Molecular cloning of a member of a third class of Shaker-family K⁺ channel genes in mammals

T. MCCORMACK, E. C. VEGA-SAENZ DE MIERA, AND B. RUDY*

Department of Physiology and Biophysics and Department of Biochemistry, New York University Medical Center, 550 First Avenue, New York, NY 10016

Communicated by R. Llinás, April 5, 1990 (received for review February 5, 1990)

ABSTRACT We report the cloning of RKShIIIA, a cDNA encoding a K⁺ channel sequence expressed in rat brain. This cDNA encodes K⁺ channel subunits that express in *Xenopus* oocytes a slow, 4-aminopyridine- and tetraethylammonium-sensitive, delayed rectifier-type K⁺ channel activated by large membrane depolarizations. This gene belongs to the Shaker (Sh) family of K⁺ channel genes, since the predicted protein has the same overall structure and shows significant homology to other members of this family. However, RKShIIIA cannot be assigned to either of the two known classes of Sh-family genes in mammals based on sequence analysis. Notable features of the RKShIIIA protein product include a probable cytoplasmic loop rich in prolines and a stretch very homologous to the *Drosophila* Shaw protein, both near the amino terminus.

Ion channels are ubiquitous membrane proteins with important and multiple functions in both excitable and nonexcitable cells (1). Potassium (K^+) channels are particularly diverse and are present in all eukaryotic cells (1, 2). In neurons specific combinations of various ion channels underlie the generation of many different signal waveforms and firing patterns and thus contribute to the complexity of neuronal information coding and integration (3).

Two unrelated sequences encoding voltage-dependent K⁺ channels have been characterized, clearly representing two distinct families of genes. The first family, the Shaker (Sh) family, encodes proteins of several hundred amino acids that are similar to those encoded in the Shaker gene in Droso*phila*, a gene that generates several products by alternative splicing (4-6), and includes several members in mammals (7-14) and in Drosophila (15). A gene encoding a voltagedependent K⁺ channel with extremely slow kinetics cloned originally from kidney (16) defines the second K^+ channel gene family. The products of the Sh-family genes consist of a core region with six hydrophobic sequences of which probably five are membrane-spanning domains. Between the third and fourth hydrophobic segments there is a sequence, thought to be membrane-spanning (17, 18), consisting of an arginine (sometimes lysine) at every third position and hydrophobic amino acids in the other positions. This motif, believed to be responsible for voltage sensing, is known as S4 and is also present in voltage-dependent Na⁺ and Ca²⁺ channels (17, 18). In contrast, the gene cloned from kidney encodes a protein with a single membrane-spanning domain and no real S4 sequence.

A comparison of the known sequences in the Sh family suggests the presence of several groups or classes (13). We now report on the cloning of a K^+ channel cDNA from rat brain (RKShIIIA) that clearly belongs to the Sh family.[§] A comparison with other members of this family suggests that this cDNA is a member of a third class of K^+ channel genes in the Sh family in mammals.

MATERIALS AND METHODS

DNA Sequencing. Sequences were obtained by the dideoxynucleotide chain-termination method using Sequenase (United States Biochemical) and plasmid DNA as template.

RNA Expression. The recombinant pBluescript (Stratagene) plasmid containing the RKShIIIA insert was linearized by digestion with *Apa* I or *Xho* I, and full-length capped RNA transcripts were synthesized with T3 polymerase as described (19). Standard methods were used to inject RNA into *Xenopus laevis* oocytes and to record ionic currents (19, 20), except that often the oocytes were not defolliculated until use. All electrophysiological recordings were carried out at 21–22°C in ND96 solution (96 mM NaCl/2 mM KCl/1.8 mM CaCl₂/1 mM MgCl₂/5 mM Hepes, pH 7.5).

Southern Blot Analysis. Digested rat genomic DNA was electrophoresed into 0.8% agarose in $1 \times$ TBE buffer (21). Southern blot transfers into Hybond (Amersham) were prepared essentially as described (21). The blots were hybridized with ³²P-labeled DNA probes synthesized by the random hexamer primer method (22).

RESULTS AND DISCUSSION

Primary Structure of RKShIIIA. The nucleotide and deduced amino acid sequence of clone RKShIIIA is shown in Fig. 1. The predicted product of the cDNA is a protein of 563 amino acids with a calculated relative molecular mass of 61,894. The conceptual translation of RKShIIIA (Fig. 1), a hydropathy analysis (not shown), and a comparison with the amino acid sequence of three classes of Sh-family genes (Fig. 2) indicate that our sequence is similar in overall structure to other Sh-family genes. Like all Sh-family gene products, RKShIIIA contains six hydrophobic regions (H1–H6) and a well conserved S4 domain.

A special feature of RKShIIIA is a 44-amino acid sequence inserted in the amino end region of the protein (enclosed in a box in Fig. 2). The sequence contains several stretches of consecutive prolines, a rare feature found in some other proteins such as the hinge region of some IgGs (25), myelin A1 basic protein (26), and synapsin (27). In addition, two serines in this sequence are surrounded by amino acids similar to those that surround serines or threonines that undergo O-glycosylation (26). According to the proposed topology of this channel (Fig. 1), these putative O-glycosylation sites are on the cytoplasmic side and thus might be involved in the regulation of phosphorylation and protein assembly and targeting (28).

We find that the fifth hydrophobic domain (H5) is consistently one of the most conserved regions among known Sh K^+ channel sequences. This sequence is probably not hydrophobic enough to constitute a membrane-spanning domain (6, 18). Furthermore, if one assumes that H5 is not membrane-spanning, the predicted topology of Sh K⁺ chan-

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.

^{*}To whom reprint requests should be addressed.

[§]The sequence reported in this paper has been deposited in the GenBank data base (accession no. M34052).

-79	ATC CTG AGA ATC TTC AAG CTT ACC CGC CAT TTC GTA GGT CTG AGA GTG CTC GGA 1116
GCCACGGTG AACATCTGGC CCACGAGAGC TTTAACTTGG TGCTGTGTTC GCCTTCCCTA	Ile Leu Arg Ile Phe Lys Leu Thr Arg His Phe Val Giv Leu Arg Val Leu Giv 372
- 20	S4
GTCATGTCTG AGCCACAGAG ATG GGC AAG ATC GAG AAC AAC GAG AGG GTG ATC CTC 36	CAC ACT CIT CCT CCC ACC AAT CAA III TIG IIG CTC ATC ATC III CTC CCT IITO
Met Civilve Tie Civilen Ann Civilero Vel tie Levi 12	His The Ley Are Ale See The Are Clu Die Ley Ley Ley Lie The Ley Are Ale The
	ATS THE LEG ANY ALE SET THE ASH GLO PHE LEG LEG LEG ILE THE PHE LEG ALS SYO
AAT GIL GUA GUC ALL AGG CAL GUA ALL TAL LUL AGU ALI CIC AAG ACC CIT CCT 90	CIG GGA GIT TIG ATA TIC GCT ACG ATG ATC TAC TAC GCT GAG CGA GTA GGG GCT 1224
Ash val bly bly inr arg his blu inr fyr arg Ser inr Leu Lys inr Leu Pro 30	Leu Gly Val Leu Ile Phe Ala Thr Met Ile Tyr Tyr Ala Glu Arg Val Gly Ala 408
	H4
GGA ACT CGE ETG GEE ETT ETE GEE TEE TET GAA EET EAG GGE GAE TGE ETG AET 144	CAA CCT AAT GAT CCC TCA GCG AGT GAG CAC ACA CAG TTC AAA AAC ATC CCC ATT 1278
Gly Thr Arg Leu Ala Leu Leu Ala Ser Ser Glu Pro Gln Gly Asp Cys Leu Thr- 48	Gin Pro Asn Asp Pro Ser Ala Ser Glu His Thr Gin Phe Lys Asn Ile Pro Ile 426
GCT GCG GGT GAC AAG CTG CAG CCG CTG CCC CCT CCG CTG TCT CCA CCG CCG CGA 198	GGT TTC TGG TGG GCT GTG GTG ACC ATG ACT ACC TTA GGC TAT GGG GAT ATG TAC 1332
Ala Ala Giy Asp Lys Leu Gin Pro Leu Pro Pro Pro Leu Ser Pro Pro Pro Arg 66	Gly Phe Trp Trp Ala Val Val Thr Met Thr Thr Leu Gly Tyr Gly Asp Met Tyr 444
_	H5
CCG CCT CCC TTG TCC CCT GTC CCC AGC GGC TGC TTC GAG GGC GGC GCA GGC AAC 252	CCC CAA ACA TGG TCA GGG ATG TTG GTG GGG GCC TTG TGT GCT CTG GCT GGA GTG 1386
Pro Pro Pro Leu Ser Pro Val Pro Ser Gly Cys Phe Glu Gly Gly Ala Gly Asn 84	Pro Gin Thr Trp Ser Gly Met Leu Val Gly Ala Leu Cys Ala Leu Ala Gly Val 462
TGE AGT TEG CAE GGT GGE AAT GGE AGE GAE CAE CET GGG GGA GGE EGE GAA TTE 306	CTG ACC ATA GCT ATG CCT GTG CCC GTC ATT GTC AAC AAT TTT GGG ATG TAC TAC 1440
Cvs Ser Ser His Gly Gly Asn Gly Ser Asn His Pro Gly Gly Gly Arn Glu Phe 102	teu Thr lie Ala Met Pro Val Pro Val lie Val Asn Asn Phe Giv Met Tvr Tvr 480
	H6
THE THE GAT FOR FAR FEA ORA GIE THE OFF TAT GIE FIE AAR TAE TOP APP. 340	THE THE GEA ATE GEE AND FAG ANA CIT FEA AGA ANA AGA AND AND THE FAC ATT CET 1/0/
The fire day doe the con day are the dec fire are the fire for the the day the 120	for los Ala Net Ala los Cia los Los Ana Ana los Ana los Mar Mar Mar His Ila Des (00
Phe Phe Asp Arg His Pro GLY VEL Phe ALE LYT VEL LEU ASH LYT LYT Arg INT 120	Ser Leu ala met ala Lys Gin Lys Leu Pro arg Lys arg Lys Lys his ile Pro 490
GGE AAG ETG EAC TGE EEE GEE GAE GTG TGT GGA EEG ETE TTE GAG GAA GAG ETG 414	CCT GCC CCT CTG GCA AGC TCA CCT ACA TTT TGC AAG ACA GAA TTA AAC ATG GCT 1548
Gly Lys Leu Mis Cys Pro Ala Asp Val Cys Gly Pro Leu Phe Glu Glu Glu Leu 138	Pro Ala Pro Leu Ala Ser Ser Pro Thr Phe Cys Lys Thr Glu Leu Asn Met Ala 516
GCA TTC TGG GGC ATC GAT GAG ACC GAC GTG GAG CCC TGC TGC TGG ATG ACC TAC 468	TGT AAC AGT ACC CAG AGT GAC ACA TGT CTG GGC AAA GAA AAC CGG CTT CTG GAA 1602
Ala Phe Trp Gly Ile Asp Glu Thr Asp Val Glu Pro Cys Cys Trp Met Thr Tyr 156	Cys Asn Ser Thr Gin Ser Asp Thr Cys Leu Giy Lys Giu Asn Arg Leu Leu Giu 534
	•
AGG CAG CAC CGG GAC GCG GAG GAG GCC CTG GAT ATC TTC GAG ACA CCC GAC CTC 522	CAT AAC AGA TCA GTG TTA TCA GGT GAC GAC AGT ACA GGA AGT GAG CCG CCA TTA 1656
Arg Gin His Arg Asp Ala Giu Giu Ala Leu Asp Ile Phe Giu Thr Pro Asp Leu 174	His Asn Arg Ser Val Leu Ser Gly Asp Asp Ser Thr Gly Ser Glu Pro Pro Leu 552
ATE GGA GGE GAE EET GGT GAT GAT GAG GAE ETA GGG GGE AAG AGA ETG GGE ATT 576	TCA CCT TCC GGA AAG GCT CCC CAT CAG ACG CTC TAG TACCAGAG ACAAAAACAG 1710
Ile Gly Gly Asp Pro Gly Asp Asp Glu Asp Leu Gly Gly Lys Arg Leu Gly Ile 192	Ser Pro Ser Gly Lys Ala Pro His Gln Thr Leu * 563
GAG GAT GCT GCG GGG CTG GGA GGA CCC GAT GGC AAG TCT GGC CGC TGG AGG AAG 630	AAGAGGGGAA ACATGTTTCC TGTTGACGAC AGGTGATTAC ACGTGCGCTT CTGATGGAGG 1770
Glu Aso Ala Ala Gly Leu Gly Gly Pro Aso Gly Lys Ser Gly Arg Tro Arg Lys 210	AATCAGGAAA GATAACTGCA AAGATGTTGT CATTACTGGT TACACGCAAG CCGAGGCCAG 1830
TTE FAC FET FEE ATE TEE GET FTE TIT GAE GAE FEE TAT TEA TEE AGA GET GET ANA	ATCICITACT TAATGACTIG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACAGAATT 1890
	ATCTCTTACT TAATGACTTG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACAGAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950
inu Gin Pro Ara Met Tro Ala Leu Phe Giu Ara Dro Tur Car Car Ara Ala Ala 228	ATCHTTACT TRATGACTIG GGAAGGCAC AAAACATGAA AGAAGTGTT GTACGGAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATC AAGGATGTA 1950 CIGGAAACTI CIGCTIGCA ATGTGACCGG AICTGACCGG TITGGCTCG GAAGAGTGC 2010
Leu Gin Pro Arg Met Trp Ala Leu Phe Glu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228	ATCHTCHTACT TAATGACTIG GGAAAGGCAC AAAACATGAA AGAAAGTT GTACAAGAATT 1890 TAATCATGGAT TATTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATGTAT 1990 CTGGAAACTT CTGCTACTGA ATGTGACCGG ATGTGACCGG TTTGTGTTCA GAAGAGTTCC 2010 CCCATTGT CACCCATTA AAGGTGTCC 2010
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228	ATCTETTACT TAATGACTIG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACAGAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CTGGAAACTT CTGCTACTGA ATGTGACCGG ATGTGACCGG TTTGTGTTCA GAAGAGTTC2 2010 CGCATTCTCT GAGGCATTTA AAGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCTT 2070 TGCGAAACCA ATATTATCACTT GTCGTGAGA GCAGATGGA GCTGGAACTG AAATGGGCTT 2070
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738	ATCTCTTACT TAATGACTIG GGAAGGCAC AAAACATGA AGAAGTGTT GTACGAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CTGGAAACTT CTGCTACTGA ATGTGACCGG ATGTGACCGG TITIGTGTTCA GAAGAGTTCC 2010 CGCAATCTT CTGCTACTGA ATGTGACCGG ATGTGACGGG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAATACCA 2130
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TIT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Vai Ser Ile Thr Thr Phe</u> 246	ATCLETTACT TAATGACTIG GGAAGGCAC AAAACATGAA AGAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CTGGAAACTT CTGCTACTGA ATGCACCGG ATGTGACGG TTGTGTTCA GAAGGATTGTA 2010 CGCATTCTCT GAGGCATTTA AAGCTTGTAA AGAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGCT 2190
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TIT GCT TCT CTG TIC TTC ATT TIG GTT TCC ATC ACA ACC TIT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Vai Ser Ile Thr Thr Phe</u> 246 H1	ATCTETTACT TAATGACTTG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATGTA 1950 CTGGAAACTT CTGCTACTGA ATGTGACCGG ATGTGACCGG TTTGTGTTCA GAAGAGTTCC 2010 CGCATTCTCT GAGGCATTA AAGCTTGTA AGAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGCTCTGA ATGGCTCCT TCGGGCCTAT 2250
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TIT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TIT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AMA AAC AAG ACA GAG CCA GTC 792	ATCHTCHTACT TAATGACTIG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CTGGAAACTT CTGCTACTGA ATGTGACCGG ATGTGACCGG TTTGGTGTCA GAAGAGTTGC 2010 CGCATTCTCT GAGGCATTTA AAGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCTT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCCAC AAGGCTCTGA ATGGCTCCTC TCGGGGCTCAT 2250 GGTGGTTTG CTGGAAATTA CAGAGGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTI ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cys Ley</u> Glu Thr Nis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264	ATCLETTACT TAATGACTIG GGAAGGCAC AAAACATGAA AGAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGGG ACATGACGGG GCTGGAACTG GAAGAGTTGC 2010 CGCATTCTCT GAGGCATTTA AAGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATAACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAG AATGCCCCAC AAGGCTCTGA ATGGCTCCT CCGGCCTCAT 2250 GGTGGTTTG CTGCAAATTA CGAGGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTCA 2310 TITCTTTTAT TTGTACATTA CGTGTTGTAT CTACAAACC CGATTGCCTC ATTTCTTGC 2370
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cys Ley</u> Glu Thr Wis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264	ATCLETTACT TAATGACTIG GGAAGGCAC AAAACATGAA AGAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGGG ATGTGACGGG TTTGGTGTCA GAAGGATTGTA 2010 CGCATTCTCT GAGGCATTTA AAGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2010 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCCACC TTCTCACAGT CTATAGCAC AGGGAGCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAG AAATGCCCAC AAGGCTCTGA ATGGCTCCTC TCGGGCCCAT 2250 GGTGGTTTTG CTGCAAATA CAGAGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTTCC 2310 TTTCTTTTA TTGTACATTA CAGGGGCAG TGAACCACT GTAGTGCCT ATTTCTTCC 2370 AACCACTTA TATACCGCT TGTGAACTAT GTTGTAAACT TTTAGTGTCT ACATAGAAAT 2430
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Vai Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cys Ley</u> Giu Thr Wis Giu Ala Phe Asn Ile Vai Lys Asn Lys Thr Giu Pro Val 266 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846	ATCTETTACT TAATGACTTG GGAAGGCAC AAAACATGAA AGAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATGGAAGT TTAGCCATTC AAGGATTGTA 1950 CTGGAAACTT CTGCTGCTGA ATGTGACCGG ATGTGACGCGG TTTGGTGTTCA GAAGGATTGTA 2010 CGCATTCTCT GAGGCATTTA AAGCTTGTA AGAAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAG AAATGCCCAC AAGGCTCTGA ATGGCTCCTC TCGGGCTCTZ 2500 GGTGGTTTG CTGCAAACTA CAGAGGCCAG TGATCCAAT GTAGTATAGG TCTGTTTTCA 2310 TITCTTTTAT TTGTACATTA CAGAGGCCAG TGATCCAAT GTAGTATAGG TCTGTTTTCA 2310 TITCTTTTAT TTGTACATTA CGTGTTGTAT CTACAAACAC CGATGCCTC ATTTCTTGC 2570 AAACCACTTA TATATCCGCT GTGGAACTAT GTTGTAAACT TTTAGTGTCT ACATAGAAT 2430 CTAACCATTT CTAGTGCAG TATTCCATTA GTTCTGGAAC GCTGTTACTA TCTAGAAGAC 2490
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TIT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TIT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 <u>H1</u> TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cvs Leu</u> Glu Thr Wis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala Leu 282	ATCTETTACT TAATGACTTG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATGTA 1950 CIGGAAACTT CTGCTACTGA ATGTGACCGG ATGTGACGG TTTGGTGTCA GAAGAGTTGC 2010 CGCATTCTCT GAGGCATTA AAGCTTGTA AGAAATGGTG GCTGGAACTG AAATGGGCTT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGGCCAG AGGGCCAC ATGCATCCT TCGGGGCTCAT 2250 GGTGGTTTG CTGCAAATTA CAGAGGGCAG TGAATCCAAT GTAGTAAGG TCTGTTTTCA 2310 TITCCTTTAT TIGTACATTA CGTGTGGTAT CTACAAACCAC CGATGCCTC ATTTTCTTGC 2370 AACCACTTA TATACCGCT GTGGAACTAT GTTGTAAACT TTTAGTGTCT ACATAGAAAT 2430 CTAACCACTTT CTAGTCATG TATTCATTA GTTCTGGAAC GCTGTTACTA TCAGAGAC 2490 AGCACTACCAC AGTATCATC GTATAGGTCT TITGTCATG TTAGCAGTAT TATAAATATT 2550
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTI ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AMA AAC AMG ACA GAG CCA GTC 792 <u>Cys Leu</u> Glu Thr Nis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GMA ATC GMA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala Lys 282	ATCLETTACT TAATGACTIG GGAAGGGAC AAAACATGAA AGAAGTGTT GTACGAAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATTGAGAT TTAGCGATTC AAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGG ACATTGAGAT TTAGCGATTGTA A950 CIGGAACTT CIGCTACTGA ATGTGACGGG ACATGACGG GCTGGAACTG GAAGAGTTCC 2010 CGCATTCTCT GAGGCATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACGAG CCAGAGTCCA CAAAACAACT GTAATAACCA 2130 AGCATCGTG AGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGCTCTGA ATGGCTCCTC TCGGGCTCAT 2250 GGTGGTTTTG CTGCAAATA CGAGGGCGA TGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310 THTCTTTTAT TTGTACATTA CGTGTTGTAT CTACAAACC CGATTGCCTC ATTCTTTGC 2370 AAACCACTTA TATATCCGCT TGTGAACTAT GTTGTAAACT TTTAGTGTCT ACATAGAAAT 2430 CTAACCACTTA CTAGTCATG TATTCATTTA GTTCTGGACC GCTGTTACCA 2450 AGCACTCACC AGTATCACTC TGTGAACTAT TTTGTGCATG TTGCTGATATAGC 2450 AGTCATCACC AGTATCACTC TGGAGATATT TTTGTCATTG TTAGCTCAAT TATAATATT 2550 AMAATATAA ATTCTGGCAA TGAGAATATT TTTTATTAA ATGATCAGG AMAATGCCA 2610
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Jie Ala Phe Ala Ser Leu Phe Phe Jie Leu Val Ser Jie Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cys Leu</u> Giu Thr Wis Giu Ala Phe Asn Ile Val Lys Asn Lys Thr Giu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala <u>Leu</u> 282 ACA TAT GTG GAA GGA GTG TGT GTG GTG GTG TGT ACT TTT GAA TTT TTA GTC GCT 900	ATCLETTALE TAATGACTIG GGAAGGCAC AAAACATGAA AGAAGTGTT GTACGAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATGAGAAT TTAGCCATTC AAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGG ACATGAGCG GCTGGAACTG AAGGATTGTA 2010 CGCATTCTCT GAGGCATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2010 CGCATGTGTGCAT GGGTCCACC TTCTCACAGT CTATAGCAC AGGGAGCCAC AATCATTGTC 2190 AGCATCTGTG AAGCTTCTAA AAATGCCCAC AAGGGCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAAG AAATGCCCAC AAGGGCCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAAG AAATGCCCAC AAGGGCCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAAG AAATGCCCAC AGGGCCCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAAG AAATGCCCAC AGGGCCCCCC ATTCATTGC 2370 MACCACTTA TATACCGCT TGTGAACTAT GTTGTAAACC CGATTGCCTC ACAAGAAAT 2430 CTAACCACTTA TATACCGCT TGTGAACTAT GTTGTGAACAC GCTGTTACTA TATAAATAT 2550 AGCTACTCACC AGTATCACC GTATAAGTCT TTTGTCATG TACCTCAAT TATAAATAT 2550 AAAATAGCAC AGTATTCACT GAGAATATT TITTATTAA ATGCCCAAT AAAGAGGATAT 2670
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cys Ley</u> Glu Thr Wis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 266 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala <u>Ley</u> 282 ACA TAT GTG GAA GAG GTG TGT GTG GTG TGG TTT ACT TTT GAA TTT TTA GTC CGT 900 Thr Tyr Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg 300	ATCLETTACT TAATGACTIG GGAAGGCAC AAAACATGAA AGAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATGAGAT TTAGCATTC AAGGATTGTA 1950 CIGGAAATCT CIGCITGGA ATGAGACGGA ATGTGACGG TTAGCGATTC AAGGATTGTA 1950 CIGGAAATCT CIGCITGA ATGTGACCGG ATGTGACGCG TTAGTGTTCA GAAGAGTTCC 2010 CGCATTCTCT GAGGCATTTA AAGCTTGTAA AGAATGGTG GCTGGAACTG AAATGGGCTT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCACCC TTCTCACAGT CTATAGCAC AGGGAGCCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGCACCAC AGGGACCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGCACCAC AGGGACCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AGGGCCGCAC AGGCATCCTC TCCGGCTCAT 2250 GTGGTTTTG CTGCAAATTA CAGAGGCCAG TGAACCAAT GTAGTATAGG TCTGTTTTCA 2310 TITCTTTTAT TTGTACATTA CAGGAGGCAG TGAACCAAT GTAGTATAGG TCTGTTTTCC 2370 AAACCACTTA TATACCGCT TGTGAACTAT GTIGTGAACT TTTAGTGTCT ACATAGAAAT 2430 CTAACCACTTA TATATCCGCT TGTGAACTAT GTIGTGAACG CCTGTTACTA TCTAGAAGAC 2490 AGTCATCACC AGTATTCATC GTAGAGTATT TTTTGTCATTG TTACCTCAAT TATAAATTAT 2550 AMAATATAA ATTCTGGAAT GAGAATATT TTTTATTAA ATGATCAAGG AMAATGTCA AAGGGATAT CAAGGAGTAT CGTGTAGCT TTGGGCTA TGAAGGGGTG GTTTGTCTC 2730
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TIT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Ara Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>GVE Ley</u> Glu Thr Wis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 266 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala <u>Leu</u> 282 ACA TAT GTG GAA GGA GTG GTG GTG GTG GTG GT	ATCTETTACT TAATGACTTG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACGAAATT 1890 TATCATGGAT TITTGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CTGGAAACTT CTGCTACTGA ATGTGACCGG ATGTGACGCGG TTTGTGTTCA GAAGGATTGTA 1950 CGCATTCTCT GAGGCATTTA AAGCTTGTA AGAAATGGTG GCTGGAACTG AAATGGCCT 2010 CGCATTCTCT GAGGCATTTA AAGCTTGTA AGAAATGGTG GCTGGAACTG AAATGGCCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGCTCTGA TGGGACCTC TCCGGCTCT 2250 GGTGGTTTG CTGCAAATTA CAGAGGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTCA 2310 TITCTTTTAT TIGTACATTA CGTGTGTACT ATGCACCC CATTGCCTC ACTGGAACT AACCACCTTA TATATCCGCT GTGGAACTAT GTGTAAACCA CGATTGCCTC ACTAGAAAT 2630 CTAACCATTT CTTAGTCATG TATTCATTTA GTTCTGGAAC GCTGTTACTA TCTAGAGAAC 2490 AGCCATCACC AGTATTCAC GTATAGTCT TTTGTCATTG TTACCTCAAT TATAAATATT 2550 MAAATATAA ATTCTGGCAA TGAGAATATT TTTTTATTAA ATCATCAAGGA AAAAGGATAT 2670 TCTTCAATGG ATATCTTAA AATTATCC TAAAAGTCT ATTTGCGTAA AAAGGGATAT 2670 TCTTCAATGG ATATCTTTAA TGGTGAACAT TTGGGCTAA TGAAGGGTGT CTTTGCCCC 7270 TCTTCAATCG TTGCGAACT TTGGTGAACG TTTGTGGCTAA TGAAGGGTGT 2730
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC ATT ATT GTT AMA AAC AAG ACA GAG CCA GTC 792 <u>Cys Ley</u> Glu Thr Nis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala Ley 282 ACA TAT GTG GAA GGA GTG TGT GTG GTG TGG TTT ACT TTT GAA TTT TTA GTC CGT 900 <u>Thr Tyr Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg</u> 300 H2 ATT GTT TCT CGC CCC AAT AMA CTT GAG TTT GAC ATC ATT T56 AC ATC ATT 954	ATCLETTACT TAATGACTIG GGAAGGGAC AAAACATGAA AGAAGTGTT GTACGAAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATGAGAAT TTAGCGATT CAAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGG ACATGAGAGG GCTGGAACTG AAAGGATGTC 2010 CGCATCCTT GAGGATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATAACCA 2130 AGCATCGTG AAGCTTCAAG AAATGCCCAC TACAGAGCCAC ATTCATTGC 2190 AGCATCGTG AAGCTTCAAG AAATGCCCAC AAGGGAGCCAC ATTCATTGC 2370 AACACCATTA TIGTACATA CGTGTTGAT CTACAAACC CGATGCCCC ATTCATTGC 2370 AAACCACTTA TATACCCGT TGTGAACTAT CTGCAACTA GTGGTATAGG TCTGTTTTCA 2310 TITCTTTAT TIGTACATTA CGTGTTGAT CTACAAACC CGATGCCCC ATTCATTGC 2370 AAACCACTTA TATACCCGT TGTGAACTAT GTTGTAACAC CGGTTACCA ACTGAGAAGAZ 2430 CTAACCACTTA TATACCGCT TGTGAACTAT GTTGTAACTAC TTTAGTGCCTA TCTAGAAGAA 2430 CTAACCACTTA TATACCACT GGAACTAT TTTGTGCATG TTGCTGATT ACATGAGAGA 2450 AGCCATCACC AGTATCACTC TGTGAACTAT TTTGTGCATG ATGCCCAAT AAAAGGAATA 2550 AMAATATAA ATTCTGGCAA TGAGAATATT TTTTTATTAA ATGATCAAGG AAAATGCCA 2640 TATATAGTAG AATTATCAA AATTATCC TAAAATGTCT ATTTCCATA AAAGGAATAT 2570 TCTCTCAATGG AATTCTTTTT TGGTGACAA TGGGTGAATGC ATTTTTCCATGA AAAGGGATAT 2670 TCTCTCAATGG ATTCTTTTT TGGTGAGGAT TTTGTGGCTAA TGAAGGGTGT GTTTGTCTC 2730 ATCACTGCTG TAGTGAAAC TGTGTAAGAG TTTTCATGTA GAAGGGTGT GTTTGTCTC 2730 ATCCCCCTGG TAGTGAAACT TTGTGAACTA TGCTCAAT AAAGGATAT 2670 TCTTCAATGG ATTCTTTTT TGGTGAAGAG TTTTCCATTA GAAGGGTGT GTTTGTCTC 2730 ATCCCCTGG TAGTGAAAC TGTGTAAGAG TTTTCCAATG GCTAATGCA AAATGTCCA 2490
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC ATT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cys Ley</u> Glu Thr Wis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala Lev 282 ACA TAT GTG GAA GGA GTG TGT GTG GTG GTG TGG TTT ACT TTT GAA TTT TTA GTC CGT 900 <u>Thr Tyr Val Glu Giy Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg</u> 300 H2 ATT GTT TTC TCG CCC AAT AAA CTT GAA CTC ATA TTG AAC ATC ATT 954 Ile Val Phe Ser Pro Asn Lys Leu Glu Phe Ile Lys Asn Lys Leu Asn Ile Ile 318	ATCLETTALE TAATGACTIG GGAAGGCAC AAAACATGAA AGAAGTGTT GTACGAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATGAGAAT TTAGCCATTC AAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGG ACATGAGCG GTTGGACTG AAGGATTGTA 2010 CGCATTCTT CAGGCATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2010 CGCATGTGCAT GGGTCCACC TTCTCACAGT CTATAGCAC AGGGAGCCA AATGATGTC 2130 AGCATCTGTG AAGCTTCTAA AAATGCCCAC AAGGGCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAAG AAATGCCCAC AAGGGCCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAAG AAATGCCCAC AAGGGCCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCTAAG AAATGCCCAC AGGGCCCTC AGGGCCCTC TCGGGCTCAT 2250 GGTGGTTTTG CTGGCAATTA CAGAGGCCAG TGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310 TITCTTTTAT TTGTACATTA CGTGTTGTAT CTACAAACAC CGATTGCCTC ACATAGAAAT 2430 CTAACCACTTA TATACCGCT TGTGAACTAT GTTGTGAACG CGTGTTACTA TCTAGAGAAT 2430 CTAACCACTTA TATACCGCAT GAGAATATT TTTTTTTATA ATGCTCCAAT TATAAATATT 2550 AAGCACCCACG AGTATCACA AGGATATT TTTTTTTATA ATGATCAAG AAATGTCG 2610 TATATAGTAG AATGTCT TTGCGGAGC TTGTGAAGGT ATGAAGGGTGT 2670 ICTTCAATCG ATTCTTTT ICGTGAAGC TTTTGTGGCTAA TGAAGGGGTGT CTGTTTGCC2370 AAGCACTGCT AGTTCAAGC TGTGTAAGGC TTGTGGCTAA TGAAGGGTGT 2670 ICTTCAATCG ATTCATTAC AATTATACC TAMAATGTCT ATTTGCATA AAGAGGATAT 2670 ICTTCAATCG ATTCCTTTTT ICGTGAAGG TTTTCTTCT TGCTTAATCA ATATTGCA 2800 AGTCCTTAG TTCCCTTGGG ATTGCTTAT TGTGGAAGG TTTGCCTCAA TGAAGGGTGT GTTGTCTCC 2730 ATCCCTCTAG TTCCCTGGG ATTCGGAATA TATCCTTATG ATAAAAATCC 2790 AGTCCTTAG TTCCCCTGGG ATTCGGAATA TATCTCTATG CCTAATTAA AATCTCCCCTGA AAAGGGATAT 2670 ICTTCAATCG ATGCCTTTTT ICGTGAAGG TTTTCTTCT
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC ATA TTT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cvs Lev</u> Giu Thr Wis Giu Ala Phe Asn Ile Val Lys Asn Lys Thr Giu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala Lev 282 ACA TAT GTG GAA GGA GTG TGT GTG GTG GTG TGG TTT ACT TTT GAA TTT TTA GTC CGT 900 <u>Thr Tyr Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg</u> 300 <u>H2</u> ATT GTT TTC TGG CCC AAT AAA CTT GAG TTC ATC ATT GAA CT ATT GAC ATC ATT 954 <u>Ile Val Phe</u> Ser Pro Asn Lys Leu Glu Phe Ile Lys Asn Leu Lev Asn Ile Ile 318	ATCLETTACT TAATGACTIG GGAAGGCAC AAAACATGA AGAAGTGTT GTACGAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATGAGAT TTAGCATTC AAGGATGTA 1950 CIGGAAATCT CIGCTACTGA ATGTGACGG ACATGAGAGT TTAGCCATTC AAGGATGTA 2010 CGCATTCTCT GAGGCATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2010 CGCATGCTGT GGGCCACCC TTGTGGACGAG CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 AGCGTGTGCAT GGGTCCACCC TTCTCACAGT CTATAGCAC AGGGAGCCAC AATCATTGC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGCTCTGA ATGGCTCCTC TCGGGCTCAT 2350 GTGGTTTTG CTGCAAATA CAGAGGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310 TITICTTTTAT TTGTACATTA CAGGGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310 TITICTTTTAT TTGTACATTA CAGGGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310 CTAACCACTTA TATACCGCT TGTGAACTAT GTGTGAACCAC CGATGCCTC ACATAGAAAT 2430 CTAACCACTTA TATACCGCT TGTGAACTAT GTGTGAACCAC GCTGTTACTA TATAAATAT 2550 MAACCACTTA TATACCGCAT TGAGAATAT TTTTGTCATG TTACCCCAAT TATAAATATT 2550 MAAATATAA ATTCTGGCAA TGAGAATATT TTTTGTGATG TTACCCCAAT TATAAATAT 2570 TCTTCAATGG ATTCCTTTG TGGGAGCAT TGTGGCTAA TGAAGCGGTG GTTTGTCTC 2730 ATCACTGCTG TAGTTGAAAC TGTGTAAGGC TTTGTGGCTAA TGAAGGGTGT GTTTGCTCC 2730 ATCACTGCTG TAGTTGAAAC TGTGTAAGG TTTTGTGATG TGAGCGGTG GTTTGCTC 2730 ATCACTGCTG TAGTTGAAAC TGTGTAAGAG TTTTGTGATG CTAAAGCGGTG GTTTGCTCC 2730 ATCACTGCTG TAGTTGAAAC TGTGTAAGGG TTTGTGGCTAA TGAAGCGGTG GTTTGCTCC 2730 ATCACTGCTG TAGTTGAAAC TGTGTAAGGG TTTTCGCTTAAC ATATTTCC AAATGCCT TGGGCAATAA CTTGGGCAA TGAGGGATGA TGCCTAATG CTAATGCA TATATATACA ATTATTCC AAATGCCT TTGTGGCTAA TGAGGAATAT ACCCCCTGG ATTCGCTGAA TGCCTAATGA CTGTGAACGAGG 2910 CCTGGTTACCA TCCCATGGAG ATCCCAATG AAATGCCAAT TGAGCGAATAGC 2910 CACGTTACTA ACCAAATGAA GCCAATGAAAAAGG CCAAAGGCAATGA CTGGAATAAC CTGGAATAAC
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TIT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Ara Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC AAT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Gvs Ley</u> Glu Thr Wis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 266 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala <u>Ley</u> 282 ACA TAT GTG GAA GGA GTG TGT GTG GTG TGG TTT ACT TTT GAA TTT TTA GTC GGT 900 <u>Thr Tyr Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg</u> 300 H2 ATT GTT TTC TCG CCC AAT AAA CTT GAG TTC ATC ATA AAT CTA TTG AAC ATC ATT 954 <u>Ile Val Phe</u> Ser Pro Asn Lys Leu Glu Phe Ile Lys Asn Leu <u>Leu Asn Ile 11e</u> 318 GAC TH GTG GCC ATC GTC CTC TIC TAG GTG GGG GTG GTG ATC ATT GGG GTG GTG TTT TTA GTC GGT GTG TTT 1008	ATCLETTACT TAATGACTIG GGAAAGGCAC AAAACATGAA AGAAAGTGT GTACGAAATT 1890 TATCATGGAT TITIGCCTGC TGAAAAGGGG ACATGAGAAT TTAGCGATT CAAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGGG ACATGAGCGG GCTGGAACTG GAAGAGTGC 2010 CGCAATCTCT GAGGCATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACATC GTAATAGCAC 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGCT 2370 AGCATCTGTG AAGCTTCAAG AATGCCCAC TAGCAC AGGGAGCCAC ATTCATTGCT 2370 AGCATCTGT GAGCTTCAAG GATGCCCAC AAGGCTCTGA ATGGCTCCTC TCGGGCTCAT 2250 GGTGGTTTTG CTGCAAATTA CAGAGGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310 TITCTTTTAT TIGTACATTA CGTGTTGTAT CTACAAACAC CGATGCCTC ATTTCTTGC 2370 AACCCACTTA TATATCCGCT TGTGAACTAT GTTGTAACAC CGATGCCTC ATTTTCTGC 2370 CTAACCACTTA TATATCGCT TGTGAACTAT GTTGTAACAC CGATGCCTC ATTTTCTGC 2370 CTAACCACTTA TATATCGGCT AGGAACTAT TTTGTGCATG TCACAAGAAAC 2490 AGTCATCACC AGTATTCATC GTATAGTCT TTTGTGCATG TTACCTCAAT TATAATATT 2550 AAAATATAA ATTCTGGGCAA TGAGAATATT TTTTGTGCTT ATTTGCATA AAAGGAATA 2430 CTAACCGATTG TATCCTTTT TGCGGAGCT TTGTGGCTAA TGAAGCGTGT GTTGGCTTA CAATGCAAT 2430 ATCCTCTGG ATTCCTTTTT TCGTGAGGCT TATGGCTAA TGAAGGCTGT GTTGTGCTC 2730 TCTTCCAATCG ATTCCTTTTT TCGTGAGGCT TGTGGCTAA TGAAGGCTGT GTTGTGCTTC 2730 ATCCCTTAG TCCCCGGG ATTCGGAATA TATCTCTATG CCTAATTATA AACCGCTGT 2750 TATCACTGCTG TACTCAAACT GTGTAAGGG CATGCCTAAT CGTTGTGCTA 2750 ATCCCTTAG TCCCCGGG ATTCGGAATA TATCCTAATG CCTAATTATA AACCCCTGTA 2850 TCGTGTACCA TTGCCCGGG ATCCGAATA AGCAATGGC CGTGTGCTAA AGCCAATGGA 2910 CATGGTTACA ACTGAATGA AGAATAAAG TCAAACGATC TGGGGAAAA CTTGAATGCA 2920 ATCCCTTAG ACCGAATGA AGAATAAAG TCAAACGATC TGGGGAAAA CTTGAATGCA 2920
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC ATT ATT GTT AMA AAC AMG ACA GAG CCA GTC 792 <u>Cys Ley</u> Glu Thr Nis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala Ley 282 ACA TAT GTG GAA GGA GTG TGT GTG GTG TGG TTT ACT TTT GAA TTT TTA GTC CGT 900 <u>Thr Tyr Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg</u> 300 H2 ATT GTT TCT CG CCC AAT AMA CTT AGG TTC ATA ATT CTA TTG AAC ATC ATT 954 <u>Ile Val Phe</u> Ser Pro Asn Lys Leu Glu Phe Ile Lys Asn Ley Ley Asn Ile Ile 318 GAC TTT GTG GCC ATC CTC CCC TTC TAC TTA GAG GTG GGA CTC AGC GGG CTG TCT 1008 Arg Dhe Ser Pro Asn Lys Leu Glu Yal Cyl Val Cyl Val GGA CTC AGC GGG CTG TCT 1008	ATCLETTACT TAATGACTIG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACGAAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGGG ACATGGACGG GCTGGAACTG AAAGGGTTGA 2010 CGCATTCTCT GAGGCATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACGAG CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 AGCATCTGTG AAGCTTCAAG AAAGCCCAC TACGAGTCCA CAAAACAACT GTAAATACCA 2130 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGCTCTGA ATGGCTCCTC TCGGGCTCAT 2250 GGTGGTTTTG CTGCAAATA CAGAGGCGAG TGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310 THTCTTTAT TTGTACATTA CGTGTTGTAT CTACAAACAC CGATTGCCTC ATTCTTTGC 2370 AAACCACTTA TATATCCGCT TGTGAACTAT GTTGTAAACT TTTAGTGTCT ACATAGAAAT 2430 CTAACCACTTA TATATCCGCT TGTGAACTAT GTTGTGAACC TTTTGTGTCAAC TTTAGAGACT 2550 AGCCACTGAC AGTTCTCATC GTATATGTCT TTGTGCTTA TCAGAGAGC 26400 AGTCATCACC AGTATTCACTC GTATATGTCT TTGTGCTAA TGATCAGG AAAATGTCAG 2610 TATATAGTAG AATATTATCA AATTATCC TAAATGTCT TTGTGGCTAA TGAGAGATAT 2570 ATCCTGTGTGTG TGTGTAACG TGTGTAAGAG TTTTGTGGCTAA TGAGAGATAT 2670 TCTTCAATGG AATTCTTTTT TTGTGGCAAT ACTCTAATT CAAGAGATAT 2670 TCTTCTAATGG ATTCTTTTT TGGTGAACTAT TCTGTGGCTAA TGAGGATAT 2670 ACTCCTTGG TAGTGAACC TGTGTAAGAG TTTTCTATCA ATATTTCCA AAAGGGATAT 2670 TCCTTGATCG TAGTGAACC TGTGTAAGAG TTTTCATTA ACCCCTGTA 2250 AACCCCTGTA GTGTGAACGA TCTCAAATA ACCCCTGTA CGTTGGTCT 2730 ATCCTCTTAG TTGTGGACCA TCTCAAGGG CATGCCTAAT CGGTTGATCA ACCCCTGTA 2250 ACCCCTGTA ACTCCCTGGG ATTCTCAATA TACTCTAATG ATATTTCCA ATATTTCCA ATCCTGTGT TTTGTGACCA TCTCAAGGG CATGCCTAAT CGTTGGTAT AACCCCTGTG 2850 TCGTGTGACCT TTTGTGACCA TCTCAAGGG CATGCCTAAT CGTTGGATA AGCCCATGGA 2910 CATGTTATAA ACTGAATGA AGAATAAAAG TCAAACGATC TGGGGATAA CTTGAATAA ACGC
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC ATT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cys Ley</u> Glu Thr Wis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala Ley 282 ACA TAT GTG GAA GGA GTG TGT GTG GTG GTG TGT GTG ATT ACT TTT GAA TTT TTA GTC CGT 900 <u>Thr Tyr Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg</u> 300 H2 ATT GTT TTC TCG CCC AAT AAA CTT GAG TTC ATC ATA TGT AGA ATC ATT 954 <u>Ile Val Phe</u> Ser Pro Asn Lys Leu Glu Phe Ile Lys Asn Leu Ley Asn Ile Ile 318 GAC TTT GTG GCC ATC CTC CCC TTC TAC TTA GAG GTG GGA CTC AGC GGG CTG TCT 1008 <u>Asp Phe Val Ala Ile Lew Pro Phe Tyr Ley Glu Val Gly Lew Ser Giy Lew Ser</u> 336	ATCTETTACT TAATGACTIG GGAAAGGCAC AAAACATGAA AGAAAGTGT GTACGAAATT 1890 TATCATGGAT TITIGCCTGC TGAAAAGGG ACATGAGAAT TTAGCATTC AAGGATGTA 1950 CIGGAAACTT CIGCTACTGA ATGTGACGG ACATGAGCGG GTTGGAACTG AAAGGATGTC 2010 CGCATTCTCT GAGGCATTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCAAGGAA ATATTCTGCT TGTCGACAGA CCAGAGTCCA CAAAACAACT GTAAATACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGCCCCC ATTCATGTC 2190 AGCACTCGTG AAGCTTCAAG AAAGCCCCAC AAGGCTCTGA ATGGCTCCTC TCGGGCTCAT 2250 GGTGGTTTTG CIGCAAATA CGAGGGCGA GGAATCCAAT GTAGTATAGG TCTGTTTTCA 2310 TTTCTTTAT TTGTACATTA CGTGTTGTAT CTACAAGAC CGATTGCCTC ACTAGAAAT 2430 CTAACCACTTA TATACCGCT GTGAACTAT GTTGTAAACC CGATTGCCTC ACTAGAAAT 2430 CTAACCACTTA TATATCCGCT TGTGAACTAT GTTGTGAACC CGTGTTACTA TCTAGGAAGAC 2490 AGCCACTCAC AGTATCTCGCAA TGGGATGATT TTTTTTTTTA ATGTGTCT ACATAGAAAT 2550 AAAATGTGGG ATGTCTTTTT TGGGCAATA TATTATTACC AATGTCTGCATG GTTGTGATGA ATGCTCATT GTGGGAGTAT 2670 ICTCLAATGG ATTCCTTTTT TGGGGAACTAT TGTGGGAGG GTTTGTCTCACG AAAATGTCGG AGCATTACC AGTTGTCGGA TGGGAATATT TTTTTTTTATA ATGATCAAGA AAAGGCAAT 2670 ICTCLAATGG ATGCTTTTT TGGTGAACG TTTGTGGCTAA TGATGCAGG GTTGTGCTC Z730 ATCCCTGGG TAGTGGAAC TGGGTAAGGG TTTTCATCT TGCTAATGA ATATTGCCTGAT ATCCCTGGG ATGCGGG ATTGTGAAGA TATCTCTTAG CCTAATTAA AACCCCCTGT 2850 ICGTGTACCT TTGTGCAGG ATTGTGAAGA TATCCTATG CGTGAATGA ACCCCCTGTA 2850 ICGTGTACCT TTGTGCAGG ATTGTGAAGA TATCCTATG CGTGAATGA ACCCCCTGTA 2850 ICGTGTACCT TTGTGCACG ACGAATAAAAG TCAACGGATC TGGGGATAAA CTTGAATGA 2910 CATGTTACA ACTGAAATGA AGAATAAAAG TCAAACGATC TGGGGATAAA CTTGAATGC 2970 ATCCTGATTAA ACTGAAATGA AGAATAAAAG TCAAACGATC TGGGGATAAA CTTGAATGA ATGAT
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TTT ATT GCT TTT GCT TCT CTG TTC TTC ATT TTG GTT TCC ATC ACA ACC TTT 738 <u>Arg Phe lie Ala Phe Ala Ser Leu Phe Phe lie Leu Val Ser lie Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC ATT GTT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>Cys Leu</u> Giu Thr Wis Giu Ala Phe Asn Tie Val Lys Asn Lys Thr Giu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 TI GL CTG GAG ACA CAC GAG GTT TCC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 TAC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 TAC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 TAC TAT GTG GAA GGA GTG TGT GTG GTG GGT TTG ACT TTT GAA TTT TTA GTC CGT 900 Thr Tyr Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg 300 H2 ATT GTT TTC TCG CCC AAT AAA CTT GAG TTC ATC AAA AAT CTA TTG AAC ATC ATT 954 <u>Tie Val Phe</u> Ser Pro Asn Lys Leu Glu Phe IIe Lys Asn Leu Leu Asn IIe IIe 318 GAC TTT GTG GCC ATC CTC CCC TTC TAC TTA GAG GTG GGA CTC AGC GGG CTG TCT 1008 <u>Asp Phe Val Ala IIE Leu Pro Phe Tyr Leu Glu Val Giy Leu Ser Giy Leu</u> Ser 336 H3	ATCTETTACT TAATGACTTG GGAAAGGCAC AAAACATGAA AGAAAGTGTT GTACGAAATT 1890 TATCATGGAT TITIGCCTGC TGAAAATGGG ACATGAGAT TTAGCCATTC AAGGATTGTA 1950 CTGGAAAGTT CTGCTACTGA ATGTGACGG ACATGAGCGG GTTGGAACTG AAGGATTGTA 2010 CGCATTCTCT GAGGCATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2010 CGCATGCTGT GGGCCACCC TTCTCACAGT CTATAGCAC AGGGAGCCA AATCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AGGGAGCCAC AATCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGGCCCAC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AGGGGCCCCC ATTCATTGTC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC GGAACTCAC GTGGTATAGG TCTGTTTTCC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AGGGGCCCCC ATTCATTGTC 2190 AGCATCTGTG AGCTTCAAG AAATGCCCAC GGATGCCTC ACAAGAACT CTGGGGTCTTZ CTGGCAATGAC TATGCGCAAT GTGTATAGG CCGATTGCCT ACATGGAAAT 2430 CTAACCACTTA TATACCGCT TGTGAACTAT GTGTGAACC CGATTGCCT ACATAGAAAT 2430 CTAACCACTTA TATACCGCA TGGAACTAT GTTGTGAACC GCTGTTACTA TCTAGAAGAC 2490 AGTCATCACC AGTATCACTG GTATAGTCT TTTGTCATTG TTACCTCAAT TATAATATT 2550 AMAATATAA ATTCTGGCAA TGAGAATATT TTTTTTTTAA ATGATCAAGG AAATGTCAG 2610 TATATGGTAG ATTCTTGCAAT GTGTAAAGTCT TTGTGGCTAA AAAGGGATAT 2670 TCTTCAATCG ATTCCTTTG TGCGGACGAT TGGGGATAA CTTGTGCTAC AAAGGGATAT 2670 TCTTCAATCG ATTCCTTTTT TCGGGAGCAT TGGGGATAA TGAAGGGTGT GTTTGCTCC 2730 ATCACTCGTG TACTCGGGGA ATTATATCC TAAAATGTCT ATGAAGCGTGT GTTGTCTC 2730 ATCACTGCTG TACTCGGGGA ATTCACAGGG CATGCCTAAT GGAGATAGC 2610 TATATGCTAG ATCCTTTTT TCGGGAGCG TTGTGGCTAA TGAAGCGGTG GTTTGCTCC 2730 ATCACTGATACA ATTCATCC ATTATACC TAAAATGTCT TGGTAACA ATATTCCAG 2790 ATCCCTTAG TTCCCCTGGG ATTCGGAGG CATGCCTAAT CGTTGGGATA AGCCAATGGA 2910 CACGTTACTA ACTGAAATGA AGAATAAAAG TCAACCACTC TGGGGATAA CTTGAATGCT 2770 ATCTGAATGA ATGAT
Leu Gin Pro Arg Met Trp Ala Leu Phe Giu Asp Pro Tyr Ser Ser Arg <u>Ala Ala</u> 228 AGG TIT ATT GCT TIT GCT TCT CTG TTC TTC ATT TIG GTT TCC ATC ACA ACC TTT 738 <u>Ara Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu Val Ser Ile Thr Thr Phe</u> 246 H1 TGC CTG GAG ACA CAC GAA GCT TTC ATT ATT GTT AAA AAC AAG ACA GAG CCA GTC 792 <u>GYB Ley</u> Glu Thr Wis Glu Ala Phe Asn Ile Val Lys Asn Lys Thr Glu Pro Val 264 ATC AAC GGC ACC AGC GCT GTT CTC CAG TAT GAA ATC GAA ACG GAT CCT GCC TTG 846 Ile Asn Gly Thr Ser Ala Val Leu Gin Tyr Glu Ile Glu Thr Asp Pro Ala Ley 282 ACC TAT GTG GAA GAG GTG TGT GTG GTG TGT ACT TIT GAA TTT TTA GTC CGT 900 <u>Thr Tyr Val Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg</u> 300 H2 ATT GTT TTC TCG CCC AAT AAA CTT GAG TTC ATC ATA AAT CTA TTG AAC ATC ATT 954 <u>Ile Val Phe</u> Ser Pro Asn Lys Leu Glu Phe Ile Lys Asn Leu L <u>eu Asn Ile 118</u> 318 GAC TTT GTG GCC ATC CTC CCC TTC TAC TTA GAG GTG GGA CTC AGC GGG CTG TCT 1008 <u>Asp Phe Val Ala Ile Leu Pro Phe Tyr Leu Glu Val Giy Leu Ser Giy Leu</u> Ser 336 H3 TCC AAA GCG GCT AAA GAT GTG CTC GCC TTT CTC AGG GTG AGT AGG TTT GTG AGG 1062	ATCTETTACT TAATGACTTG GGAAAGGCAC AAAACATGAA AGAAAGTGT GTACGAAAT 1890 TATCATGGAT TITGCCTGC TGAAAATGGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 1950 CTGGAAACTT CTGCTACTGA ATGTGACGG ACATTGAGAT TTAGCCATTC AAGGATTGTA 2010 CGCAATCTT CTGCTACTGA ATGTGACGGG ACATTGACGCG TTTGTGTTCA GAAGAGTTCC 2010 CGCAATCTT GAGGCATTTA AGCTTGTAA AGAAATGGTG GCTGGAACTG AAATGGGCT 2070 TGCCCAAGGAA ATATTCTGCT TGTCGCACGA CCGAGGTCCA CAAAACAACT GTAATAACCA 2130 ACGTGTGCAT GGGTCCACCC TTCTCACAGT CTCATAGCAC AGGGAGCCAC ATTCATTGC 2190 AGCATCTGTG AAGCTTCAAG AAATGCCCAC AAGGCTCTGA ATGCGTCCTC TCGGGCTCAT 2250 GGTGGTTTTG CTGCAAATTA CAGAGGGCAG TGAATCCAAT GTAGTATAGG TCTGTTTCA 2230 TTTCTTTAT TTGTACATTA CGTGTTGTAT CTACAAACAC CGATTGCCTC ATTTCTTGC 2370 AAACCACTTA TATATCCGCT TGTGAACTAT GTTGTAAACAC CGATTGCCTC ATTTTTCTGC 2370 CTAACCACTTA TATATCCGCT TGTGAACTAT GTTGTAAACAC CGATTGCCTC ATTTTCTGC 2370 CTAACCACTTA TATATCCGCT TGTGAACTAT GTTGTAACAC CGATTGCCTC ATTTTCTGC 2370 AACCACTTA TATATCCGCT TGTGAACTAT GTTGTAACAC CGATTGCCTC ATTTTCTGC 2370 CTAACCACTTA TATATCCGCT TGTGAACTAT TTTGTGCAATG CCTGTTACCA 2490 CTCAACCATTA CTAGTCATG TGAAAGTCT TTTGGGAAC GCTGTTACCA TATAAATATT 2550 AAAATATAA ATTCTGGGCAA TGAGAATATT TTTTGTGCAT AAGGAATAT 2470 TGTCTCAATCG ATTCCTTTT TTGGTGAACCA TGGGGATAA ATGTCAGG AAAATGTCAG 2400 TTGTTAATGG ATTCCTTTT TTGGTGAACCA TTATGGAAC CCTGTTACTA TATAAATATT 2570 TGTTGCACG ATTCCTTTT TTGGTGAACCA TTATGGAACAC TTTGTGCATA AAGGAATAT 2570 TCTTCAATCG ATTCCTTTT TTCGTGAACCA TGAGAATATT TTTTTATAA ATGATCAAGG ATTTTCCAA ATTATATCC AATATATAA ATTCCGGGA TGAGAATATT TTTTGAGCTA TATTGCAAGA ATGTTCCAGCTG TATTGTCTTAC ATTCTGAACGA TGTGTAACA ATATTTCCAG 2790 ATCCCTTAG TTCCACTGGG ATTCGAATA TACCCAATC GTTGTGATAA AGCCATGCA 2690 CATGTTACA ACTGAATGA AGAATAAAG TCAAACGATC TGGGATAAA CTTGAATGC 2970 ATCTGTTAGA ATGAATGA AGAATAAAG TCAAACGATC TGGGGATAAA CTTGAATGCC 2970 ATCTGATTAA ATGAT

FIG. 1. Nucleotide and deduced amino acid sequence of RKShIIIA. The name given to this clone attempts to indicate its relationship to other K⁺ channel sequences. The R indicates the clone is from rat, K that it encodes a K⁺ channel, Sh that it belongs to the Sh family of genes, III that it is a member of the third known class of genes in this family, and A that it is the first known member of this class in rat. We have not indicated tissue of origin in the name, because other sequences have been found in more than one tissue (e.g., RCK4 and RHK1, found in brain cortex and heart, respectively, are the same sequence). The numbers indicate the nucleotide and amino acid positions with the chosen initiation codon as 1. The ATG triplet chosen as the initiation codon is the second found in the sequence, but it produces the longest possible open reading frame and predicts a protein that is highly homologous to other Sh K⁺ channel sequences. The first stop codon in frame is shown by an asterisk. Hydrophobic domains (H1-H6) and the S4 motif are underlined. Consensus sequences for N-glycosylation (23) are indicated with filled squares. Only two of these (Asn-259 and -266) are extracellular in the expected topology of the protein assuming that the amino terminus is intracellular (because of a lack of signal-peptide sequence) and that H1-H4, S4, and H6 are membrane-spanning domains (6, 18). The clone containing the RKShIIIA insert was obtained by screening (21) a rat brain cDNA library in λ ZAP (Stratagene) with DNA probes derived from the amino-terminal end of a partial clone obtained previously. Hybridization to nitrocellulose replicas of cultures was carried out in 50% (vol/vol) formamide/ $5 \times$ SSPE at 42°C (21). Another isolate was partially sequenced. It is identical to RKShIIIA except that in the 3' end instead of a T as the last nucleotide it contains a stretch of 10 A residues. The initial partial clone was obtained by screening, at low stringency [20%] formamide/ $5 \times$ SSPE at 30°C], an oligo(dT)-primed rat brain λ gt11 cDNA library with an oligonucleotide probe corresponding to the first 60 bases of the (H5) domain of the Drosophila Shaker H4 cDNA (4). The recombinant pBluescript plasmid containing the RKShIIIA insert was obtained from the purified λ ZAP clone by *in vivo* excision using Stratagene protocols.

nel subunits will be similar to that of the homologous repeats in voltage-dependent Na^+ and Ca^{2+} channels (18). We propose that H5 enters the membrane partially and helps to form the K^+ selectivity filter. If a K^+ channel is a tetramer of Sh-family gene products (18), the center of the H5 sequences could form a ring containing 12 hydroxyl groups. These

	Ų	
RKShillA:	MGK IENNERVILNVGGTRHETYRSTLKTLPGTRLALLASSE	41
Shaw:	-NLIN MDSENVKAKI-ASR-TEAL	43
DRK1:	-T-HGSRSTSSLPPEPHEIVRS KACSRRLAVLWRDRRGK-RDCN	60
RCK1:	-TVNSGENADEASAAPGHPQDGSYPRQADHDDH-CCVI-IS-L-FQLKAQF-N-L-GNPKKRM	70
RKShIIIA:	PGGDCLTAAGDKLGPLPPPLSPPPRPPPLSPVPSGCFEGGAGHCSSHGGNGSDHPGGGREFFFDRHPGVF	111
Shaw:	ANY-PIL N-Y	62
DRK1:	THDSL-QVCD-YSLED N-YA-	88
RCK1:	RYF-P-R N-YNRPS-	89
RKShIIIA:	AYVLNYYRTG KLNCPADVCGPLFEELAFWGIDETDVEPCCWNTYRONRDAEEALDIFETPDLIGGDPG	180
Shaw:	-QTQ-T-AVLDRLDTEK-S	131
DRK1:	TSIF RNMEEN-ALS-SQDYIYL-SQAR-H-KKEGNNEELKR-AET-REREGE	157
RCK1:	DAI-YQS-GR-RR-VN-PLDM-SIK-YELG-EAM-KFREDEGFIKEEERPLPE	146
	H1	
RKShIIIA:	DDEDLGGKRLGIEDAAGLGGPDGKSGRURKLOPRIMALFEDPYSSRAARFIAFASLFFILVSITTFCLET	250
Shaw:	EE-LARKFGFEEDYYK-TIS W-GENK1-SDENKT-GVV-VCILSK-	195
DRK1:	EFONTCC AEKRKKL-D-L-K-NVKIL-II-INVL-TIALS-N-	206
RCK1:	KEY-RQV-LY-EGPYIV-VWVIVI	188
	H2	
RKShVIIIA	:NEAFNIVKNKTEPVINGTSAVLQYEIETD PALTYVEGVCVVWFTFEFLVRIVFSPNKLEFIKNL	314
Shaw:	-POMRVPIVRNIT-KTANGSNGWFLDK-QTNAN I-FF-I-CNAFISFIS	263
DRK1:	LPELQSLDEFGQSTD N-Q-AHAIAH-Y-L-FLSK-WK-F-GP	257
RCK1:	LPELKDD-DF-GTINRIDNTTVI-TSNIFTD PEFITL-11SLVFFAC-S-TD-F1	253
	<u> </u>	
RKShIIIA:	LNIIDFVAILPFYLEVGLSGUSSKAAKDVLG FLRVVRFVRILRIFKLTRHFVGLRVLGHTLRAST	379
Shaw:	VYI-T-SIDLV-ORFA-HLENAD I-EFFSIIH-LVSS-KI-IQ-FA	326
DRK1:	ALLY-VTIF-TESHKSVLOFONVRR VVOIF-INLAST-QSFR-Y	324
RCK1:	M-FII-YFITL-TEIAEGEGNGKGEGATSLA	323
	• • H4 H5	
RKShIIIA:	NEFLLLIIFLALGVLIFATNIYYAERVGAQPNOPSASENTQFKNIPIGFMAVVTNTTLGYGDNYPQTMS	449
Shaw:	K+LTVFVIVSLV+-IQPNND-NSL-LLVA-K+YI	.389
DRK1:	LGLM-IMSSLVFFKDEDD -KSASTIVIK-LL	385
RCK1:	R-LGFFIIL-SSAV-F-AEE-E SH-SSDLSV-IG	384
		E 10
KKSHITIA:		219
Snaw:		437
DKK1:	KI-G-G-CIV-C-I-ISEF-KEGKRGE-AIKK-EALERAKRG-IVSHWADAFAKSIE	433
RUNI:	. K 1	-)4
RKShIIIA:	TQSDTCLGKENRLLEHNRSVLSGDDSTGSEPPLSPSGKAPHQTL*	563
Shaw:	PG-GPHS-PMGSGGTGP-RMNNKTKDLV-PKSDMAFSFD*	498
DRK1:	NHD I VVEKNGESI AKKDKVQDNHLSPNKWKWTKRALSETSSSK (+355)	498
RCK1:	MEIEEDMINISIAHYRGANIRTGNCTA-DQNCVNKSKLLTDV*	495

FIG. 2. Comparison of predicted amino acid sequences of RK-ShIIIA and other Sh K⁺ channel genes. The complete sequences of RKShIIIA, Shaw, and RCK1 are shown. DRK1 extends 355 amino acids further at the carboxyl end. Identical amino acids are shown with dashes. The sequences were aligned to maximize long stretches of homology. Gaps in the sequence are shown as blanks. The beginning and the end of the central core region used for calculations of identity are marked by arrows. The H1–H6 and S4 sequences and a 44-amino acid insert peculiar to RKShIIIA are boxed. The leucines in heptad repeat ("leucine zipper" motif) adjacent to the S4 domain are indicated with filled circles (24). In RKShIIIA the fourth leucine is replaced by phenylalanine.

hydroxyls could provide the water-like environment which dehydrates K^+ ions selectively (1, 29).

Functional Properties of RKShIIIA Channels. In vitro transcribed RNA from RKShIIIA expresses K⁺ channels when injected into Xenopus oocytes. Large currents absent in uninjected oocytes are observed in oocytes injected with RKShIIIA transcripts (Fig. 3A-C). The reversal potential of these currents, determined from tail-current analysis (19), depends on external K⁺ concentration as expected for a K^+ -selective channel. A reversal potential of -85 mV is obtained in 2 mM external K^+ and -55 mV in 10 mM (data not shown). These values are close to the expected equilibrium potential for K^+ (-95 mV and -55 mV, respectively), assuming an internal K^+ concentration of 90 mM, and indicate a very high selectivity for K^+ ions (1). The currents expressed by RKShIIIA activate slowly, starting at about -10 mV, and their rise time decreases with increasing depolarization (Fig. 3 A-C). These currents inactivate very slowly (Fig. 3B) and activate with a delay (Fig. 3C). All these are features of delayed rectifier-type K^+ channels (1, 2).

A conductance-voltage curve obtained from five experiments is presented in Fig. 3D. The conductance begins to rise



FIG. 3. Electrophysiological properties of ion currents expressed by RNA transcripts of RKShIIIA. (A-C) Currents recorded under voltage clamp. Depolarizing pulses, from -80 to +40 mV with 10-mV increments, were delivered at a rate of every 4 sec (A and C) or 8 sec (B) from a holding potential of -100 mV. The oocytes were bathed in ND96 solution and the electrodes were filled with 3 M KCl. (D)Plot of normalized conductance (conductance at the indicated potential divided by the maximum conductance) obtained from five oocytes and three different preparations of RNA. In one case (open circles) depolarizing pulses were given in 5-mV increments. The conductance was calculated (1) by dividing the steady-state current at the end of 500-msec depolarizing pulses by the driving force using a reversal potential of -85 mV.

between -15 and -10 mV and increases steeply to a maximum value. Sometimes a decline in conductance is observed with large depolarizations. Half-maximal conductance is obtained at around 13 mV and the limiting slope of the curve is estimated to be 7-7.5 mV for an *e*-fold change in conductance. The reduced steepness exhibited by RKShIIIA, compared to *Drosophila* Shaker channels (19), may be due to the smaller number of basic residues in the S4 domain (six vs. seven). However, amino acids in other regions of the protein may also influence the steepness of conductance-voltage relations (13).

The currents expressed by RKShIIIA are blocked 70– 100% (depending on the membrane potential) by 1 mM tetraethylammonium, 5 mM Ba²⁺, and 1 mM 4-aminopyridine, three well known K⁺ channel blockers. The blockage of RKShIIIA channels by 4-aminopyridine seems to exhibit open-channel block behavior (1), contrary to what has been observed for 4-aminopyridine blockade of K⁺ channels in squid giant axon (30). Specifically, the degree of blockage of RKShIIIA currents by 4-aminopyridine increases with increasing depolarization, the rise time of the currents becomes faster, and at some voltages (e.g., at +20 mV) a timedependent inactivation is induced (Fig. 4). Furthermore, repetitive depolarization produces further blockage (see *Inset* in Fig. 4B), as opposed to the gradual release of block observed for the delayed rectifier of the squid giant axon (30).

 K^+ Channel Classes. Based on sequence homology, and hence evolutionary relatedness, of previously known Shfamily sequences, two classes of genes (13) can be defined (classes I and II in Table 1). Members of a given class, in different species, are far more similar to each other than to members of different classes in the same species (Table 2). This suggests that these two classes originated from some ancestral sequence before the flies and mammals diverged (13–15). RKShIIIA is more similar to the *Drosophila* Shaw gene (15) than to any other known K⁺ channel sequence (Table 2) or any sequence in the GenBank data base (January



FIG. 4. 4-Aminopyridine block of RKShIIIA currents. Records of ion currents obtained as in Fig. 3 for depolarizing pulses to the indicated voltages are shown before (A) and after (B) the addition of 0.8 mM 4-aminopyridine. The *Inset* in B shows the currents during a depolarizing pulse to 20 mV applied at a frequency of 1/sec obtained after the experiment shown in B was finished.

1990), particularly in a long stretch near the amino terminus that is almost identical between these two sequences (Fig. 2). Furthermore, genomic DNA analysis suggests that several genes of the same class, with similar homologies to Shaw, are present in the rat genome (Fig. 5). However, as shown in Table 2, the overall homology between RKShIIIA and Shaw is significantly less than that among class I or class II genes (e.g., between RCK1 in rat and H4 in *Drosophila*, or between DRK1 in rat and Shab in *Drosophila*). Outside the stretch in the amino terminus, RKShIIIA is more similar to Shaw than to other Sh genes in the sixth hydrophobic domain (H6) and the sequence immediately following H6, but it is more similar to RCK1 or DRK1 in the region around the S4 domain (Fig. 2).

RKShIIIA is clearly different from class I or class II genes and thus we consider it a member of a third Sh gene class in mammals. Class I and II genes have vertebrate and invertebrate counterparts. RKShIIIA and Shaw are probably related genes; however, their evolutionary relatedness is different from that between class I and class II genes. Thus, their assignment to the same class of genes is tentative, and definite assignment must await further evolutionary analysis.

The physiological significance of these classes is not known. Inactivation is clearly not a defining parameter because, for example, class I includes channels that inactivate very fast and channels that inactivate very slowly and have delayed rectifier properties (Table 3). Differences in these properties apparently can be achieved without major changes in the core region. From the electrophysiological studies done thus far, we can distinguish two parameters of predicted physiological significance which appear to vary among classes (Table 3): class I members activate quickly and at relatively low (but not subthreshold) voltages, while

Table 1. S	haker fam	ily of K ⁺	channel	genes
------------	-----------	-----------------------	---------	-------

Class	Rat	Human	Drosophila
I	RCK1 = RBK1 (10, 8)	HuKI (12, 13)	Shaker (4-6)
	RCK5 = RBK2 (10, 9)	HuKIV (12, 13)	
	RCK4 = RHK1 (10, 11)	HuKII (12, 13)	
	RCK3 (10)	HuKIII (12, 13)	
		HuKV (12, 13)	
		HuKVI (12, 13)	
Π	DRK1 (14)		Shab (15)
Ш	RKShIIIA		Shaw? (15)
	(this paper)		
IV	· /		Shal (15)
V			Shaw? (15)

This table was modified from ref. 13. A homologue of RCK1, RBK1, and HuKI is known in mouse (MBK1, ref. 7).

Table 2. Fraction of identical amino acids in analogous positions between pairs of Sh K^+ channel proteins

	Shaw	DRK1	Shab	RCK1	H4
RKShIIIA	0.52	0.42	0.41	0.41	0.40
Shaw		0.41	0.40	0.41	0.41
DRK1			0.72	0.42	0.40
Shab				0.39	0.37
RCK1					0.76

The fractions of identical amino acids in overlapping sequences in the core region were calculated using an alignment like the one shown in Fig. 3 but which included also Shab and H4. H4 is one of the products of the *Drosophila* Shaker gene (4). Shab (15) and Shaw (15) are *Drosophila* genes, and DRK1 (14) and RCK1 (10) are rat genes.

class II and class III channels (DRK1 and RKShIIIA) activate relatively slowly and require large-depolarizations. RK-ShIIIA currents activate at particularly high voltages. These differences might explain why there seems to be much more expression of class I genes than of class II or III genes in brain, since most neuronal K^+ channels activate in the voltage range of class I channels. The three gene classes also differ in the number of positively charged residues in the S4 domain (Table 3).

Other properties such as developmental regulation, cellular localization, posttranslational modification, or participation in heteromultimer formation (31, 32) might turn out to be important properties of Sh channel classes. For example, RKShIIIA and Shaw lack a consensus sequence for cAMPdependent phosphorylation present in most Sh K⁺ channel sequences.

Functional Significance. One major task in the future of K^+ channel gene research is to understand the physiological significance of the various K^+ channel subunits that are being cloned. Because the functional properties of a given K^+ channel will depend on whether it is a homo- or heteromultimer of Sh-family gene products (19, 31, 32), on the interactions of this aggregate with smaller subunits (20, 33), and on posttranslational modifications (2), the results of the expression of a particular Sh gene will depend on the cellular context. It is interesting in this regard that the properties



FIG. 5. Southern analysis of rat genomic DNA. DNA was digested with an excess of EcoRI or Sac I. The Southern blots were hybridized (21) under conditions of relatively low stringency (30% formamide/5× SSPE at 37°C) with one of two probes. The first probe (A) contains the region that is most similar between Shaw and RKShIIIA (nucleotides 301-501); the second probe (B) contains the sequence encoding the H3 through H4 domains (nucleotides 964-1296), where Shaw and RKShIIIA are less similar. A single strong hybridization band and several faint but discrete hybridization bands are observed with both probes, suggesting the presence of several genes homologous to both regions of RKShIIIA. HindIII-digested λ DNA was used as size markers. Kb, kilobases.

Table 3. Electrophysiological properties of three classes of Sh K^+ channels

Class	Channel (ref.)	S4 <i>Z</i> *	$k_{\rm in}$, [†] msec	V _{on} ,‡ mV	t _{on} ,§ msec
I	H4 (19)	7	4	-40	5
	RCK1 (10)	7		-50	15
	RCK5 (10)	7		-55	6
	RHK1 (11)	7	30	-40	10
	RCK3 (10)	7	1000	-55	14
II	DRK1 (14)	5	_	-25	120
III	RKShIIIA	6	_	-10	80

Approximate values were determined from published data, at room temperature. All the experiments used a similar bathing solution except for those on DRK1, which used Co^{2+} instead of Ca^{2+} .

*Number of positive charge residues in S4.

[†]Time constant of inactivation at 0 mV. A dash implies inactivation is very slow or very small.

[‡]Activation voltage, the approximate voltage at which the conductance starts to rise.

[§]Activation rate at 0 mV determined as time to peak for rapidly inactivating currents, and time to rise between 10% and 90% for slowly inactivating and noninactivating currents.

observed here are very similar to those of one of the K⁺ currents (I_{Ks}) observed in *Xenopus* oocytes injected with whole brain poly(A)⁺ RNA (34). RKShIIIA channels activate at about the highest voltages one may expect physiologically. If RKShIIIA is expressed in the plasma membrane, and if its properties resemble those in the oocyte, we expect this channel to be particularly important in the termination of prolonged depolarizations (above 0 mV) such as plateau potentials or in the regulation of the frequency and duration of spike bursts (3). A channel such as RKShIIIA would play little role in subthreshold phenomena or in shaping the waveform of fast Na⁺ spikes.

Note Added in Proof. After this paper was submitted we became aware of a sequence, obtained from NG108 cells, just published by Yokoyama *et al.* (35). Their sequence (NGK2) is similar (but not identical) to that of RKShIIIA and appears to be the same as that of a related cDNA (bn13) from PC12 pheochromocytoma cells (Vega *et al.*, ref. 36). Sequence analysis suggests that NGK2 (or bn13) and RKShIIIA are two different members of the same gene class as defined here. Interestingly, although the amino acid sequence differences between NGK2 and RKShIIIA are similar to those found among class I gene members, their electrophysiological properties are very similar.

The research presented here would not have been possible if B.R. had not been given a joint appointment at the Department of Biochemistry at New York University Medical Center. We thank N. Godson (Chairman of the Department of Biochemistry) and R. Llinás (Chairman of the Department of Physiology and Biophysics) for their support and J. W. Lin for his tremendous help in many aspects of this work. We thank R. Dunn (McGill University) and L. Buck (Columbia University) for their cDNA libraries. B.R. would especially like to thank S. Kamb, L. Iverson, K. McCormack, M. Ramaswami, M. Mathew, M. Gautam and Prof. N. Davidson for their trust and patience in teaching him many of the techniques used here and would like to dedicate this paper to them. This research was supported by U.S. Public Health Service Grant GM26976.

- 1. Hille, B. (1984) *Ionic Channels of Excitable Membranes* (Sinauer, Sunderland, MA).
- 2. Rudy, B. (1988) Neuroscience 25, 729-750.

- 3. Llinás, R. (1988) Science 242, 1654-1664.
- Kamb, A., Tseng-Crank, J. & Tanouye, M. A. (1988) Neuron 1, 421-430.
- Schwarz, T., Tempel, B., Papazian, D., Jan, Y. N. & Jan, L. Y. (1988) Nature (London) 331, 137-142.
- Pongs, O. N., Kecskemethy, N., Muller, R., Krah-Jentgens, I., Baumann, A., Kiltz, H., Canal, I., Llamazares, S. & Ferrus, A. (1988) EMBO J. 7, 1087–1096.
- Tempel, B. L., Jan, Y. N. & Jan, L. Y. (1988) Nature (London) 332, 837–839.
- Christie, M. J., Adelman, J. P., Douglass, J. & North, A. J. (1989) Science 244, 221–224.
- 9. McKinnon, D. (1989) J. Biol. Chem. 264, 8230-8236.
- Stuhmer, W., Ruppersberg, J., Schroter, K., Sakmann, B., Stocker, M., Giese, K., Perschke, A., Baumann, A. & Pongs, O. (1989) EMBO J. 8, 3235–3244.
- Tseng-Crank, J., Tseng, G.-N., Schwartz, A. & Tanouye, M. A. (1990) FEBS Lett., in press.
- Kamb, A., Weir, M., Rudy, B., Varmus, H. & Kenyon, C. (1989) Proc. Natl. Acad. Sci. USA 86, 4372–4376.
- Mathew, M. K., Ramaswami, M., Gautam, M., Kamb, A., Rudy, B. & Tanouye, M. A. (1989) *Neurosci. Abstr.* 15, 540.
- Frech, G. C., VanDongen, A. M. J., Schuster, G., Brown, A. M. & Joho, R. H. (1989) Nature (London) 340, 642-645.
- 15. Butler, A., Wei, A., Baker, K. & Salkoff, L. (1989) Science 243, 943-947.
- Takumi, T., Ohkubo, H. & Nakanishi, S. (1988) Science 242, 1042-1045.
- 17. Numa, S. (1987) Biochem. Soc. Symp. 52, 119-143.
- 18. Catterall, W. A. (1988) Science 242, 50-60.
- Iverson, L., Tanouye, M., Lester, H., Davidson, N. & Rudy, B. (1988) Proc. Natl. Acad. Sci. USA 85, 5723-5727.
- Rudy, B., Hoger, J. H., Lester, H. A. & Davidson, N. (1988) Neuron 1, 649-658.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY).
- 22. Feinberg, A. P. & Vogelstein, B. (1983) Anal. Biochem. 132, 6-13.
- 23. Marshall, R. D. (1974) Biochem. Soc. Symp. 40, 17-26.
- McCormack, K., Campanelli, J. T., Ramaswami, M., Mathew, M. K., Tanouye, M. A., Iverson, L. E. & Rudy, B. (1989) *Nature (London)* 340, 103.
- 25. Putnam, F. W. (1987) in *The Plasma Proteins*, ed. Putnam, F. W. (Academic, New York), Vol. 5, pp. 49-140.
- 26. Eylar, E. H. (1972) Adv. Exp. Med. Biol. 32, 215-240.
- Sudhof, T. C., Czernik, A. J., Kao, H., Takei, K., Johnston, P. A., Horiuchi, A., Kanazir, S. D., Wagner, M. A., Perin, M. S., Camilli, P. D. & Greengard, P. (1989) Science 245, 1474-1480.
- Hart, G. W., Haltiwanger, R. S., Holt, G. D. & Kelly, W. G. (1989) Ciba Found. Symp. 145, 102–118.
- Bezanilla, F. & Armstrong, C. M. (1972) J. Gen. Physiol. 60, 588-608.
- Yeh, J. Z., Oxford, G. S., Wu, C. H. & Narahashi, T. (1976) J. Gen. Physiol. 68, 519-539.
- Haugland, F. N. & Wu, C. F. (1990) J. Neurosci. 10, 1357– 1371.
- McCormack, K., Lin, J. W., Ramaswami, M., Tanouye, M. A., Iverson, L. E. & Rudy, B. (1990) *Biophys. J.* 57, 209 (abstr.).
- Rehm, H. & Lazdunski, M. (1988) Proc. Natl. Acad. Sci. USA 85, 4919–4923.
- 34. Hoger, J. H., Ahmed, I., Davidson, N., Lester, H. & Rudy, B. (1987) Neurosci. Abstr. 13, 177.
- Yokoyama, S., Imoto, K., Kawamura, T., Higashida, H., Iwabe, N., Miyata, T. & Numa, S. (1989) FEBS Lett. 259, 37-42.
- Vega-Saenz de Miera, E., Sen, K., Serodio, P., McCormack, T. & Rudy, B. (1990) Neurosci. Abstr. 16, in press.