

The Human Aquaporin-5 Gene

MOLECULAR CHARACTERIZATION AND CHROMOSOMAL LOCALIZATION*

(Received for publication, November 6, 1995, and in revised form, January 17, 1996)

M. Douglas Lee†§¶, Kushal Y. Bhakta||, Surabhi Raina||, Raluca Yonescu**,
Constance A. Griffin**, Neal G. Copeland‡‡, Debra J. Gilbert‡‡, Nancy A. Jenkins‡‡,
Gregory M. Preston‡||, and Peter Agre‡||§§

From the †Departments of Medicine, §Division of Pulmonary and Critical Care Medicine, ||Biological Chemistry, and **Pathology, Johns Hopkins University School of Medicine, Baltimore Maryland 21205-2185 and the ‡‡Mammalian Genetics Laboratory, ABL-Basic Research Program, NCI-Frederick Cancer Research and Development Center, National Institutes of Health, Frederick, Maryland 21702-1201

The cDNA for the fifth mammalian aquaporin (AQP5) was isolated from rat, and expression was demonstrated in rat salivary and lacrimal glands, cornea, and lung (Raina, S., Preston, G. M., Guggino, W. B., and Agre, P. (1995) *J. Biol. Chem.* 270, 1908–1912). Here we report the isolation and characterization of the human AQP5 cDNA and gene. The AQP5 cDNA from a human submaxillary gland library contains a 795-base pair open reading frame encoding a 265-amino acid protein. The deduced amino acid sequences of human and rat AQP5 are 91% identical with 6 substitutions in the 22-amino acid COOH-terminal domain. Expression of human AQP5 in *Xenopus* oocytes conferred mercurial-sensitive osmotic water permeability (P_f) equivalent to other aquaporins. The human AQP5 structural gene resides within a 7.4-kilobase *SalI-EcoRI* fragment with four exons corresponding to amino acids 1–121, 122–176, 177–204, and 205–265 separated by introns of 1.2, 0.5, and 0.9 kilobases. A transcription initiation site was identified 518 base pairs upstream of the initiating methionine. Genomic Southern analysis indicated that AQP5 is a single copy gene which localized to human chromosome 12q13; this coincides with the chromosomal locations of the homologous human genes *MIP* and *AQP2*, thus confirming 12q13 as the site of an aquaporin gene cluster. The mouse gene localized to distal chromosome 15. This information may permit molecular characterization of AQP5 expression during normal development and in clinical disorders.

Discovery of the aquaporin family of water transporters provided a molecular explanation for osmotically driven water transport (P_f)¹ across cell membranes of mammalian and plant

tissues (reviewed by Chrispeels and Agre (1994)). When expressed in *Xenopus* oocytes, aquaporins confer large increases in P_f without conducting small molecules or ions. AQP1 is present in red cells, renal proximal tubules, lung, and other tissues (reviewed by Agre *et al.* (1993)). The major intrinsic protein of lens (MIP, AQP0) was recently confirmed as a weak water channel (Mulders *et al.*, 1995a). cDNAs encoding AQP2 through AQP5 were isolated by homology cloning (reviewed by Knepper (1994)). In response to vasopressin, AQP2 (Fushimi *et al.*, 1993) is targeted to the apical surface of renal collecting duct principal cells (Nielsen *et al.*, 1993). AQP3 is located at the basolateral membranes of renal collecting ducts (Ishibashi *et al.*, 1994, Ma *et al.*, 1994; Echevarria *et al.*, 1994). AQP4 is the major water channel in brain (Jung *et al.*, 1994; Hasegawa *et al.*, 1994). AQP5 is expressed in rat salivary and lacrimal glands, corneal epithelium, and lung (Raina *et al.*, 1995).

The human genes for several aquaporins have recently been characterized (Moon *et al.*, 1993; Uchida *et al.*, 1994; Pisano and Chepelinsky, 1991; Inase *et al.*, 1995; Mulders *et al.*, 1995b; Yang *et al.*, 1995). Identification of the Colton blood group antigens on the first exofacial loop of AQP1 led to the identification of three Colton-null individuals who had different mutations in AQP1, but total lack of AQP1 is not associated with an apparent phenotype (Smith *et al.*, 1994; Preston *et al.*, 1994). In contrast, a subset of patients with nephrogenic diabetes insipidus have mutations of the AQP2 gene resulting in a lack of clinical response to vasopressin (Deen *et al.*, 1994; van Lieburg *et al.*, 1994). Mice with mutations in MIP (AQP0) suffer from congenital cataracts (Shiels and Bassnett, 1996). Disease relevance of AQP3 and AQP4 remains to be determined. These studies reported here were undertaken to elucidate the gene structure and chromosomal localization of AQP5 as initial steps needed to understand regulation of the gene and to identify possible linkages to human disease.

MATERIALS AND METHODS

Isolation of Salivary cDNA Clones—The 1018-bp rat AQP5 cDNA containing the coding region plus 109-bp 5'-UTR and 114-bp 3'-UTR (Raina *et al.*, 1995) was labeled with [α -³²P]dCTP (Amersham) by random DNA labeling (Boehringer Mannheim). This probe was used to screen 1.5×10^5 plaques of an adult human submaxillary gland λ gt11 cDNA library (Clontech) under conditions of moderate stringency (0.5 \times SSC, 0.1% SDS, 55 °C). Two positive plaques were purified. DNA inserts were released with *EcoRI*, gel-purified, and ligated into pBS II KS(-) (Stratagene). Both strands were sequenced by the dideoxynucleotide termination method (U. S. Biochemical Corp.). A 1348-bp cDNA containing the entire coding region was utilized for all future studies.

* This work was supported in part by National Institutes of Health Grants EY11239, HL33991, HL48268, DK32753, and AI36973 and by the National Cancer Institute, DHHS, under contract with ABL. This work was presented at the Annual Meeting of the American Society of Cell Biology, Washington, D. C., December 10, 1995 (Lee, M. D., Preston, G. M., Bhakta, K. Y., Raina, S., Yonescu, R., Griffin, C. A., and Agre, P. (1995) *Mol. Biol. Cell* 6, 306 (abstr.)). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) U46566–U46569.

¶ Recipient of an Allen & Hanbury Pulmonary Fellowship Award.

§§ To whom correspondence should be addressed: Dept. Biological Chemistry, Johns Hopkins University School of Medicine, 725 N. Wolfe St., Baltimore, MD 21205-2185. Fax: 410-955-3149.

¹ The abbreviations used are: P_f , coefficient of osmotic water perme-

ability; UTR, untranslated region; RFLP, restriction fragment length polymorphism; bp, base pair(s); kb, kilobase(s); PIPES, 1,4-piperazineethanesulfonic acid.

Functional Expression in Oocytes—An 850-bp *SfiI*-*HindIII* fragment of the human *AQP5* cDNA containing the entire open reading frame along with 15-bp 5'-UTR and 40-bp 3'-UTR was ligated into the *BglII* site of the *Xenopus* expression construct pXβG, which contains the *HindIII*-*PstI* insert of pSP64T in pBS II KS(+) (Preston *et al.*, 1992). Capped cRNA was synthesized *in vitro* after digestion with *XbaI* (Preston *et al.*, 1993). Oocyte swelling was measured after cRNA expression (Raina *et al.*, 1995).

Isolation of Genomic DNA Clones—A human genomic DNA library in λEMBL-3 was generously provided by Dr. Jeremy Nathans, Johns Hopkins University (Wang *et al.*, 1992). The ³²P-labeled rat *AQP5* cDNA probe was used to screen 5 × 10⁵ plaques under conditions of moderate stringency. DNA from plaque-purified recombinants was digested with *EcoRI* and *SalI*, gel-purified, and ligated into pBS II KS(-).

Restriction Mapping and Determination of Exon-Intron Boundaries—The 7.4-kb *EcoRI*-*SalI* genomic clone was digested with restriction enzymes; DNA fragments were analyzed by Southern analysis with ³²P-labeled cDNA probes obtained from the 1348-bp human submaxillary cDNA clone. Nucleotide sequences corresponding to 5'- and 3'-UTR were determined using oligonucleotide primers to sense and antisense strands. Exon-intron boundaries were identified using oligonucleotide primers designed from coding regions of the human *AQP5* cDNA. Exon 1-intron 1 (sense, 310–326), 5'-ATCCTCTACGGTGTGGC-3'; intron 1-exon 2 (antisense, 389–405), 5'-CTCCACCACCATGGCCT-3'; exon 2-intron 2 (sense, 493–509), 5'-ATTGGCCTGTCTGTAC-3'; exon 3-intron 2 (antisense, 585–601), 5'-GGCTGAACCGATTTCATG-3'; exon 3-intron 3 (sense, 565–581), 5'-TCTTTTGGCCCTGCGGT-3'; exon 4-intron 3 (antisense, 683–699), 5'-ACTCAGGCTCAGGGAGT-3'. Sizes of introns 1–3 were determined by polymerase chain reaction using the intron flanking primers (30 cycles: 1 min 94 °C; 1 min 47 °C; 2 min 72 °C). Polymerase chain reaction products were electrophoresed into a 0.8% agarose gel with 1-kb DNA ladder standards.

Primer Extension—An oligonucleotide primer (designated S1) complementary to nucleotides +94 to +71 was end-labeled with [γ-³²P]ATP, and 1 × 10⁶ cpm were hybridized to 10 μg of total human lung or submaxillary RNA at 60 °C in 40 mM PIPES (pH 6.4), 1 mM EDTA, 400 mM NaCl, 80% formamide for 12 h. Following ethanol precipitation of hybridized products, primer extension was performed in 50 mM Tris-Cl (pH 7.6), 60 mM KCl, 10 mM MgCl₂, 2.5 mM dNTP, 1 mM dithiothreitol, 1 unit/μl RNase inhibitor, and 250 units Moloney murine leukemia reverse transcriptase (Life Technologies, Inc.) for 2 h at 37 °C. Synthesized products were analyzed on a 10% polyacrylamide sequencing gel alongside an adjacent sequencing ladder generated using the same oligo as a primer and a 600-bp *SalI*-*AclI* genomic fragment as a template.

Genomic Southern Analysis—Ten micrograms of human genomic DNA from leukocytes of two unrelated individuals was digested with *EcoRI*, *BamHI*, or *HindIII*, electrophoresed into a 0.8% agarose gel, and transferred to a nylon membrane (Genescreen Plus, DuPont NEN). Southern analyses were performed using the human *AQP5* cDNA as a probe (Moon *et al.*, 1993).

Chromosomal Localization and *In Situ* Hybridization—The plasmid containing the 7.4-kb genomic clone of human *AQP5* (Fig. 2) was nick-translated with biotin-14 dATP (Life Technologies, Inc.); 41% incorporation was determined by ³H incorporation. Chromosomal spreads of normal male lymphocytes cultured with BrUrd (Bhatt *et al.*, 1988) were used for fluorescence *in situ* hybridization (Moon *et al.*, 1993) and were photographed with color slide film (Kodak Ektachrome 400HC).

Interspecies Mouse Backcross Mapping—Mating of F₁ females (C57BL/6J × *Mus spretus*) and males (C57BL/6J) generated 205 N₂ mice which were used to map the *Aqp5* locus (Copeland and Jenkins, 1991). DNA isolation, restriction enzyme digestion, agarose gel electrophoresis, and Southern transfer to Hybond N⁺ (Amersham) nylon membranes were described (Jenkins *et al.*, 1982). The 1442-bp *EcoRI*-*XhoI* fragment of rat *AQP5* cDNA (Raina *et al.*, 1995) was ³²P-labeled with dCTP by nick translation (Boehringer Mannheim) and used to probe blots washed to a final stringency of 0.5 × SSCP, 0.1% SDS, 65 °C. An 11.0-kb fragment was detected in *SphI*-digested C57BL/6J DNA and a 9.4-kb fragment in *M. spretus* DNA. Presence or absence of the 9.4-kb *SphI* *M. spretus*-specific fragment was followed in backcross mice. Description of the probes and RFLPs for the loci linked to *Aqp5* including platelet-derived growth factor, endothelial cell (*Pdgfcd*), wingless-related Moloney murine tumor virus integration site 1 (*Wnt1*), and retinoic acid receptor γ (*Rarg*) were reported (Brannan *et al.*, 1992). Recombination distances were calculated by computer program SPRETUS MADNESS (Green, 1981). Gene order was determined by minimizing the number of recombination events required to explain the allele distribution pattern. References for the human map positions of cited loci are from GDB (Genome Data Base), a computerized data base of

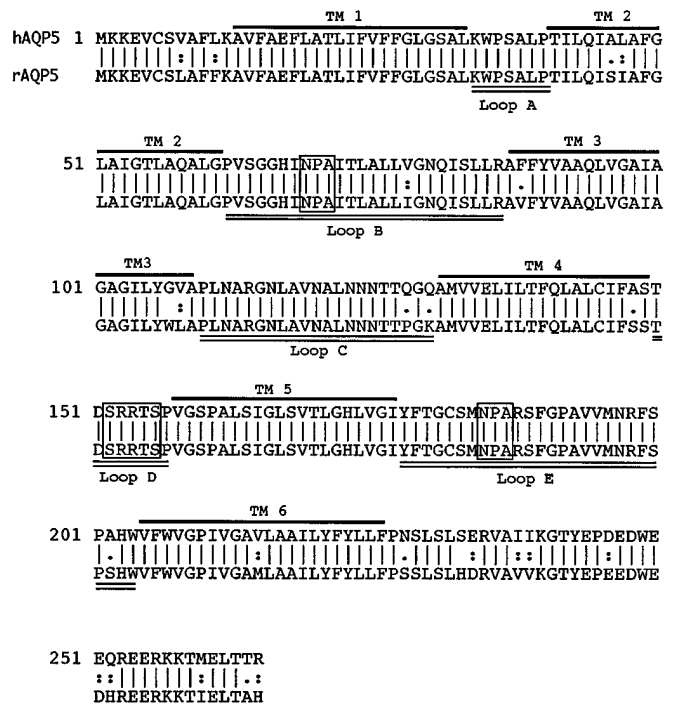


FIG. 1. Comparative alignment of deduced amino acid sequences of human and rat *AQP5* by GAP program analysis. Conserved NPA and cAMP-protein kinase motifs are enclosed in rectangles. Presumed bilayer-spanning domains are indicated by single bold lines (TM 1–TM 6). Polypeptide sequences within connecting loops A–E are double-underlined. Identical amino acids are joined with straight lines, while dissimilar amino acids are connected with dots or colons based on polarity and charge.

human linkage information maintained by the William H. Welch Medical Library, Johns Hopkins University (Baltimore, MD).

RESULTS

Isolation and Analysis of Human *AQP5* cDNA—The rat *AQP5* cDNA was used to isolate a recombinant from a human submaxillary gland cDNA library. The 1348-bp insert contained an open reading frame encoding a 265-amino acid protein. Further sequencing revealed the presence of 513 bp of 5'-UTR and only 40 bp of 3'-UTR. No other ATG translation initiation signals were identified within the 5'-UTR; the 3'-UTR did not contain a polyadenylation consensus sequence. Human and rat *AQP5* cDNAs were compared by GAP analysis program (Fig. 1). The deduced amino acid sequences were 91% identical. Like all members of the aquaporin family, tandem repeats of the NPA motif (Asn-Pro-Ala) were present in rat and human cDNAs. The consensus for cAMP-protein kinase A phosphorylation (SRRTS) present in rat *AQP5* was also in the human cDNA. Greatest divergence between the rat and human cDNAs was found within the 22-residue carboxyl-terminal domain where six substitutions were identified.

Transmembrane water flow by human *AQP5* was evaluated by expression in *Xenopus* oocytes and measurement of osmotic induced swelling. Oocytes injected with 5 ng of *AQP5* cRNA or 50 nl of water were incubated for 3 days at 18 °C. Oocytes were then transferred from 200 to 70 mosM modified Barth's solution, and increases in volume were detected by videomicroscopy. The human *AQP5* oocytes exhibited a 20-fold increase in P_f when compared to the water-injected oocytes; this increase in P_f was blocked by treatment with 1 mM HgCl₂, and the block was reversed by incubation with β-mercaptoethanol (Fig. 2). Thus, the human *AQP5* cDNA encodes an aquaporin which is functionally equivalent to the rat homolog.

Structure of the Human *AQP5* Genomic Locus—Three differ-

ent genomic clones were isolated from a human genomic DNA library using the rat cDNA as a molecular probe. The insert from one isolated plaque was released with *EcoRI* and *SaII* and subcloned. The restriction map for *AQP5* (Fig. 3) was confirmed by digestion with single restriction enzymes or combinations of enzymes followed by Southern analysis (data not shown).

Nucleotide sequencing with primers corresponding to the human *AQP5* cDNA revealed the exon-intron boundaries (Fig. 4). Exons 1–4 corresponded to amino acids 1–121, 122–176, 177–204, and 205–265. Nucleotide sequences of these exons were identical to sequences obtained from the human submaxillary cDNA. Exon-intron class 0 boundaries were identified for all four exons. Using exon-specific sense and antisense oligonucleotide primers, introns 1–3 were determined by polymerase chain reaction to be 1.2, 0.5, and 0.9 kb (Fig. 3). Sequencing of 500 bp of the 3'-flanking sequence of the human *AQP5*

genomic clone revealed a polyadenylation consensus sequence 490 bp from the last amino acid of AQP5 (Fig. 4).

Analysis of the 5'-flanking region of the human *AQP5* gene was undertaken to determine the site of transcription initiation and identify possible regulatory elements (Fig. 5A). Sequence obtained from the genomic clone was identical with that found in the 5'-UTR of the human *AQP5* cDNA. Using human lung mRNA as a template, a single transcription start site was identified 518 bp upstream from the translation start site of the *AQP5* gene (Fig. 5B). This site was 5 bp upstream of the 5' terminus of the cDNA clone isolated. A major band of identical size was identified using human submaxillary RNA (data not shown). RNase protection using antisense RNA to the DNA region of interest confirmed the presence of a major band of the same approximate size as that seen with primer extension (data not shown). Several common response elements (Fig. 5A) were identified upstream of the transcription initiation site (Prestridge, 1991; Ghosh, 1990). A defined TATA consensus was not identified within the 406 bp of 5'-flanking sequence upstream of the transcription initiation site (Fig. 5A).

Genomic Southern analyses were performed to determine if *AQP5* exists as a single copy gene. Human leukocyte DNA from two unrelated individuals was digested with restriction enzymes and hybridized at high stringency with the coding region to the human *AQP5* cDNA (Fig. 6). Identical hybridization patterns were found in the DNA from the two individuals. Based on the restriction map for the *AQP5* genomic clone (Fig. 3), the anticipated sizes of *BamHI*-digested DNA are 1.7 and 2.9 kb, which is consistent with the hybridization pattern ob-

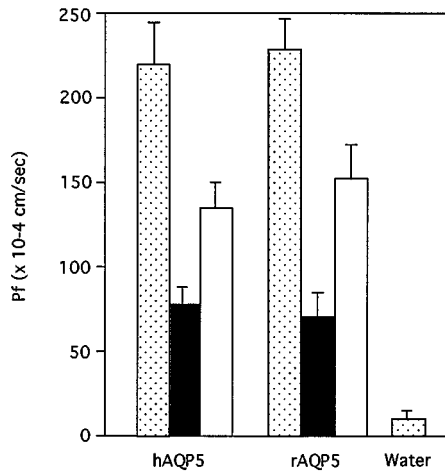


FIG. 2. Osmotic water permeability of oocytes expressing human and rat *AQP5*. Oocytes were injected with 5 ng of indicated cRNAs or 50 nl of water. Oocyte swelling was determined as described under "Materials and Methods." Depicted are mean values \pm S.D. of 4–6 oocytes (stippled bars), oocytes incubated for 5 min in 1 mM HgCl₂ (black bars), or oocytes incubated for 5 min in 1 mM HgCl₂ and subsequently incubated for 30 min in 5 mM β -mercaptoethanol (open bars).

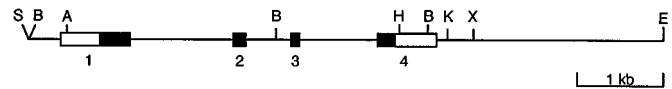


FIG. 3. Restriction map and exon-intron organization of human *AQP5*. The 7.4-kb *EcoRI*-*SaII* fragment of genomic DNA was used to determine the sites of exons 1–4 using ³²P-labeled probes prepared from the human *AQP5* cDNA (see "Materials and Methods"). Black rectangles represent coding regions, and open rectangles represent untranslated regions established by cDNA and genomic sequencing (see text). S = *SaII*, B = *BamHI*, A = *AccI*, H = *HindIII*, K = *KpnI*, X = *XhoI*, E = *EcoRI*.

5' utr	1 M	2 K	3 K	1	119 V	120 N	121 A	5' splice donor
..ctcgcgccgggcccccgccacc	ATG	AAG	AAG		GTC	AAC	GCG	gtgagtgcctgggggggtgggagc..
3' splice acceptor	122 L	123 N	124 N	2	174 L	175 V	176 G	5' splice donor
..gtcctaaccgctatccccttgacg	CTC	AAC	AAC		CTT	GTC	GGA	gtgagcagaccgacattgggctggg..
3' splice acceptor	177 I	178 Y	179 F	3	202 A	203 H	204 W	5' splice donor
..tctccaccgacctgtctctatccag	ATC	TAC	TTC		GCT	CAC	TGG	gtgagtctgtcccttcccctggctc..
3' splice acceptor	205 V	206 F	207 W	4	263 T	264 T	265 R	3' utr
..ctgactcgtgccctgtctccaccag	GTT	TTC	TGG		ACC	ACC	CGC	tgaccagtgtca477 bp aataaa..

FIG. 4. Exon-intron boundaries of the human *AQP5* gene. Nucleotide sequences surrounding the coding regions of each exon (numbered black boxes) along with corresponding 5' splice-donor and 3' splice-acceptor regions were determined by dideoxynucleotide sequencing using oligonucleotide primers generated from the *AQP5* cDNA. Partial 5'- and 3'-untranslated sequences with a polyadenylation consensus sequence (underlined) are represented.

A

sp1
aaggggaaccggcctgggagagggcgccctccggggtccgttgcctagtcagggtactg -346

cccagctaccgggcctgagagttgcaacggtcgggagaggtggcaccgtgcccacttt -285

tcccaaaactccagccttcccaagcccagaagctccccggcccagccgagctggccacgt -224

AP-2 sp1
 cggacggcggagcggcctcgaggacccagccggcgccgcggcggccgcccggcggaggtgaggg -163

sp1
 aggtgagcggcgccgacctgagggagcagcatcattccgccaccagcgggggtgagggcggg -102

tcaggatgctccggctgcaggagggaaagggagagctggaccaaaaagcccgaagagagaa -41

AGCTAGACGCCCGAGGTCGG
 aaggggaagcccgccacgggagcggtaagggcgccggAGCTAGACGCCCGAGGTCGG +21

AGTGAAGCGCCGGACCGAGCCCGTCTCCAGGGAGTCGGGGGGCAGCCGACCCGAGGAG +82

AGCGCGGGAGCCAACCTGGCGCATCATGCGCAGGGCCCGGGACGCTGGGCCGGTCTACAC +143

CGCCCGCTGGTGACGTGGCCGGAGCGGGCCGGCTGCCCGCCGGGGGGCGGGGTC +204

GCGCCGGGTGCTGCTGAGACGAGGAGCGGGCCGGCCCGCGAGCCCTGGAGCCCTCCAA +265

CCC GCCGCCTGGCCCTGAGCGTAGAGCCGCCCTGCCGCCCGCCGCCGCCCCGC +326

GCCCGCCCGCCCGCCCCATATAGCGCGCCCGCAGGAGGGCCCGCCGACGGCCGACGCT +387

CGGAGTGGGCGGGACAGTGGCGGGCCCGCCAGCCAGGCCCCGCCCCCGCCGCATCC +448

ACCTCTCCGCGCTCGCACCCAACGGCCGCCCGCCGGCAGCTCGCGCGGGCCCCC +509

GCGGCCACC ATG AAG AAG +527
 M K K

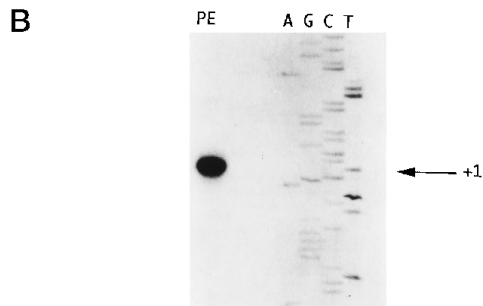


FIG. 5. Analysis of 5'-flanking region and primer extension of human *AQP5* gene. *A*, nucleotide sequence of 5'-flanking region. Nucleotide sequences of genomic DNA were obtained by dideoxynucleotide sequencing. The site of transcription initiation is indicated by an arrow and designated +1. Characteristic Sp1 and AP-2 sites are underlined. The first three codons of the *AQP5* cDNA are in bold italicized letters, and corresponding amino acid designations are below each codon. *B*, identification of transcription initiation by primer extension. A synthetic primer designated S1 (see "Materials and Methods") was end-labeled and hybridized to 10 μ g of total RNA from human lung overnight at 42 $^{\circ}$ C. Primer extension was performed as described under "Materials and Methods," and the isolated products were treated with RNase A (5 μ g, U. S. Biochemical Corp.). The final product (PE) was analyzed on a 10% polyacrylamide gel alongside sequencing products obtained by using a *SalI-AccI* genomic subclone (see Fig. 3) as a template and S1 as a sequencing primer to generate antisense DNA sequence.

served from the genomic Southern (Fig. 6). Moreover, digestion of the original phage clone with *EcoRI* yielded a fragment of approximately 7.7 kb (not shown) similar to that seen with human DNA (Fig. 6). Like the other aquaporins, *AQP5* exists as a single copy gene (Moon *et al.*, 1993; Pisano and Chepelinsky, 1991; Uchida *et al.*, 1994; Yang *et al.*, 1995).

Human Chromosomal Localization of *AQP5* Gene—The 7.4-kb genomic clone of the human *AQP5* gene was used as a molecular probe for *in situ* hybridization with chromosomal spreads from normal male lymphocytes (Fig. 7*A*). Analysis of 71 metaphase cells showed that 28 cells (40%) had at least one pair of signals involving both chromatids of a single chromosome. When these 28 metaphases were photographed on color

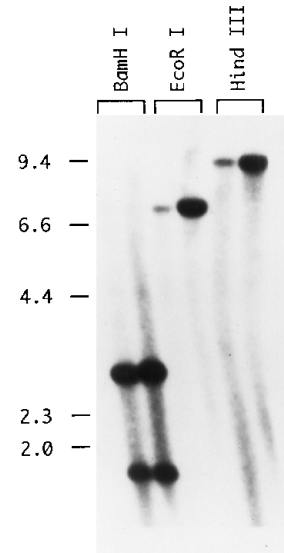


FIG. 6. Genomic Southern analysis with human *AQP5* cDNA. Ten μ g of DNA from two unrelated individuals was digested with *Bam*HI, *Eco*RI, or *Hind*III. Southern analysis was undertaken with a 32 P-labeled probe corresponding to the coding region of human *AQP5*.

slide film, 30 paired signals were seen. All signals were located on the long arm of a C-group chromosome (numbers 6–12). To determine the specific chromosome and band location of the signals, the hybridized slides were G-banded by fluorescence plus Giemsa, photographed, and aligned with the color slides to determine sub-band location (Fig. 7*B*). Twenty eight signals were analyzed after banding and were located on chromosome 12, band q13 (25 signals) and q12 (2 signals).

Mouse Chromosomal Localization of *Aqp5* Gene—Mouse genetic studies may provide another approach for analysis of the patterns of aquaporin gene expression. Mouse chromosomal localization of *Aqp5* was determined by interspecies backcross analysis using progeny derived from matings of ((C57BL/6J \times *M. spretus*) F1 \times C57BL/6J). This interspecies backcross mapping panel has been typed for over 2000 loci that are well distributed among all autosomes and the X chromosome (Copeland and Jenkins, 1991). C57BL/6J and *M. spretus* DNAs were digested with several enzymes and analyzed by Southern blot hybridization for informative RFLPs using the rat *AQP5* cDNA probe. The 9.4-kb *SphI* RFLP for *M. spretus* was used to follow segregation of the *Aqp5* locus in backcross mice. The mapping results identified *Aqp5* in the distal region of mouse chromosome 15 linked to *Pdgyf*, *Wnt1*, and *Rarg*. Although 104 mice were analyzed for every marker as shown in the segregation analysis (Fig. 8), up to 181 mice were typed for some pairs of markers. Each locus was analyzed in pairwise combinations for recombination frequencies using the additional data. Ratios of the total numbers of mice exhibiting recombinant chromosomes to the total number of mice analyzed for each pair of loci and the most likely gene order are: centromere-*Pdgyf*-7/112-*Wnt1*-0/151-*Aqp5*-1/181-*Rarg*. Note that *Wnt1* and *AQP5* were typed in common in 151 animals, suggesting that the two loci are within 2.0 centimorgans of each other (upper 95% confidence limit).

The interspecies map of mouse chromosome 15 was compared to a composite linkage map that reports map locations of many uncloned mouse mutations (Mouse Genome Data Base, a computerized data base maintained in The Jackson Laboratory, Bar Harbor, ME). *Aqp5* mapped to a region of the composite map that did not include any known mouse mutations with a phenotype consistent with defects in the tissues of *Aqp5* expression (not shown). Nevertheless, the distal region of

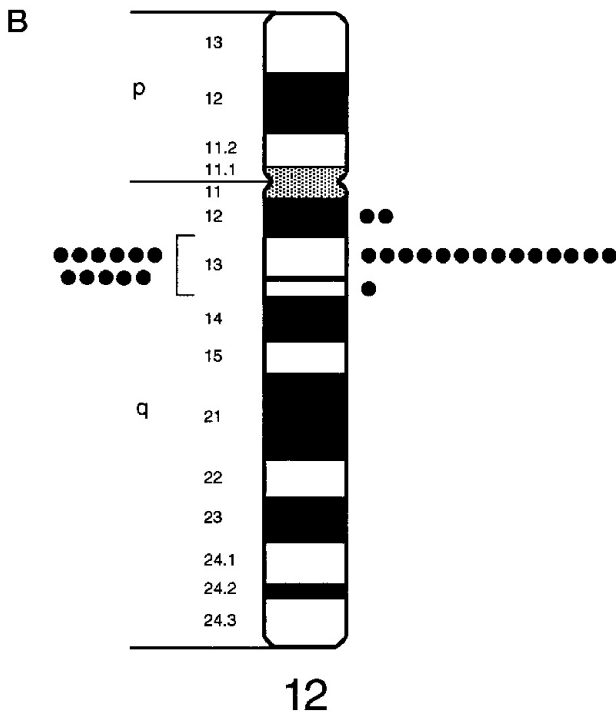
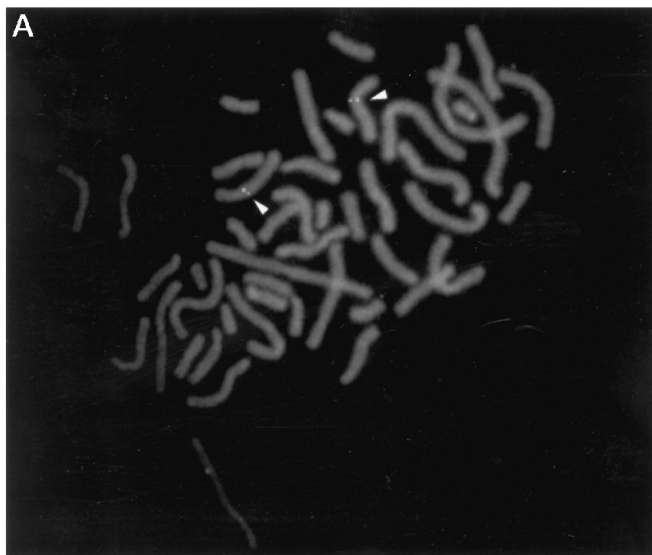


FIG. 7. **Chromosomal localization of human AQP5 genomic locus.** *A*, representative photograph demonstrating chromosomal spread from normal male lymphocytes hybridized with the 7.4-kb human AQP5 genomic clone. *In situ* hybridization of banded metaphase chromosomes localized all signals to the long arm of chromosome 12 where 25 of 27 signals were located on band 13. *B*, ideogram of human chromosome 12 showing localization of signals to 12q12–13. Each dot represents a paired signal seen on metaphase chromosomes. Signals clearly located on a single band are shown to the right. Signals which could not be localized to a single band are shown to the left.

mouse chromosome 15 shares homology with human chromosomes 22q and 12q (Fig. 7). In particular, *Wnt1* has been placed on human 12q13. The tight linkage between *Wnt1* and *Aqp5* suggests that *Aqp5* resides on 12q, which was confirmed by our *in situ* hybridization (Fig. 7).

DISCUSSION

These studies have defined the genomic organization of human AQP5. The AQP5 gene is structurally similar to MIP (AQP0), AQP1, and AQP2 (Pisano and Chepelinsky, 1991; Moon *et al.*, 1993; Uchida *et al.*, 1994). The lack of a defined

<i>Pdgfrc</i>	■ □	□ ■	□ ■	□ ■	□ ■	□ ■	□ ■
<i>Wnt1</i>	■ □	■ □	□ ■	□ ■	□ ■	□ ■	□ ■
<i>Aqp5</i>	■ □	■ □	■ □	■ □	□ ■	□ ■	□ ■
<i>Rarg</i>	■ □	■ □	■ □	■ □	■ □	■ □	■ □
	52	45	4	3	0	0	0

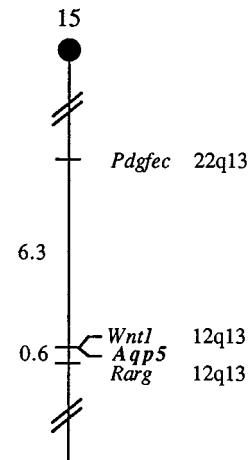


FIG. 8. **Aqp5 maps to the distal region of mouse chromosome 15 by interspecies backcross analysis.** *Top*, segregation patterns of *Aqp5* and flanking genes in 104 backcross animals typed for all loci. More than 104 animals were typed for individual pairs of loci (see text). Each column represents the chromosome identified in the backcross progeny that was inherited from the (C57BL/6J × *M. spretus*) F₁ parent. The shaded boxes represent the presence of a C57BL/6J allele, and white boxes represent the presence of a *M. spretus* allele. The number of offspring inheriting each type of chromosome is listed at the bottom of each column. A partial chromosome 15 linkage map showing the location of *Aqp5* in relation to linked genes is represented at the bottom of the figure. Recombination distances between loci in centimorgans are shown to the left of the chromosome, and the positions of known loci in human chromosomes are shown to the right.

TATA consensus within the region of the transcription initiation site of AQP5 is the first example of a TATA-less promoter in the aquaporin gene family. Transcription begins approximately 30 bp downstream of the TATA box in most eukaryotic genes (reviewed by Sawadogo and Sentenac, 1990). A growing number of eukaryotic genes lack a typical TATA box (reviewed by Roeder (1991) and Weis and Reinberg (1992)), but some of these genes contain an initiator element (Inr) that serves to organize the various transcription factors at the site of transcription initiation (Smale and Baltimore, 1989). Analysis of the AQP5 promoter failed to reveal any close sequence similarities with previously recognized Inr families (Weis and Reinberg, 1992), although there is currently no clear Inr consensus sequence (O'Shea-Greenfield and Smale, 1992). Studies directed at defining the active promoter regions of the AQP5 gene are currently underway to clarify this issue. The recognition of developmental expression patterns, identification of possible pharmacological modulation of expression, and definition of cellular and subcellular sites of expression may provide additional questions which may be answered by studies of AQP5 gene regulation.

Precise chromosomal localizations of candidate genes may permit linkage to mutant phenotypes. AQP1 has been mapped to human chromosome 7p14 (Moon *et al.*, 1993). The genes for MIP (AQP0) and AQP2 have both been localized to human chromosome 12q13 (Saito *et al.*, 1995). Although the chromosomal localization site for the AQP3 gene was reported to be

7q36 (Inase *et al.*, 1995), this has been corrected to 9p12-21 (Mulders *et al.*, 1995b). Nevertheless, localization of *AQP5* to human chromosome 12q13 is well supported: (i) fluorescence *in situ* hybridizations with large genomic DNA probes is highly reliable; (ii) *AQP5* colocalized with *Wnt1* to the distal arm of mouse chromosome 15 which corresponds to human chromosome 12q; (iii) colocalization of *AQP5* to the same human chromosomal region as *MIP* (*AQP0*) and *AQP2* identifies 12q13 as the site of an aquaporin gene cluster. To date, no obvious mutations in mice or humans suggesting *AQP5* dysfunction have been mapped to the *AQP5* gene locus. It is likely that *AQP5* participates in the generation of pulmonary secretions, saliva, and tears, as well as prevention of corneal edema. Thus, definition of the *AQP5* gene structure may aid in the identification of the role of *AQP5* in normal physiology and may possibly reveal clinical disorders related to this protein.

The chromosomal clustering of *MIP* (*AQP0*), *AQP2*, and *AQP5* may reflect similar modes of regulation of the proteins, since *AQP5* is most closely related to *AQP0* and *AQP2* at the amino acid level, and preliminary evidence indicates that phosphorylation is involved in the regulation of both *MIP* and *AQP2* (Kuwahara *et al.*, 1995; Ehring *et al.*, 1991). *AQP1* and *AQP3* are not part of this gene cluster and are believed to be constitutively activated (reviewed by Agre *et al.* (1995)). The conserved cAMP-protein kinase A consensus sequences within rat and human *AQP5* is similar to the consensus present in the vasopressin-regulated *AQP2*, suggesting that *AQP5* may be under neurohormonal control. Moreover, *AQP5* is expressed in salivary glands and lacrimal glands where rapid fluid shifts occur in response to adrenergic stimuli (reviewed by Nauntofte (1992)). The mechanism by which *AQP2* is targeted to the cell surface in response to phosphorylation has been the subject of intense recent investigation (reviewed by Agre *et al.* (1995)). Studies designed to determine the mechanism of *AQP5* protein regulation as well as *AQP5* gene expression are currently underway.

Acknowledgments—We thank Giuseppe Calamita, Landon King, and William B. Guggino for helpful suggestions. We also thank Mary Barnstead and Anita Hawkins for excellent technical assistance.

REFERENCES

- Agre, P., Preston, G. M., Smith, B. L., Jung, J. S., Raina, S., Moon, C., Guggino, W. B., and Nielsen, S. (1993) *Am. J. Physiol.* **265**, F463–F476
- Agre, P., Brown, D., and Nielsen, S. (1995) *Curr. Opin. Cell Biol.* **7**, 472–483
- Bhatt, B., Burns, J., Flannery, D., and McGee, J. (1988) *Nucleic Acids Res.* **16**, 3951–3961
- Brannan, C. I., Gilbert, D. J., Ceci, J. D., Matsuda, Y., Chapman, V. M., Mercer, J. A., Eisen, H., Johnston, L. A., Copeland, N. G., and Jenkins, N. A. (1992) *Genomics* **13**, 1075–1081
- Chrispeels, M. J., and Agre, P. (1994) *Trends Biochem. Sci.* **19**, 421–425
- Copeland, N. G., and Jenkins, N. A. (1991) *Trends Genet.* **7**, 113–118
- Deen, P. M. T., Verdijk, M. A. J., Knoers, N. V. A. M., Wieringa, B., Monnens, L. A. H., van Os, C. H., and van Oost, B. A. (1994) *Science* **264**, 92–95
- Echevarria, M., Windhager, E. E., Tate, S. S., and Frindt, G. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 10997–11001
- Ehring, G. R., Lagos, N., Zampighi, G. A., and Hall, J. E. (1991) *J. Membr. Biol.* **126**, 75–88
- Fushimi, K., Shinichi, U., Hara, Y., Hiratya, Y., Marumo, F., and Sasaki, S. (1993) *Nature* **361**, 549–552
- Ghosh, D. (1990) *Nucleic Acids Res.* **21**, 3113–3118
- Green, E. L. (1981) *Genetics and Probability in Animal Breeding Experiments*, pp. 77–113, Oxford University Press, New York
- Hasegawa, H., Ma, T., Skach, W., Matthay, M. A., and Verkman, A. S. (1994) *J. Biol. Chem.* **269**, 5497–5500
- Inase, N., Fushimi, K., Ishibashi, K., Uchida, S., Ichioka, M., Sasaki, S., and Marumo, F. (1995) *J. Biol. Chem.* **270**, 17913–17916
- Ishibashi, K., Sasaki, S., Fushimi, K., Uchida, S., Kuwahara, M., Saito, H., Furukawa, T., Nakajima, K., Yamaguchi, Y., Gojobori, T., and Marumo, F. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6269–6273
- Jenkins, N. A., Copeland, N. G., Taylor, B. A., and Lee, B. K. (1982) *J. Virol.* **43**, 26–36
- Jung, J. S., Bhat, R. V., Preston, G. M., Guggino, W. B., Baraban, J. M., and Agre, P. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 13052–13056
- Knepper, M. A. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 6255–6258
- Kuwahara, M., Fushimi, K., Terada, Y., Bai, L., Marumo, F., and Sasaki, S. (1995) *J. Biol. Chem.* **270**, 10384–10387
- Ma, T., Frigeri, A., Hasegawa, H., and Verkman, A. S. (1994) *J. Biol. Chem.* **269**, 21845–21849
- Moon, C., Preston, G. M., Griffin, C. A., Jabs, E. W., and Agre, P. (1993) *J. Biol. Chem.* **268**, 15772–15778
- Mulders, S. M., Preston, G. M., Deen, P. M. T., Guggino, W. B., van Os, C. H., and Agre, P. (1995a) *J. Biol. Chem.* **270**, 9010–9016
- Mulders, S. M., Weghuis, D. O., van Boxel, J. A. F., van Kessel, A. G., Echevarria, M., van Os, C. H., and Deen, P. M. T. (1995b) *Aquaporins and Epithelial Water Transport International Symposium*, p. 58 (abstr.), University of Manchester Press, Manchester, UK
- Nauntofte, B. (1992) *Am. J. Physiol.* **263**, G823–G837
- Nielsen, S., DiGiovanni, S. R., Christensen, E. I., Knepper, M. A., and Harris, H. W. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 11663–11667
- O'Shea-Greenfield, A., and Smale, S. T. (1992) *J. Biol. Chem.* **267**, 1391–1402
- Pisano, M. M., and Chepelinsky, A. B. (1991) *Genomics* **11**, 981–990
- Preston, G. M., Carroll, T. P., Guggino, W. B., and Agre, P. (1992) *Science* **256**, 385–387
- Preston, G. M., Jung, J. S., Guggino, W. B., and Agre, P. (1993) *J. Biol. Chem.* **269**, 1668–1673
- Preston, G. M., Smith, B. L., Zeidel, M. L., Moulds, J., and Agre, P. (1994) *Science* **265**, 1585–1587
- Prestridge, D. S. (1991) *Comput. Appl. Biosci.* **7**, 203–206
- Raina, S., Preston, G. M., Guggino, W. B., and Agre, P. (1995) *J. Biol. Chem.* **270**, 1908–1912
- Roeder, R. G. (1991) *Trends Biochem. Sci.* **16**, 402–408
- Saito, F., Sasaki, S., Chepelinsky, A. B., Fushimi, K., Marumo, F., and Ikeuchi, T. (1995) *Cytogenet. Cell Genet.* **68**, 45–48
- Shiels, A., and Bassnett, S. (1996) *Nature Genet.* **12**, 212–215
- Smale, S. T., and Baltimore, D. (1989) *Cell* **57**, 103–113
- Smith, B. L., Preston, G. M., Spring, F. A., Anstee, D. J., and Agre, P. (1994) *J. Clin. Invest.* **94**, 1043–1049
- Sawadogo, M., and Sentenac, A. (1990) *Annu. Rev. Biochem.* **59**, 711–754
- Uchida, S., Sasaki, S., Fushimi, K., and Marumo, F. (1994) *J. Biol. Chem.* **269**, 23451–23455
- van Lieburg, A. F., Verdijk, M. A., Knoers, V. V., van Essen, A. J., Proesmans, W., Mallmann, R., Monnens, L. A., van Oost, B. A., van Os, C. H., and Deen, P. D. (1994) *Am. J. Hum. Genet.* **55**, 648–652
- Wang, Y., Macke, J. P., Merbs, S. L., Zack, D. J., Klaunberg, B., Bennett, J., Gearhart, J., and Nathans, J. (1992) *Neuron* **9**, 429–440
- Weis, L., and Reinberg, D. (1992) *FASEB J.* **6**, 3300–3309
- Yang, B., Ma, T., and Verkman, A. S. (1995) *J. Biol. Chem.* **270**, 22907–22913