bones from Mules Ear Peak Cave, Texas, have been dated at 1,500 to 3,000 B.P. based on nonassociated archeological artifacts (*4*) but a radiocarbon date now indicates a late Pleistocene age (Table 2). Although historic sightings of condors are known from Arizona and Utah in the late 1800s (*5*), it is thought that these represented an early historic range expansion when large herds of cattle, horse, and sheep were introduced by European settlers to the inland West beginning in the 1700s; cattle are the primary food of condors today (*7, 9*). These herds probably allowed the condor to re-expand its range inland by feeding on the carcasses of these animals, just as it had on those of late Pleistocene megafauna. If condors had survived in the inland West prehistorically, their bones should be conspicuous among the many thousands of bird bones now known from archeological sites in Arizona, New Mexico, Colorado, and Utah; none have been reported (*15*). The absence of Holocene, prehistoric records of *G. californianus* in the inland West supports the above hypotheses.

At Grand Canyon, the extinction of condors in the late Pleistocene is relevant to suggestions on releasing condors there as part of the condor recovery program (*5*). Since condors could not survive in the canyon at the close of the Pleistocene, and when modern flora and faunal communities were gradually established by 8,500 B.P. (*16*), it is unlikely that it could survive there today unless supplemental food supplies are provided on a regular basis.

Other questions regarding condor extinctions still remain, such as why it did not survive in coastal regions of Florida and the eastern United States, and why it apparently did not occur in the Great Plains where large herds of bison existed throughout the Holocene. Answers to these questions may relate to habitat, climatic, and temperature limitations for breeding condors that require additional investigation.

REFERENCES AND NOTES

1. *Gymnogyps amplus*, a fossil species described by L. Miller [*Univ. Calif. Publ. Geol.* **6**, 385 (1911)], is slightly larger than but osteologically similar to modern *G. californianus* and many avian paleontologists recognize it as no more than a temporal subspecies of the latter species. I concur with this synonymy.
2. C. König, in *Vulture Biology and Management*, S. R. Wilbur and J. A. Jackson, Eds. (Univ. of California Press, Berkeley, 1983), pp. 153–171.
3. T. R. Van Devender and R. D. Worthington, *Natl. Park Service Trans. Proc. Ser. 3* (1977), pp. 85–106. One whole but unspecified bone of *Gymnogyps californianus* from a pit with no stratigraphic control at Howell's Ridge Cave provided a conventional date. Van Devender (personal communication) also recently obtained a date (AA-1276) on condor bone from a packrat midden from Maravillas Canyon, Texas.
4. L. Miller, *Condor* **33**, 32 (1931); A. Wetmore, *ibid.*, p. 76; A. Wetmore and H. Friedmann, *ibid.* **35**, 37 (1933); R. deSaussure, *Plateau* **29**, 44 (1956); P. W. Parmalee, *J. Ariz. Acad. Sci.* **5**, 204 (1969).
5. A. M. Rea, *Envir. Southwest* **92**, 8 (1981).
6. D. D. Simons, in *Vulture Biology and Management*, S. R. Wilbur and J. A. Jackson, Eds. (Univ. of California Press, Berkeley, 1983), pp. 470–494.
7. S. R. Wilbur, *North Am. Fauna* **72** (1978).
8. N. F. R. Snyder, R. R. Ramey, F. C. Sibley, *Condor* **88**, 228 (1986).
9. C. B. Koford, *The California Condor* (National Audubon Society, Washington, DC, 1953).
10. A. M. Rea and L. L. Hargrave, *Grand Canyon Nat. Hist. Assoc. Mono. 6* (1984), pp. 77–92. The radiocarbon sample was from a right coracoid, AO 922, housed at the Museum of Northern Arizona, Flagstaff.
11. This date was calculated through weighted averaging of the three dates from Sandblast Cave, which include the youngest from the canyon, after the method of A. Long and B. Rippeteau [*Am. Antiq.* **39**, 205 (1974)].
12. D. J. Meltzer and J. I. Mead, in *Environments and Extinctions: Man in Late Glacial North America*, J. I. Mead and D. J. Meltzer, Eds. (Center for the Study of Early Man, University of Maine, Orono, 1985), pp. 145–173.
13. J. I. Mead *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 836 (1986); P. S. Martin, R. S. Thompson, A. Long, in *Environments and Extinctions: Man in Late Glacial North America*, J. I. Mead and D. Meltzer, Eds. (Center for the Study of Early Man, University of Maine, Orono, Maine, 1985), pp. 5–14.
14. D. W. Steadman and P. S. Martin, in *Quaternary Extinctions: A Prehistoric Revolution*, P. S. Martin and R. G. Klein, Eds. (Univ. of Arizona Press, Tucson, 1984), pp. 466–477.
15. C. R. McKusick, in *The Hohokam*, E. W. Haury, Ed. (Univ. of Arizona Press, Tucson, 1976), appendix 6, pp. 374–377; P. W. Parmalee, *Contrib. Sci. Nat. Hist. Mus. Los Angeles County 330* (1980), pp. 237–250; S. D. Emslie, *Hum. Ecol.* **9**, 305 (1981); A. Ferg, *Kiva* **50**, 111 (1985). One exception is a reference to a bone "fragment resembling California condor, *Gymnogyps californicus*" (sic) from Boulder Springs Rock Shelter, Arizona, a Patayan archeological site dated at A.D. 900 to 1150 [R. H. Hevly, M. L. Heuett, S. J. Olsen, *J. Ariz. Nev. Acad. Sci.* **13**, 67 (1978)]. All bones from the provenience specified for this specimen were examined at the Museum of Northern Arizona, Flagstaff, where they are housed. No bones of *G. californianus* were located though one fragment of the base of an artiodactyl skull is very pneumatic and may have been mistaken for bird.
16. K. Cole, *Science* **217**, 1142 (1982); *Am. Nat.* **125**, 289 (1985); J. I. Mead, *Quaternary Res.* **15**, 311 (1981).
17. L. Miller, *Condor* **62**, 70 (1960).
18. J. I. Mead and A. M. Phillips, III, *Southwestern Nat.* **26**, 257 (1981).
19. D. Donahue and B. Gore, personal communication.

20. This research was funded by grants from the National Geographic Society, Grand Canyon Natural History Association, and Cave Research Foundation. Fossils were collected from caves in Grand Canyon under a permit from Grand Canyon National Park. All material is housed at the University of Arizona Laboratory of Paleoenvironmental Studies, Tucson. I thank rock climbers L. Coats, T. Coats, and D. Dawson for their invaluable assistance and skills in Grand Canyon. Field assistance was also provided by J. and E. Mead, L. Stevens, S. Bassett, S. Heavenrich, R. and J. Raymond, L. Rayburn, R. Haratani, and C. Brown. J. Mead and L. Kiff confirmed identifications of extinct mountain goat bones and condor eggshell fragments, respectively, and T. R. Van Devender kindly allowed publication of radiocarbon data from Maravillas Canyon, Texas. D. Weaver, R. C. Euler, S. Olson, and A. Harris generously provided fossil condor bone and tissue from their institutional collections for use in radiocarbon analysis. I especially thank B. Gore and D. Donahue for their expertise with the radiocarbon analysis. P. S. Martin, D. Steadman and S. David Webb read an earlier draft of the paper and provided many useful suggestions.

5 March 1987; accepted 3 June 1987

Sequence of a Probable Potassium Channel Component Encoded at *Shaker* Locus of *Drosophila*

BRUCE L TEMPEL,* DIANE M. PAPAZIAN, THOMAS L. SCHWARZ, YUH NUNG JAN, LILY YEH JAN

Potassium currents are crucial for the repolarization of electrically excitable membranes, a role that makes potassium channels a target for physiological modifications that alter synaptic efficacy. The *Shaker* locus of *Drosophila* is thought to encode a K^+ channel. The sequence of two complementary DNA clones from the *Shaker* locus is reported here. The sequence predicts an integral membrane protein of 70,200 daltons containing seven potential membrane-spanning sequences. In addition, the predicted protein is homologous to the vertebrate sodium channel in a region previously proposed to be involved in the voltage-dependent activation of the Na^+ channel. These results support the hypothesis that *Shaker* encodes a structural component of a voltage-dependent K^+ channel and suggest a conserved mechanism for voltage activation.

PHYSIOLOGICAL STUDIES OF *Shaker* mutants suggest the the *Shaker* locus of *Drosophila melanogaster* encodes a structural component of a voltage-dependent K^+ channel that conducts a transient current, the A current (*1*). We have characterized genomic DNA and several complementary DNA (cDNA) clones from the *Shaker* region (*2*). Two of these cDNA clones, ShA1 and ShA2, differ only in length. They appear to have at least seven exons spread over 65 kilobases (kb) of genomic DNA; molecular rearrangements corresponding to five *Shaker* mutations are found within this same expanse of genomic DNA. These results suggest that the ShA cDNA clones represent a processed transcript that

Howard Hughes Medical Institute and Department of Physiology, University of California, San Francisco, CA 94143-0724.

*To whom correspondence should be addressed.

is likely to be affected by several *Shaker* mutations.

Here we report the sequence of the cDNA clones, ShA1 and ShA2. Because little biochemical information is available for any K^+ channel, proof that these cDNA clones encode a structural component of a K^+ channel must await functional reconstitution of the A current. However, two features of the predicted primary sequence suggest that these cDNA clones may encode a channel component. First, there are multiple hydrophobic sequences characteristic of integral membrane proteins. Second, the predicted *Shaker* product is homologous to the vertebrate Na^+ channel in a region that may be responsible for voltage-dependent activation. These features suggest that the cDNA clones correspond to the *Shaker* gene product and encode a structural component of a K^+ channel.

The strategy used to sequence ShA1 and ShA2 is shown in Fig. 1. Conceptual translation of ShA1, the longer cDNA at 2928 bp, revealed stop codons in all frames at both ends (Fig. 2A). In between, however, is one long open reading frame. This frame is oriented in the genome so that the 5′ end of the message would be at the end of the locus nearest to the centromere. These results were confirmed by sequencing the independent isolate ShA2, which was 2591 bp long and contained the complete open reading frame present in ShA1. Within the open reading frame, the two were identical except for two single base changes resulting in two conservative amino acid changes in the predicted protein sequence (Fig. 3). A deletion of 2 bp was found in the untranslated 3′ end. These nucleotide changes were probably due to natural variation within the population of flies from which the cDNA library was made.

Occasionally, cDNA clones contain the sequences of introns that had not been removed during messenger RNA (mRNA) maturation (*3*). To determine whether the stop codons 5′ to the open reading frame in ShA were found in a properly processed exon or if they might reflect the presence of an unexcised intron, we sequenced portions of the genomic DNA and compared it to the sequence of the cDNA clones. The interval between the putative exons at positions +64 kb and +78 kb [figure 1 in (*2*)] is a legitimately spliced intron, flanked in the genome by appropriate consensus splice sequences (Fig. 2B). Since, in the exons on either side of this splice site, stop codons exist in all three frames preceding the long open reading frame, we conclude that the 5′ end of the coding region is indeed complete. Similarly, a correctly processed splice site has been located in the predicted 3′ untranslated region (Fig. 3), suggesting that these cDNA clones contain the entire protein coding sequence.

The first AUG codon in the open reading frame may be the translational start site. The sequence immediately surrounding it (TGGCAUG–) differs from the consensus sequence for a ribosome binding site compiled from known initiation sequences in *Drosophila* (ĀĀĀĀĈĈAUG–). This consensus sequence, however, is not absolutely required for proper initiation of translation, since many known initiation sites do not match the consensus (*4*). The next methionine codon is no closer to agreement with the initiation site consensus and is a less attractive alternative because it does not occur until halfway through the open reading frame. A definitive assignment of the translation start site must await the determination of the amino-terminal sequence of an isolated *Shaker* protein.

If translation is initiated at the first AUG in the open reading frame, the sequence predicts that the *Shaker* cDNA's encode a protein of 616 amino acids with a molecular mass of 70,200 daltons (Fig. 3). Possible A channel components of similar size (65 to 75 kD) have been identified in vertebrate brain by cross-linking studies with dendrotoxin, an A channel blocker (*5*).

The *Shaker* product predicted by the sequence of ShA1 and ShA2 has the properties of an integral membrane protein. Hydropathy analysis (*6*) indicates that the protein has a central hydrophobic core, with

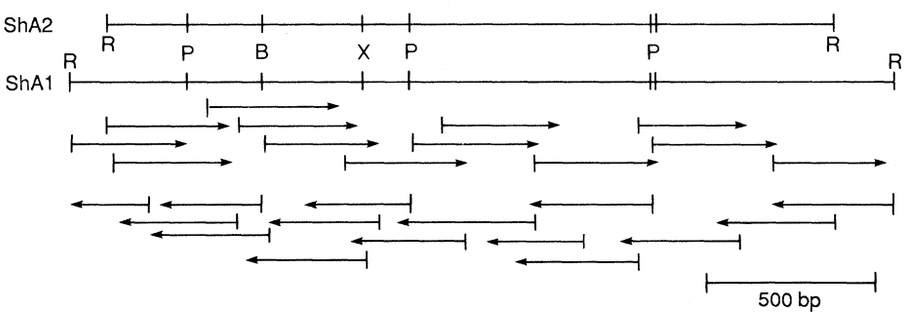

Fig. 1. Restriction map and sequencing strategy of the cDNA clones ShA1 and ShA2. The Eco RI restriction sites at the ends of the cDNA's reflect the cloning insertion sites in λgt11. R, Eco RI; B, Bam HI; X, Xba I; P, Pst I. No sites were found for Hind III, Bgl II, Sac I, or Sma I. DNA was sequenced by the dideoxy termination method (*23*) in M13 vectors (*24*). Initially, several internal restriction fragments were subcloned and sequenced. Oligonucleotide primers were synthesized in order to complete the sequencing of both strands and to allow the determination of the complete sequence in Eco RI subclones where microfragments would not be missed (for example, the adjacent Pst I sites at bp 1705–1716). ShA1 and ShA2 are similarly oriented here and in Fig. 2.

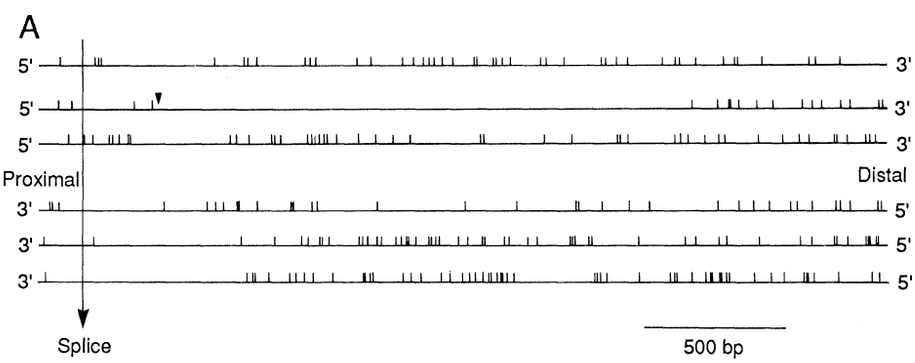

Fig. 2. (**A**) Determination of open reading frame in ShA1. Conceptual translation in all six possible frames revealed only one long open reading frame of 1848 bp. Stop codons are indicated by vertical ticks and the first methionine in the open reading frame is marked by an arrowhead. The splice site indicated by a vertical arrow represents the 12- to 14-kb intron between exons present at genomic walk positions +78 and +65 kb [figure 1 in (*2*)]. Stop codons are present in three frames between the splice site and the predicted translational start methionine. (**B**) Genomic DNA and cDNA sequence in the area surrounding the splice site shown in (A). The eukaryotic splice site consensus sequences were determined by S. Mount (*25*). Sequence from exons is printed in capital letters; from introns in lower case.

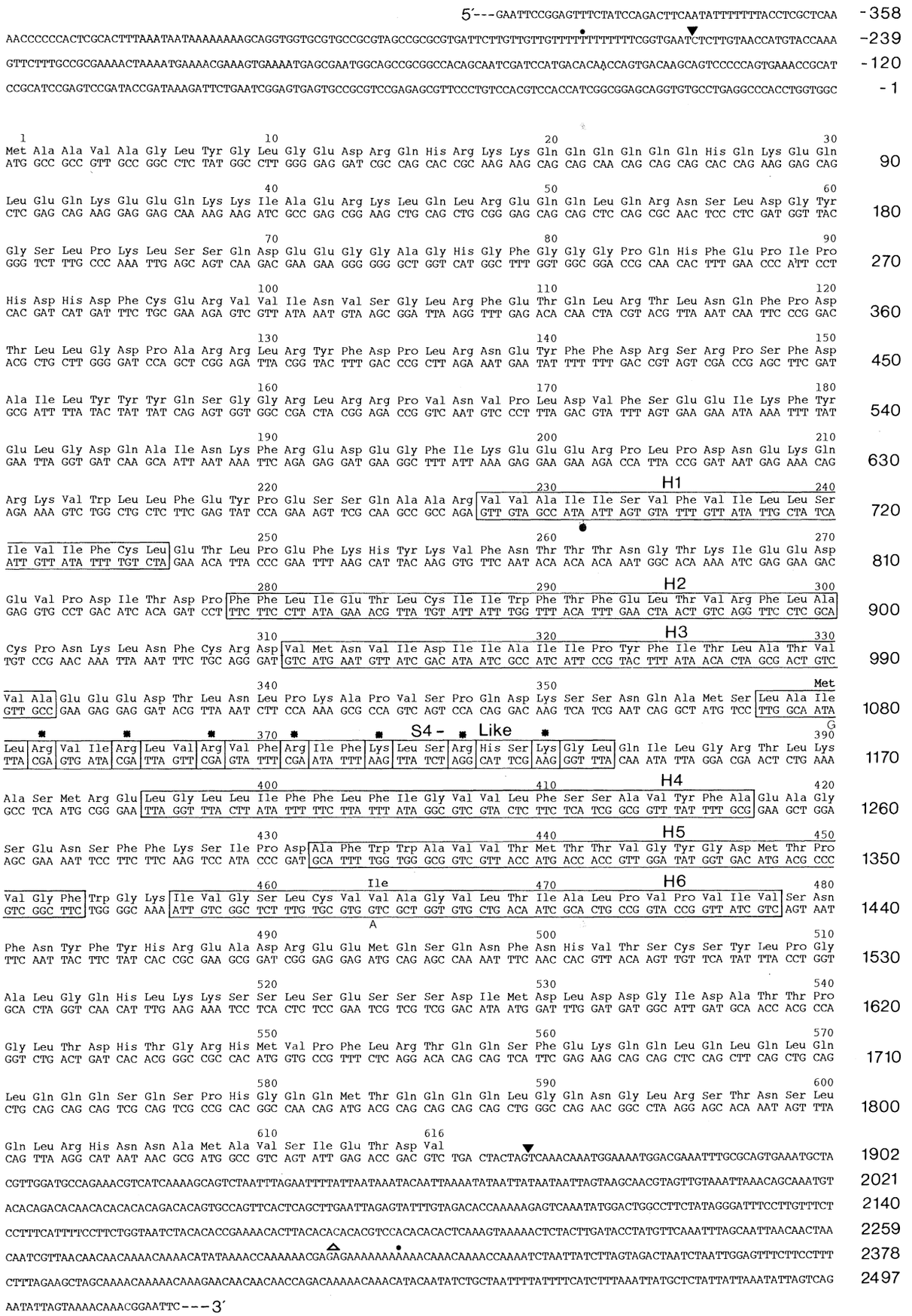
5′---GAATTCCGGAGTTTCTATCCAGACTTCAATATTTTTTTACCTCGCTCAA -358
AACCCCCCACTCGCACTTTAAATAATAAAAAAAAGCAGGTGGTGCGTGCCGCGTAGCCGCGCGTGATTCTTGTTGTTGTTTTTTTTTTTTCGGTGAATCTCTTGTAACCATGTACCAAA -239
GTTCTTTGCCGCGAAAACTAAAATGAAAACGAAAGTGAAAATGAGCGAATGGCAGCCGCGGCCACAGCAATCGATCCATGACACAACCAGTGACAAGCAGTCCCCCAGTGAAACCGCAT -120
CCGCATCCGAGTCCGATACCGATAAAGATTCTGAATCGGAGTGAGTGCCGCGTCCGAGAGCGTTCCCTGTCCACGTCCACCATCGGCGGAGCAGGTGTGCCTGAGGCCCACCTGGTGGC -1
Met Ala Ala Val Ala Gly Leu Tyr Gly Leu Gly Glu Asp Arg Gln His Arg Lys Lys Gln Gln Gln Gln Gln Gln His Gln Lys Glu Gln
ATG GCC GCC GTT GCC GGC CTC TAT GGC CTT GGG GAG GAT CGC CAG CAC CGC AAG AAG CAG CAG CAA CAG CAG CAG CAC CAG AAG GAG CAG 90
Leu Glu Gln Lys Glu Glu Gln Lys Lys Ile Ala Glu Arg Lys Leu Gln Leu Arg Glu Gln Gln Leu Gln Arg Asn Ser Leu Asp Gly Tyr
CTC GAG CAG AAG GAG GAG CAA AAG AAG ATC GCC GAG CGG AAG CTG CAG CTG CGG GAG CAG CAG CTC CAG CGC AAC TCC CTC GAT GGT TAC 180
Gly Ser Leu Pro Lys Leu Ser Ser Gln Asp Glu Glu Gly Gly Ala Gly His Gly Phe Gly Gly Gly Pro Gln His Phe Glu Pro Ile Pro
GGG TCT TTG CCC AAA TTG AGC AGT CAA GAC GAA GAA GGG GGG GCT GGT CAT GGC TTT GGT GGC GGA CCG CAA CAC TTT GAA CCC ATT CCT 270
His Asp His Asp Phe Cys Glu Arg Val Val Ile Asn Val Ser Gly Leu Arg Phe Glu Thr Gln Leu Arg Thr Leu Asn Gln Phe Pro Asp
CAC GAT CAT GAT TTC TGC GAA AGA GTC GTT ATA AAT GTA AGC GGA TTA AGG TTT GAG ACA CAA CTA CGT ACG TTA AAT CAA TTC CCG GAC 360
Thr Leu Leu Gly Asp Pro Ala Arg Arg Leu Arg Tyr Phe Asp Pro Leu Arg Asn Glu Tyr Phe Phe Asp Arg Ser Arg Pro Ser Phe Asp
ACG CTG CTT GGG GAT CCA GCT CGG AGA TTA CGG TAC TTT GAC CCG CTT AGA AAT GAA TAT TTT TTT GAC CGT AGT CGA CCG AGC TTC GAT 450
Ala Ile Leu Tyr Tyr Tyr Gln Ser Gly Gly Arg Leu Arg Arg Pro Val Asn Val Pro Leu Asp Val Phe Ser Glu Glu Ile Lys Phe Tyr
GCG ATT TTA TAC TAT TAT CAG AGT GGT GGC CGA CTA CGG AGA CCG GTC AAT GTC CCT TTA GAC GTA TTT AGT GAA GAA ATA AAA TTT TAT 540
Glu Leu Gly Asp Gln Ala Ile Asn Lys Phe Arg Glu Asp Glu Gly Phe Ile Lys Glu Glu Glu Arg Pro Leu Pro Asp Asn Glu Lys Gln
GAA TTA GGT GAT CAA GCA ATT AAT AAA TTC AGA GAG GAT GAA GGC TTT ATT AAA GAG GAA GAA AGA CCA TTA CCG GAT AAT GAG AAA CAG 630
H1
Arg Lys Val Trp Leu Leu Phe Glu Tyr Pro Glu Ser Ser Gln Ala Ala Arg Val Val Ala Ile Ile Ser Val Phe Val Ile Leu Leu Ser
AGA AAA GTC TGG CTG CTC TTC GAG TAT CCA GAA AGT TCG CAA GCC GCC AGA GTT GTA GCC ATA ATT AGT GTA TTT GTT ATA TTG CTA TCA 720
Ile Val Ile Phe Cys Leu Glu Thr Leu Pro Glu Phe Lys His Tyr Lys Val Phe Asn Thr Thr Thr Asn Gly Thr Lys Ile Glu Glu Asp
ATT GTT ATA TTT TGT CTA GAA ACA TTA CCC GAA TTT AAG CAT TAC AAG GTG TTC AAT ACA ACA ACA AAT GGC ACA AAA ATC GAG GAA GAC 810
H2
Glu Val Pro Asp Ile Thr Asp Pro Phe Phe Leu Ile Glu Thr Leu Cys Ile Ile Trp Phe Thr Phe Glu Leu Thr Val Arg Phe Leu Ala
GAG GTG CCT GAC ATC ACA GAT CCT TTC TTC CTT ATA GAA ACG TTA TGT ATT ATT TGG TTT ACA TTT GAA CTA ACT GTC AGG TTC CTC GCA 900
H3
Cys Pro Asn Lys Leu Asn Phe Cys Arg Asp Val Met Asn Val Ile Asp Ile Ile Ala Ile Ile Pro Tyr Phe Ile Thr Leu Ala Thr Val
TGT CCG AAC AAA TTA AAT TTC TGC AGG GAT GTC ATG AAT GTT ATC GAC ATA ATC GCC ATC ATT CCG TAC TTT ATA ACA CTA GCG ACT GTC 990
Met
Val Ala Glu Glu Glu Asp Thr Leu Asn Leu Pro Lys Ala Pro Val Ser Pro Gln Asp Lys Ser Ser Asn Gln Ala Met Ser Leu Ala Ile
GTT GCC GAA GAG GAG GAT ACG TTA AAT CTT CCA AAA GCG CCA GTC AGT CCA CAG GAC AAG TCA TCG AAT CAG GCT ATG TCC TTG GCA ATA 1080
G
S4 - Like
Leu Arg Val Ile Arg Leu Val Arg Val Phe Arg Ile Phe Lys Leu Ser Arg His Ser Lys Gly Leu Gln Ile Leu Gly Arg Thr Leu Lys
TTA CGA GTG ATA CGA TTA GTT CGA GTA TTT CGA ATA TTT AAG TTA TCT AGG CAT TCG AAG GGT TTA CAA ATA TTA GGA CGA ACT CTG AAA 1170
H4
Ala Ser Met Arg Glu Leu Gly Leu Leu Ile Phe Phe Leu Phe Ile Gly Val Val Leu Phe Ser Ser Ala Val Tyr Phe Ala Glu Ala Gly
GCC TCA ATG CGG GAA TTA GGT TTA CTT ATA TTT TTC TTA TTT ATA GGC GTC GTA CTC TTC TCA TCG GCG GTT TAT TTT GCG GAA GCT GGA 1260
H5
Ser Glu Asn Ser Phe Phe Lys Ser Ile Pro Asp Ala Phe Trp Trp Ala Val Val Thr Met Thr Thr Val Gly Tyr Gly Asp Met Thr Pro
AGC GAA AAT TCC TTC TTC AAG TCC ATA CCC GAT GCA TTT TGG TGG GCG GTC GTT ACC ATG ACC ACC GTT GGA TAT GGT GAC ATG ACG CCC 1350
Ile H6
Val Gly Phe Trp Gly Lys Ile Val Gly Ser Leu Cys Val Val Ala Gly Val Leu Thr Ile Ala Leu Pro Val Pro Val Ile Val Ser Asn
GTC GGC TTC TGG GGC AAA ATT GTC GGC TCT TTG TGC GTG GTC GCT GGT GTG CTG ACA ATC GCA CTG CCG GTA CCG GTT ATC GTC AGT AAT 1440
A
Phe Asn Tyr Phe Tyr His Arg Glu Ala Asp Arg Glu Glu Met Gln Ser Gln Asn Phe Asn His Val Thr Ser Cys Ser Tyr Leu Pro Gly
TTC AAT TAC TTC TAT CAC CGC GAA GCG GAT CGG GAG GAG ATG CAG AGC CAA AAT TTC AAC CAC GTT ACA AGT TGT TCA TAT TTA CCT GGT 1530
Ala Leu Gly Gln His Leu Lys Lys Ser Ser Leu Ser Glu Ser Ser Ser Asp Ile Met Asp Leu Asp Asp Gly Ile Asp Ala Thr Thr Pro
GCA CTA GGT CAA CAT TTG AAG AAA TCC TCA CTC TCC GAA TCG TCG TCG GAC ATA ATG GAT TTG GAT GAT GGC ATT GAT GCA ACC ACG CCA 1620
Gly Leu Thr Asp His Thr Gly Arg His Met Val Pro Phe Leu Arg Thr Gln Gln Ser Phe Glu Lys Gln Gln Leu Gln Leu Gln Leu Gln
GGT CTG ACT GAT CAC ACG GGC CGC CAC ATG GTG CCG TTT CTC AGG ACA CAG CAG TCA TTC GAG AAG CAG CAG CTC CAG CTT CAG CTG CAG 1710
Leu Gln Gln Gln Ser Gln Ser Pro His Gly Gln Gln Met Thr Gln Gln Gln Gln Leu Gly Gln Asn Gly Leu Arg Ser Thr Asn Ser Leu
CTG CAG CAG CAG TCG CAG TCG CCG CAC GGC CAA CAG ATG ACG CAG CAG CAG CAG CTG GGC CAG AAC GGC CTA AGG AGC ACA AAT AGT TTA 1800
Gln Leu Arg His Asn Asn Ala Met Ala Val Ser Ile Glu Thr Asp Val
CAG TTA AGG CAT AAT AAC GCG ATG GCC GTC AGT ATT GAG ACC GAC GTC TGA CTACTAGTCAAACAAATGGAAAATGGACGAAATTTGCGCAGTGAAATGCTA 1902
CGTTGGATGCCAGAAACGTCATCAAAAGCAGTCTAATTTAGAATTTTATTAATAAATACAATTAAAATATAATTATAATAATTAGTAAGCAACGTAGTTGTAAATTAAACAGCAAATGT 2021
ACACAGACACAACACACACAGACACAGTGCCAGTTCACTCAGCTTGAATTAGAGTATTTGTAGACACCAAAAAGAGTCAAATATGGACTGGCCTTCTATAGGGATTTCCTTGTTTCT 2140
CCTTTCATTTTCCTTCTGGTAATCTACACACCGAAAACACTTACACACACACGTCCACACACACTCAAAGTAAAAACTCTACTTGATACCTATGTTCAAATTTAGCAATTAACAACTAA 2259
CAATCGTTAACAACAACAAAACAAAACATATAAAACCAAAAAACGAGAGAAAAAAAAAAACAAACAAAACCAAAATCTAATTATCTTAGTAGACTAATCTAATTGGAGTTTCTTCCTTT 2378
CTTTAGAAGCTAGCAAAACAAAAACAAAGAACAACAACAACCAGACAAAAACAAACATACAATATCTGCTAATTTTATTTTCATCTTTAAATTATGCTCTATTATTAAATATTAGTCAG 2497
AATATTAGTAAAACAAACGGAATTC---3′

Fig. 3 (facing page). Nucleotide and deduced amino acid sequence of the *Shaker* cDNA clone, ShA1. The nucleotide sequence is numbered on the right, with +1 assigned to the first base of the predicted initiation codon. The deduced amino acid sequence is numbered and shown above the nucleotide sequence. A splice site in the 5′ untranslated region (Fig. 2) as well as one found in the 3′ untranslated region are marked with arrowheads. Hydrophobic proposed membrane-spanning sequences are boxed (H1-6). The basic residues of the S4-like region have asterisks above them. Variations in nucleotide sequence and predicted amino acid sequence found in ShA2 are given below and above the sequence determined for ShA1. A deletion of 2 bp in ShA2 is identified by an open triangle. The 5′ and 3′ ends of ShA2 are marked by dots.

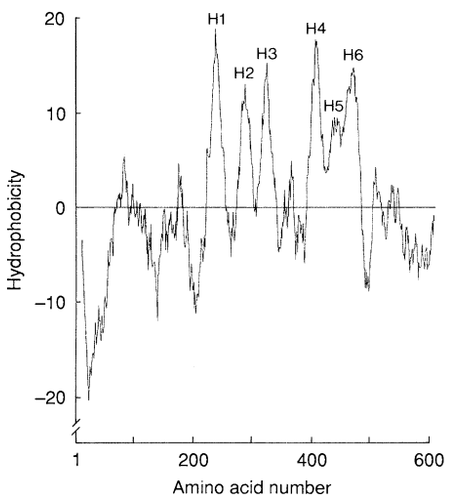

Fig. 4. Hydropathy plot of ShA1. Hydrophobicity values are from Eisenberg's consensus index (*6*). The plot is of running totals of hydrophobicity summed for windows of 19 amino acids and plotted at the position of the central amino acid in the window. The hydrophobic peaks are named as in Fig. 3.

hydrophilic domains at its amino and carboxyl termini (Fig. 4). Although there is little crystallographic information on the structure of intrinsic membrane proteins (*7*), hydrophobic stretches of 19 or more amino acids are generally believed to span the membrane (*6*). Six such stretches can be found in this sequence at residues 228–246, 279–300, 311–332, 396–417, 432–453, and 457–478 (Figs. 3 and 4). A seventh potentially membrane-spanning segment is described below.

A computer search (*8*) of the Dayhoff protein sequence database revealed that the predicted *Shaker* protein shares significant homology with the voltage-dependent Na^+ channel from electric eel (*9*). This channel conducts a rapidly activating and inactivating Na^+ current kinetically similar to the A current. The two proteins are 27 percent identical over a stretch of 121 amino acids. When conservative amino acid substitutions are included, the proteins are 47 percent homologous in the region of overlap (Fig. 5). The center of the homology is an arginine-rich sequence that has been proposed to be involved in voltage-dependent gating of the Na^+ channel (*10*). The region of homology extends into hydrophobic sequences flanking the arginine-rich sequences in both proteins. However, if regions of the predicted *Shaker* protein that do not include the arginine-rich sequence were used in the computer search, no significant homology with any protein in the database was found (*11*).

At the nucleotide sequence level, no significant homology was found between the *Shaker* cDNA's and the Na^+ channel. Since arginine can be encoded by six different codons, the lack of nucleotide homology in the arginine-rich region may easily have arisen during evolution in spite of strong pressure to conserve the protein sequence.

What is the significance of the arginine-rich *Shaker* sequence that is homologous to the Na^+ channel? The Na^+ channel is composed of four domains that are homologous to one another (*9*). A sequence called S4 by Noda *et al.* is present once in each domain. These S4 elements consist of between four and eight arginine residues (or, less frequently, lysine) at every third position; generally, the intervening amino acids are hydrophobic. The predicted *Shaker* protein has a single arginine-rich region with seven basic residues spaced at every third amino acid; the intervening residues are primarily hydrophobic. Thus the arginine-rich region of the *Shaker* gene product has the essential features of the S4 regions in the Na^+ channel (Fig. 5).

Several structural models of the Na^+ channel have proposed that the arginine-rich S4 sequence is the voltage sensor in the Na^+ channel (*10*, *12*). In these models, S4 is predicted to be a transmembrane helix in which the basic residues form a line of positive charges across the membrane. The positive charges would be stabilized within the hydrophobic interior of the protein by forming ion pairs with negative residues supplied by other transmembrane segments. In response to changes in membrane potential, each of the four highly charged S4 helices would move, inducing a conformational change that opens the channel. The fact that the S4 sequence is the most ho-

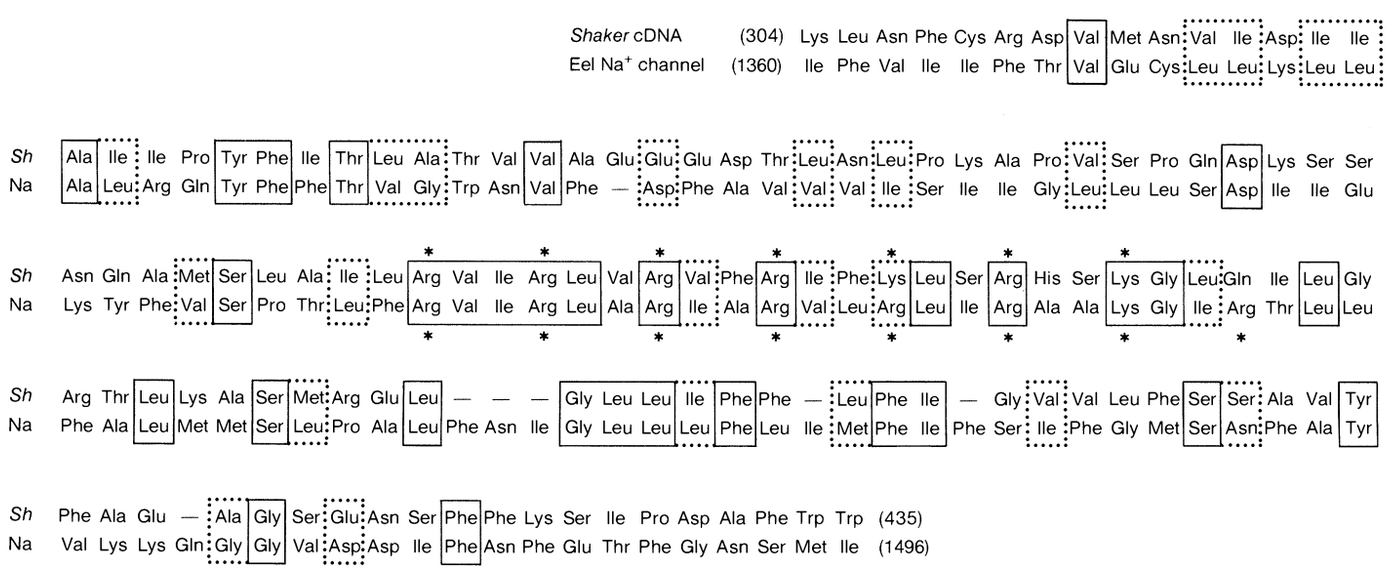

Fig. 5. Amino acid homology between the predicted ShA1 protein and the fourth domain of the electric eel Na^+ channel (*9*). The numbers of the first and last amino acids shown are given in parentheses. Identical amino acids (solid boxes) and conservative changes (dotted boxes) are shown. The amino acid groups considered conserved were [Met, Val, Ile, Leu]; [Ala, Gly]; [Ser, Thr, Gln, Asn]; [Lys, Arg]; [Asp, Glu]. Asterisks mark the basic residues of the S4 region.

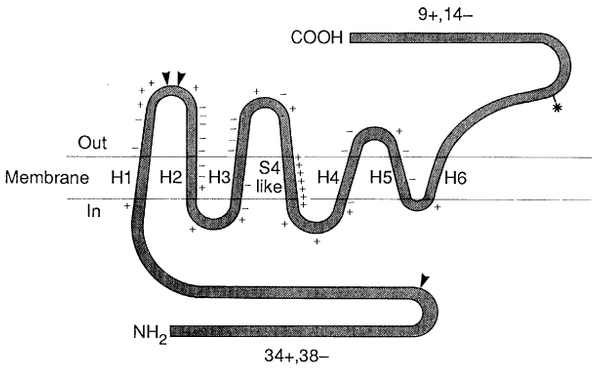

Fig. 6. Proposed model for the transmembrane topography of ShA1 protein. Potential sites for *N*-glycosylation (arrowheads) (*16*) and cyclic AMP–mediated phosphorylation (*) (*15*) are shown. Only those basic (Arg, Lys are +) and acidic (Asp, Glu are −) residues found in or near the hydrophobic regions are shown individually. The total number of charged residues found in the hydrophilic amino and carboxyl domains are indicated.

mologous portion of two voltage-gated ion channels supports the idea that these sequences are involved in the voltage-dependent activation of the channels. At the very least, the region is likely to be functionally important.

The topography of the predicted *Shaker* protein is not easily predicted from its sequence. Our tentative model is based on two features of the primary sequence (Fig. 6). First, the *Shaker* product probably lacks a signal sequence. Although the protein begins with a moderately hydrophobic region, it is only 11 amino acids long and falls just outside the limits of variability of known signal sequences (*13*). The amino-terminal of the protein is therefore likely to be cytoplasmic. Second, seven membrane-spanning regions can be postulated on the basis of hydrophobicity and the need for a transmembrane voltage sensor. With additional structural or biochemical evidence the number or type of membrane crossings may change (*10, 14*). For instance, if the potential site for adenosine 3′,5′-monophosphate (cyclic AMP)–mediated phosphorylation at residue 520 (*15*) is found to be phosphorylated in vivo, the carboxyl end would necessarily be cytoplasmic.

Several features emerge from the working model (Fig. 6). (i) Of the three potential sites for *N*-linked glycosylation [residues 102, 259, and 263 (*16*)], two are extracellular. (ii) Negatively charged residues predominate in the extracellular loops between (and especially near) proposed membrane-spanning regions. In contrast, positively charged residues predominate on the cytoplasmic loops between membrane crossings. A similar charge distribution is seen in several integral membrane proteins (*17*), although the function (if any) of this charge distribution is unknown. (iii) As proposed in models of the Na^+ channel, the S4-like region of *Shaker* is the fourth membrane-spanning region, flanked by hydrophobic regions and oriented so that it crosses the membrane from the extracellular to the cytoplasmic side.

Little is known about the subunit structure of the A channel. Since the predicted *Shaker* protein contains only one S4-like region, a logical question is whether the *Shaker* protein is one subunit of the A channel, analogous to one domain of the Na^+ channel. This question cannot be answered at present but the relatively small size of the *Shaker* protein, compared to the total size of other known channels in nerve and muscle, suggests that one *Shaker* monomer may not be sufficient to form the channel. Previous genetic observations, including allele-specific interactions (*1, 18*) have suggested that the A channel may be a multimeric protein, and that it may contain more than one copy of the *Shaker* gene product.

These considerations lead to two hypotheses for the subunit structure of the A channel. (i) The channel may be a homomultimer composed of an as yet unknown number of copies of the *Shaker* protein. (ii) It may be a heteromultimer in which the *Shaker* protein reported here combines with other subunits. These other subunits may be encoded by other products of the *Shaker* locus, perhaps represented by other cDNA clones (*2*), or by products of one or more separate loci. Known channel structures provide precedent for either a homo- or heteromultimeric structure. Gap junction channels contain six copies of a 27-kD protein (*19*). Muscle acetylcholine receptors are composed of five subunits, only two of which are identical, though all of the subunits are homologous (*20*). Since many channels are built, in one way or another, from repetitive units, it would not be surprising if additional subunits in a heteromeric A channel were homologous to the *Shaker* protein.

Thus, a genetic approach has provided a first step toward studying K^+ channels at the molecular level. It remains to be seen whether the cloning of *Shaker* can facilitate the isolation of other, related channels as might be expected on the basis of Hille's hypothesis that ion channels are evolutionarily related (*21*). The homology observed here between the predicted *Shaker* gene product and Na^+ channels suggests that these proteins may be descended from a single, voltage-sensitive ancestor. As in the case of tyrosine kinase domains or DNA-binding domains (*22*), the arginine-rich, S4-like region may constitute a voltage-gating module shared by several proteins. It will be interesting to see whether voltage-activated Ca^{2+} channels contain an S4-like sequence as well.

REFERENCES AND NOTES

1. L. Salkoff, *Cold Spring Harbor Symp. Quant. Biol.* **48**, 221 (1983).
2. D. M. Papazian, T. L. Schwarz, B. L Tempel, Y. N. Jan, L. Y. Jan, *Science* **237**, 749 (1987).
3. E. D. Gundelfinger *et al.*, in *Nicotinic Acetylcholine Receptor*, A. Maelicke, Ed. (Springer-Verlag, Berlin, 1986), p. 437.
4. D. R. Cavener, *Nucleic Acids Res.* **15**, 1353 (1987).
5. J. V. Halliwell, I. B. Othman, A. Pelchen-Matthews, J. O. Dolly, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 493 (1986); A. R. Black and J. O. Dolly, *Eur. J. Biochem.* **156**, 609 (1986).
6. D. Eisenberg, *Annu. Rev. Biochem.* **53**, 595 (1984); J. Kyte and R. F. Doolittle, *J. Mol. Biol.* **157**, 105 (1982).
7. But see J. Deisenhofer *et al.*, *Nature (London)* **318**, 618 (1985).
8. D. J. Lipman and W. R. Pearson, *Science* **227**, 1435 (1985).
9. M. Noda *et al.*, *Nature (London)* **312**, 121 (1984); M. Noda *et al.*, *ibid.* **320**, 188 (1986).
10. R. E. Greenblatt, Y. Blatt, M. Montal, *FEBS Lett.* **193**, 125 (1985); H. R. Guy and P. Seetharamulu, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 508 (1986); M. Noda *et al.*, *Nature (London)* **320**, 188 (1986); W. A. Catterall, *Annu. Rev. Biochem.* **55**, 953 (1986).
11. The Dayhoff protein sequence database searched (NBRF release 11.0, 4 December 1986) did not contain the rat brain Na^+ channels (*9*) but did contain many membrane proteins including several acetylcholine receptors, many ($Na^+ + K^+$) adenosinetriphosphatases, the anion transporter band III, and so forth.
12. C. M. Armstrong, *Physiol. Rev.* **61**, 644 (1981).
13. G. von Heijne, *J. Mol. Biol.* **184**, 99 (1985).
14. J. Finer-Moore and R. M. Stroud, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 155 (1984); B. Kleffel *et al.*, *EMBO J.* **4**, 1589 (1985).
15. B. E. Kemp *et al.*, *J. Biol. Chem.* **252**, 4888 (1977). The search consensus was (Arg or Lys)-(Arg or Lys)-X-Ser; alternatively, (Arg or Lys)-(Arg or Lys)-X-X-Ser.
16. S. C. Hubbard and R. J. Ivatt, *Annu. Rev. Biochem.* **50**, 555 (1983). The search consensus was Asn-X-(Ser or Thr), where X can be any amino acid except proline.
17. Y. Yarden *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 6795 (1986); E. G. Peralta *et al.*, *Science* **236**, 600 (1987); M. Mueckler *et al.*, *ibid.* **229**, 941 (1985).
18. L. C. Timpe and L. Y. Jan, *J. Neurosci.* **7**, 1307 (1987).
19. P. N. T. Unwin and G. Zampighi, *Nature (London)* **283**, 545 (1980); H. Goodall, *ibid.* **317**, 286 (1985).
20. S. Numa *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* **48**, 57 (1983).
21. B. Hille, *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA, 1984).
22. For instance, T. Hunter and J. A. Cooper, *Annu. Rev. Biochem.* **54**, 897 (1985); J. Berg, *Science* **232**, 485 (1986).
23. F. Sanger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463 (1977). Several regions of ShA1 and ShA2 were GC-rich or contained inverted repeats. These regions were sequenced with 7-deaza-dGTP [S. Mizusawa *et al.*, *Nucleic Acids Res.* **14**, 1319 (1986)]. In addition, deionized formamide was added to 25 percent in buffer gradient gels with 6 percent acrylamide [M. D. Biggin, T. J. Gibson, G. F. Hong, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 3963

(1983)]; to 35 percent in other gels containing 8 percent acrylamide.
24. J. Messing, *Methods Enzymol.* **101**, 20 (1983).
25. S. M. Mount, *Nucleic Acids Res.* **10**, 459 (1982).
26. We thank M. A. Doherty for synthesizing many oligonucleotides, and L. Schulte for preparing the manuscript. Supported by postdoctoral fellowships from ACS (B.L.T.), MDA and NIH (D.M.P.), MDA (T.L.S.) and by NIH grant NS15963; by Klingenstein Fellowship Award (L.Y.J.); and by Howard Hughes Medical Institute.

26 May 1987; accepted 26 June 1987

A Novel Mode of Arbovirus Transmission Involving a Nonviremic Host

LINDA D. JONES, CLIVE R. DAVIES, GORDON M. STEELE, PATRICIA A. NUTTALL

In nature, infected and uninfected arthropod vectors often feed together on an animal. In mimicking this scenario in the laboratory, uninfected vectors were found to acquire virus while cofeeding on the same host as infected vectors. However, the vertebrate host on which they fed did not develop detectable levels of virus in its blood. These observations were made with Thogoto virus, an influenza-like virus of medical and veterinary significance. *Rhipicephalus appendiculatus* ticks were used as the vector and guinea pigs as the vertebrate host. The results demonstrate that a vertebrate that is apparently refractory to infection by an arthropod-borne virus can still play an important role in the epidemiology of the virus, and they suggest a novel mode of arthropod-borne virus transmission.

THE WORLD HEALTH ORGANIZATION defines arboviruses (arthropod-borne viruses) as "viruses which are maintained in nature . . . through biological transmission between susceptible vertebrate hosts by hematophagous arthropods; they multiply and produce viremia in the vertebrates, multiply in the tissues of arthropods, and are passed on to new vertebrates by the bites of arthropods after a period of extrinsic incubation" (*1*). Biological transmission of an arbovirus is confirmed in the laboratory when the arthropod vector becomes infected by feeding on a viremic vertebrate host infected by inoculation or on a virus-containing artificial blood meal. In nature, however, infected and uninfected vectors frequently feed together on the same host. To investigate virus transmission under these conditions, we conducted studies with Thogoto (THO) virus, which is transmitted biologically by the African brown ear tick *Rhipicephalus appendiculatus* (*2*) and is structurally similar to the influenza viruses (family Orthomyxoviridae) (*3*).

Thogoto virus was originally isolated from ticks collected from cattle in Kenya (*4*) and subsequently has been detected throughout central Africa and in parts of the Middle East and southern Europe. The virus is of medical and veterinary significance (*4–6*). A laboratory colony of *R. appendiculatus*, a three-host tick of the family Ixodidae, was established by feeding the ticks on guinea pigs (female Dunkin Hartley strain, average weight 400 g) and maintaining the interfeeding stages at 28°C and a relative humidity of 85%. Donor adult ticks were infected at the preceding nymphal stage with THO virus by feeding on viremic hamsters (*2*) that had been inoculated with either the prototype IIA (*4*) or Sicilian SiAr 126 (*7*) isolates of the virus. Both isolates had been passaged through mouse brain seven times, plaque-picked three times in Vero cell cultures, and then passaged three times in BHK cell cultures. Blood levels of virus were assayed by either plaque titration in Vero cell cultures (*2*) or intracerebral inoculation of 2-day-old mice (Pathology Oxford strain); minimum detectable titers were 20 plaque-forming units (PFU) per milliliter and 1.0 $\log_{10}$ mouse median lethal dose (LD_{50}) per 0.01 ml of blood, respectively.

Infected adult *R. appendiculatus* (donors) were allowed to cofeed on guinea pigs with uninfected ticks (recipients), by placing them either together in the same retaining chamber or in different chambers, on the same animal. In the first experiments, the recipients were uninfected adults (on guinea pigs GP1 to GP4; Table 1). After engorgement, 76% of the recipient adults that were physically separated from the donor ticks had acquired THO virus (chambers 2 of GP1 and GP2). Of the recipient adults sharing the same chamber as donor adults (chambers 1, GP1 to GP4), of either the same or different sex, 64% acquired the virus. A higher proportion of recipient females (81%) compared with males (55%) became infected. There was no evidence of transvenereal infection either from male to female or vice versa. The experiment was repeated with the use of either uninfected nymphs or larvae as recipients (GP5 and GP6; Table 1): 72% of the nymphs (on GP5) and all of three pools of larvae (on GP6), assayed 12 or 10 days after engorgement, respectively, had become infected. During the period of engorgement, not one of the guinea pigs tested developed a viremia detectable by plaque assay of whole blood diluted in medium. However, by day 14, all of them showed serological conversion (50% neutralizing antibody titers of 1/32 to 1/128). The results indicate that THO virus transmission from infected to uninfected ticks can occur while ticks feed on an apparently nonviremic host.

The efficiency of this novel mode of virus transmission on guinea pigs was compared with transmission on hamsters that develop high viremic titers (Table 2). Nymphs cofeeding with infected adult ticks on guinea pigs (GP7 and GP8) had significantly higher infection rates (98%) than nymphs of the same batch fed under similar conditions on two hamsters (68%; $\chi^2 = 10.5$, $P < 0.01$). Virus was not detected in the blood of the guinea pigs (by either plaque assay or mouse inoculation) during the period of nymphal engorgement. Both hamsters, however, developed viremia, H1 reaching a maximum titer of 7.5 $\log_{10}$ mouse LD_{50} (6.0 $\log_{10}$ PFU) per 0.01 ml of blood on day 6 after tick attachment, and H2, 7.9 $\log_{10}$ mouse LD_{50} (5.9 $\log_{10}$ PFU) per 0.01 ml on day 7. Thus vector-borne virus transmission, involving a host seemingly refractory to the virus, was more efficient than transmission via a sensitive host.

To determine if THO virus isolates could replicate in guinea pigs, we inoculated four animals subcutaneously, two with the SiAr and two with the IIA isolates, at doses of 5000 PFU per animal, and then killed them on day 4 or 5 after inoculation [these conditions gave maximum viremic titers in hamsters (*3*)]. Titration by plaque assay of extracts from brain, lung, heart, liver, spleen, kidney, auxillary lymph node, and subcutaneous tissue showed no detectable virus (<10 PFU per whole organ). No virus was demonstrated in the blood by plaque assay, but, when inoculated into 2-day-old mice, blood samples from animals inoculated with the SiAr isolate (but not the IIA isolate) produced clinical signs; virus isolated from the brains of the affected mice was shown to be THO virus by plaque neutralization assay. Blood from the guinea pigs killed 4 and 5 days after inoculation had titers of <1.0

Natural Environment Research Council, Institute of Virology, Mansfield Road, Oxford OX1 3SR, United Kingdom.