleHisArgPheThrGlyThrLysSerLeuPheLeuPheTyrProTrpSerProAlaArgArgValCysValTyrIleAlaThr
                                                                                       120
         .         .       390         .         .       420         .         .       450
AACCAGTTCTTCGACTACTGCGTCATGGCCACCATTCTGTTCAACTGCATCTTCCTGGCTATGACGGAGACCGTGGAAGAGGCGGAGTAC
AsnGlnPhePheAspTyrCysValMetAlaThrIleLeuPheAsnCysIlePheLeuAlaMetThrGluThrValGluGluAlaGluTyr
        S1a                                                                             150
         .         .       480         .         .       510         .         .       540
ATCTTTCTGGCCATTTACTCTATCGAAATGGTCATAAAGATTATTGCCAAGGGCTTTCTGCTCAACAAGTACACGTATCTGCGAAACCCG
IlePheLeuAlaIleTyrSerIleGluMetValIleLysIleIleAlaLysGlyPheLeuLeuAsnLysTyrThrTyrLeuArgAsnPro
                                                                                       180
         .         .       570         .         .       600         .         .       630
TGGAATTGGCTGGACTTCGTGGTCATTACTATGCGGTACGCCACCATCGGCATGGAGGTGGGCAACCTGGCGGGGCTGCGCACTTTTCGC
TrpAsnTrpLeuAspPheValValIleThrMetArgTyrAlaThrIleGlyMetGluValGlyAsnLeuAlaGlyLeuArgThrPheArg
                S2a                                                                     210
         .         .       660         .         .       690         .         .       720
GTGCTGCGCGCCCTCAAGACGGTGTCCATCATGCCCGGATTGAAGACGATCATCAGACCGTTGCTGCACTCCGTTGCCCAGCTGGCGGAG
ValLeuArgAlaLeuLysThrValSerIleMetProGlyLeuLysThrIleIleArgProLeuLeuHisSerValAlaGlnLeuAlaGlu
                S3a                                                                     240
         .         .       750         .         .       780         .         .       810
GTCATGACGCTGACCATCTTCTGTGTGATGGTCTTCGCGCTCTTTGCCCTCCAGGTCTACATGGGCGAGCTGCGCAATAAGTGCGTGCGC
ValMetThrLeuThrIlePheCysValMetValPheAlaLeuPheAlaLeuGlnValTyrMetGlyGluLeuArgAsnLysCysValArg
                S4a                                                                     270
         .         .       840         .         .       870         .         .       900
CAGGTTCCCACCGACTGGACGAACGTCTCTCACACGGACTGGCAGATCTGGGTCAACGACACCGACAACTGGCTATATGACGAGGATGAA
GlnValProThrAspTrpThrAsnValSerHisThrAspTrpGlnIleTrpValAsnAspThrAspAsnTrpLeuTyrAspGluAspGlu
                S5a                                                                     300
         .         .       930         .         .       960         .         .       990
CTGCCCGTGCTGTGTGGTAACTTAACTGGAGCCCGCCACTGTCCTTTCGAGTACGTGTGCCTCTGCGTAGGCGAGAATCCCAATCACGGC
LeuProValLeuCysGlyAsnLeuThrGlyAlaArgHisCysProPheGluTyrValCysLeuCysValGlyGluAsnProAsnHisGly
                                                                                       330
         .         .      1020         .         .      1050         .         .       57
TACACCAACTTTGACAACTTTATGTGGTCCATGCTGACGACATTCCAGCTAATCACGCTCGACTACTGGGAGAATGTATATAACATGGTG
TyrThrAsnPheAspAsnPheMetTrpSerMetLeuThrThrPheGlnLeuIleThrLeuAspTyrTrpGluAsnValTyrAsnMetVal
                                                                                        360
```

© IRL Press Limited, Oxford, England.

```
                    1110                1140                1170
CTGGCAACATGTGGTCCTATGAGCGTCTCCTTCTTTACGGTGGTGGTTTTCTTTGGATCGTTTTATCTAATCAATCTGATGTTGGCTGTA
LeuAlaThrCysGlyProMetSerValSerPhePheThrValValValPhePheGlySerPheTyrLeuIleAsnLeuMetLeuAlaVal
                                S6a                                                         390
                    1200                1230
GTGGCGTTAAGCTACGAGGAGGAAGCGGAGATAACGAACGAGGTGAGTCCGTCAAATCTTTAT---------------------------
ValAlaLeuSerTyrGluGluGluAlaGluIleThrAsnGluValSerProSerAsnLeuTyr

                                                    1680                1710
--------------------------------CTGGCTACAGTTTCACAACTGTCTATACAAGTAGTGAGGGATCCCCTTTTCGAGCTG
                                LeuAlaThrValSerGlnLeuSerIleGlnValValArgAspProLeuPheGluLeu
                                                                Slb                         570
                    1740                1770                1800
GCCATCACACTTTGCATCGTGTTAAACACAGCCTTTTTGGCCATGGAACACCACGGGATGAGTGAAAGCTTCCGCAACGCCCTGGATGTG
AlaIleThrLeuCysIleValLeuAsnThrAlaPheLeuAlaMetGluHisHisGlyMetSerGluSerPheArgAsnAlaLeuAspVal
55                                                                                          600
                    1830                1860                1890
GGAAACAAAGTCTTTACGTCAATTTTCACGTTTGAGTGCATTGTCAAGTTAATGGCTCTTTCTAAAGATTTTTTTCTTTGCGGATGGAAC
GlyAsnLysValPheThrSerIlePheThrPheGluCysIleValLysLeuMetAlaLeuSerLysAspPhePheLeuCysGlyTrpAsn
S2b                                                                                         630
                    1920                1950                                                50
ATTTTCGACCTTCTTATCGTTACTGCGAGCTTGCTGGATATAATATTTGAGTTGGTTGATGGCTTGAGTGTCTTGCGTGGCTTGCGATTG
IlePheAspLeuLeuIleValThrAlaSerLeuLeuAspIleIlePheGluLeuValAspGlyLeuSerValLeuArgGlyLeuArgLeu
S3b                                                                                         660
                    2010                2040                2070
CTAAGGGCATTAAAATTGGCTAAATCGTGGACTACAATGAAAGTGCTGCTTAGCATTATAATCTCGACGATCGGCGCTCTTGGCAACCTC
LeuArgAlaLeuLysLeuAlaLysSerTrpThrThrMetLysValLeuLeuSerIleIleIleSerThrIleGlyAlaLeuGlyAsnLeu
S4b                                                                                         690
                    2100                2130                                                31
AGGCTCATTTTGGTCATAGTTATCTACATATTCGCTGTCATCGGCATGCAATTGTTCTCCAAAGACTATACGCCCGAGAAATTTGATCCA
ArgLeuIleLeuValIleValIleTyrIlePheAlaValIleGlyMetGlnLeuPheSerLysAspTyrThrProGluLysPheAspPro
S5b                                                                                         720
                    2190                2220                2250
GATCCAAAAGATTTTATATTAATTATATGCATACCTTTGAACAGATGGAATTTCAATGACTTCTTTCACTCATTTATGATGATCTTTCGC
AspProLysAspPheIleLeuIleIleCysIleProLeuAsnArgTrpAsnPheAsnAspPhePheHisSerPheMetMetIlePheArg
                                                                                            750
                                                    60
                    2280                2340
ATTCTGTGCGGTGAATGGATCGAGCCTCTTTGGGATTGCAATGCGAGCCGAAGAAGAGCAAGGAGCCTCACCTGCTTTGCCATATTTCTG
IleLeuCysGlyGluTrpIleGluProLeuTrpAspCysAsnAlaSerArgArgArgAlaArgSerLeuThrCysPheAlaIlePheLeu
                                                                S6b                         780
                    2370                2400                2430
CCGACCCTAGTGATGGGAAACTTTATGGTGCTTAACTTGTTCTTGGCCTTGTTGCTCAACAGTTTTAACTCCGAGGAGCTCAAGTCGAAA
ProThrLeuValMetGlyAsnPheMetValLeuAsnLeuPheLeuAlaLeuLeuLeuAsnSerPheAsnSerGluGluLeuLysSerLys
                                                                                            810
                                                                                            2970
AAAGAGGCAAGTCAA-------------------------------------------------------------GCACTGACTCGCGTG
LysGluAlaSerGln                                                             AlaLeuThrArgVal
                                                                                            990
                    3000                3030                3060
AGGACGGCTGTCTTAACGGTTGTAGATACTCCGGCCTTTGAGTGGTTTGTGTTGGTGCTGATCTTTGCGTCAAGTATCACGCTCTGCTTC
ArgThrAlaValLeuThrValValAspThrProAlaPheGluTrpPheValLeuValLeuIlePheAlaSerSerIleThrLeuCysPhe
                                Slc                                                         1020
                    3090                3120                3150
GAGGACATTAACCTGGACAAGAACAAGACGCTAAAAAGAGTTCTATACTGGATTAACTTCTCCTTCTGTCTGATATTCGTCGTTGAAATG
GluAspIleAsnLeuAspLysAsnLysThrLeuLysArgValLeuTyrTrpIleAsnPheSerPheCysLeuIlePheValValGluMet
                                                                S2c                         1050
                    3180                3210                3240
ATCCTCAAGTGGCTGGCCCTCGGATTTTCCAAGTACTTTACCAGCTTCTGGACCATACTTGATTTCATCATAGTGTTTGTATCAGTATTC
IleLeuLysTrpLeuAlaLeuGlyPheSerLysTyrPheThrSerPheTrpThrIleLeuAspPheIleIleValPheValSerValPhe
                                                                S3c                         1080
                    3270                3300
TCGCTGCTTATTGAAGAGAATGAAAACCTTAAGGTCCTAAGATCGCTGCGCACATTACGAGCTTTGCGCCCGCTGAGAGCCATCTCTCGG
SerLeuLeuIleGluGluAsnGluAsnLeuLysValLeuArgSerLeuArgThrLeuArgAlaLeuArgProLeuArgAlaIleSerArg
                                S4c                                                         1110
                    3360                3390
TGGCAAGGAATGCGGATTGTAGTAAACGCTCTCATGTATGCAATACCATCAATTTTCAATGTTCTTTTGGTTTGTTTGGTTTTTTGGTTG
TrpGlnGlyMetArgIleValValAsnAlaLeuMetTyrAlaIleProSerIlePheAsnValLeuLeuValCysLeuValPheTrpLeu
                                                                S5c                         1140
                                                    78
                    3450                3510
ATCTTCTCAATAATGGGTGTTCAGTTCTTTGGTGGTAAATTTTTAAGTGCGTTAATAAATATTATTAATCATTTCATATTCCAGGAGGTT
IlePheSerIleMetGlyValGlnPhePheGlyGlyLysPheLeuSerAlaLeuIleAsnIleIleAsnHisPheIlePheGlnGluVal
                                                                                            1170
```

```
                    3540                    3570                    3600
AACGACAAGTGGGACTGCATCGAACAGAACTACACGTGGATTAACTCTAAGATCACTTTCGACCACGTGGGATGGGCTATTTGGCCCTT
AsnAspLysTrpAspCysIleGluGlnAsnTyrThrTrpIleAsnSerLysIleThrPheAspHisValGlyMetGlyTyrLeuAlaLeu
                                                                                        1200
                    3630                    3660                    3690
CTGCAAGTGGCGACCTTCGAGGGCTGGAATGAGGTAATGGCCGACGCAGTTGATGCTCGTGGAGTGGACCTGCAACCGCGACGGGAAGCC
LeuGlnValAlaThrPheGluGlyTrpAsnGluValMetAlaAspAlaValAspAlaArgGlyValAspLeuGlnProArgArgGluAla
                                                                                        1230
                    3720                    3750                    3780
AACCTATATGCGTATATTTATTTTGTTATATTTATTGTGTGCGGATCGTTCTTCACACTGAATCTATTCATTGGAGTTATTATTGTATAAC
AsnLeuTyrAlaTyrIleTyrPheValIlePheIleValCysGlySerPhePheThrLeuAsnLeuPheIleGlyValIleIleAspAsn
                                    S6c                                                 1260
                    3810                    3840                    3870
TTTAACATGCTAAGAAGAAGTATAGAAGGAGGAGTGTTGGAAATGTTTCTCACCGAATCTCAAAAACACTATTACACGGCTATGAAAAAA
PheAsnMetLeuArgArgSerIleGluGlyGlyValLeuGluMetPheLeuThrGluSerGlnLysHisTyrTyrThrAlaMetLysLys
                                                                                        1290
                    3900                    3930                    >600
TTGGGACGAAAGAAACCACAGAAAGTTATTAAGCGACCTATAAATCATTTTTTAGCTATGTTTTATGATTTATCCAATTCGAGAAGGTTC
LeuGlyArgLysLysProGlnLysValIleLysArgProIleAsnHisPheLeuAlaMetPheTyrAspLeuSerAsnSerArgArgPhe
                                                                                        1320
                    3990                    4020                    4050
GAGATCGCGATCTTTGTACTGATTTTTCTAAATATGCTTACAATGGGCATCGAGCACTACGATCAGCCTCATGCAGTCTTCTTCATTCTG
GluIleAlaIlePheValLeuIlePheLeuAsnMetLeuThrMetGlyIleGluHisTyrAspGlnProHisAlaValPhePheIleLeu
                S1d                                                                     1350
                    4080                    4110                    4140
GAAGTGAGCAACGCCTTTTTCACCACAGTATTTGGTCTAGAAGCCATTGTGAAAATTGTCGGTCTGCGCTACCACTACTTCACAGTGCCA
GluValSerAsnAlaPhePheThrThrValPheGlyLeuGluAlaIleValLysIleValGlyLeuArgTyrHisTyrPheThrValPro
                                                                                        1380
          S2d
                    4170                    4200                    4230
TGGAACAGTGTCGACTTCCTTCTGGTGCTGGCCTCTATCTTCGGGATTCTGATGGAAGACATCATGATAGACCTGCCCATTAGTCCGACG
TrpAsnSerValAspPheLeuLeuValLeuAlaSerIlePheGlyIleLeuMetGluAspIleMetIleAspLeuProIleSerProThr
                                                                                        1410
          S3d
                    4260                    4290                    4320
TTACTTCGGGTGGTCCGCGTCTTCCGCATTGGGCGTATATTGCGGTTGATCAAGGCCGCCAAGGGGATCAGAAAGTTGCTATTCGCTCTC
LeuLeuArgValValArgValPheArgIleGlyArgIleLeuArgLeuIleLysAlaAlaLysGlyIleArgLysLeuLeuPheAlaLeu
                                                                                        1440
                                    S4d
                    4350                    4380                    4410
GTAGTGTCCCTGCCCGCCCTATTTAACATCGGAGCTCTGCTAGGACTGATCACCTTTATCTACGCAATTCTGGGCATGTCGCTGTTCGGA
ValValSerLeuProAlaLeuPheAsnIleGlyAlaLeuLeuGlyLeuIleThrPheIleTyrAlaIleLeuGlyMetSerLeuPheGly
                                    S5d                                                 1470
                    4440                    4470                    4500
AATGTCAAGCTCCAAGGTGCTCTCGATGACATGGTGAACTTCCAGACCTTCGGGCGCAGCATGCAGTTACTGTTCCAATTGATGACCTCA
AsnValLysLeuGlnGlyAlaLeuAspAspMetValAsnPheGlnThrPheGlyArgSerMetGlnLeuLeuPheGlnLeuMetThrSer
                                                                                        1500
                    4530                    4560                    4590
GCCGGGTGGAATGACGTACTTGAGTCCCTGATGATACAGCCGCCCGACTGCGACCCATTTATCCATGGGCACACGAACGGAAACTCGGGT
AlaGlyTrpAsnAspValLeuGluSerLeuMetIleGlnProProAspCysAspProPheIleHisGlyHisThrAsnGlyAsnSerGly
                                                                                        1530
                    4620                    4650                    4680
CACCCGCTTCTAGCCATTACCTACTTTACGAGCTTTATCATCATCAGCTATATGATTGTGATCAACATGTACATTGCCATTATCCTGGAA
HisProLeuLeuAlaIleThrTyrPheThrSerPheIleIleIleSerTyrMetIleValIleAsnMetTyrIleAlaIleIleLeuGlu
                                    S6d                                                 1560
                    4710                    4740                    4770
AACTTCAATCAGGCGCACCAGGAAGAGGAGATCGGCATCGTCGAGGACGATCTGGAAATGTTTTACATTCGCTGGTCAAAATACGATCCA
AsnPheAsnGlnAlaHisGlnGluGluGluIleGlyIleValGluAspAspLeuGluMetPheTyrIleArgTrpSerLysTyrAspPro
                                                                                        1590
                    4800                    4830                    4860
CATGCCACCCAATTTATACACTTCTCGCAACTGTCCGATTTTATTGCCTCTCTCGATCCACCATTGGGCATCTCGAAGCCCAATAATGTC
HisAlaThrGlnPheIleHisPheSerGlnLeuSerAspPheIleAlaSerLeuAspProProLeuGlyIleSerLysProAsnAsnVal
                                                                                        1620
                    4890                    4920                    4950
GCTTTAGTGTCATTTAATCTACCCATTTCTAAGGGTAATAAAATACACTGTCTCGACATTTTCGACGCGCTCGTTAAGCACGTGCTAGGT
AlaLeuValSerPheAsnLeuProIleSerLysGlyAsnLysIleHisCysLeuAspIlePheAspAlaLeuValLysHisValLeuGly
                                                                                        1650
                    4980                    5010                    5040
CATGTCGAGGAGACCGATAACTTCAAACAGTTGCAGGAACAAATGGATGTCAAGTTCAAGAAACAGTTTCCAACGCGCAAAGAATTGGAA
HisValGluGluThrAspAsnPheLysGlnLeuGlnGluGlnMetAspValLysPheLysLysGlnPheProThrArgLysGluLeuGlu
                                                                                        1680
```

```
INTRONS
a1:     10        20        30        40        50        60        69   a2:   10        20        30        40
GTGAGTTGTG CAAGAAGTCT GACTTTATAA TCAGAAGTGC GATCTAATCG ATGTTTTTCT TCGTTTTAG  GTATATAACA TGGTTGGTTT ATGAGGTTCC TTATGATGCG
                                                   Possible alternative splice site ^
        50        60        69 b1:    10        20        30        40        50   55        b2:   10        20
GGGGTCTAAT TCAGGGACAA TTCTTCAAG  GTGAGTGTAC ATTCCCGACC ACCAGAATAA ACCTAATTTT TTCTTGCCCT CAAAG     GTGAGAGGAC ATTACCTAGT

        30        40        50 b3:    10        20        31          b4:   10        20        30        40
TTTCAATTTC GGATTTCATA TTCATTCCAG GTGCCAAGGC AAGACCAAAA GCAAATCTGA G        GTATGGGCCC TACGGCTAAT AACTAATAAC TGAAATTACT

        50        60 c1: (estimated length)      30        40        50        60        70        80        90
TAAATTGATA TTATTTCCAG GTAAGTTGAC CAATTATTGG TCTCCCGATA GCTGCACCAC AACCTTCTCA GACATTACGT GCATTCCGAT GAAGAAAGCT ATCTGCGTAG

       100       110       120       130       140       150       160       170       180       190       200
GAACTTCCGT CACTACATCT CGTTCATTCT CTATCACTAC TGATCTATTT TATCTTTTTG GGGAGGAGTT GGTCAGACCG AGCAGTAGGT CCAAGCCGTG TGAGCTGTGT

       210       220       230                 250       260       270       280       290       300       310
TTAATACAAT CTGGTCTTCT TTAATGAAAT TTCAAG--   --CTTAAGCA GATAATTCCA CTCCACCATC TATACACCAT TTGACCACGA CCACTTGGCA CTTTCCCGGC

       320       330       340       350       360       370   376   c2:   10        20        30        40
ATATTGAGAC GACATGTGTC CATCACGTTA TTGCTGCTGG CATTTCATCG ATCATTTCTA TTCCAG    GTAGGGCGAG CTACTGCCGA TTACTGTAAG TATTCTGGTG

        50        60        70        78  d2:   10        20        30        40        51
GATGTTCAGA TGCAGCTTCA CATCAATCAA CACCTGAG  GTAAGTCGCT GAATATTGCT GCTTATATGA GTGTTAATGA CCCTTCCCGA G
```

Legend: Nucleotide sequence of the four homologous domains (a to d) of the Drosophila presumptive sodium channel gene. The deduced amino acid sequence is shown below the coding regions of the gene. The sizes of introns are shown at carets which indicate their positions in the coding sequence. Partial sequences of the presumed cytoplasmic regions 5′ to homology domain a, and 3′ to homology unit d, are also included. The sequence of the (presumed) cytoplasmic regions connecting homology domains a and b, and b and c, is not given. The absent regions are indicated by gaps in the nucleotide sequence. The nucleotide sequence of nine of the ten introns present in the homology domains is given following the sequence of the coding regions. The coding regions of homology domains a, b and c are numbered so that the numbers at the beginning of each domain coincide with the numbering at the beginning of each domain of the eel sequence (2).

Acknowledgements

These studies were supported by NIH grant RO1 NS24785-01, NSF grant BNS-8311024, and a grant from the Klingenstein Fund. We thank G. Mandel and R. Goodman for the eel cDNA probe used to isolate the Drosophila gene, and C. Smith, D. Pauron and J. Garcia, for technical aid and helpful comments.

References

1. Salkoff, L., Butler, A., Wei, A., Scavarda, N., Giffen, K., Ifune, C., Goodman, R., and Mandel, G. (1987) Science in press.
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., 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.