The human AQP4 gene: Definition of the locus encoding two water channel polypeptides in brain

MINGQI LU*,†, M. DOUGLAS LEE*,†, BARBARA L. SMITH*, JIN SUP JUNG*, PETER AGRE*,†, MARIAN A. J. VERDIJK*, GERARD MERKK*, JOHAN P. L. RUSS†, and PETER M. T. DEEN‡

*Departments of Biological Chemistry and Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205-2185; and Departments of †Human Genetcs and ‡Cell Physiology, University of Nijmegen, 6500 HB, Nijmegen, The Netherlands

Communicated by Victor McKusick, Johns Hopkins Hospital, Baltimore, MD, July 15, 1996 (received for review May 31, 1996)

ABSTRACT The aquaporin family of membrane water transport proteins are expressed in diverse tissues, and in brain the predominant water channel protein is AQP4. Here we report the isolation and characterization of the human AQP4 cDNAs and genomic DNA. Two cDNAs were isolated corresponding to the two initiating methionines (M1 in a 323-aa polypeptide and M23 in a 301-aa polypeptide) previously identified in rat [Jung, J. S., Bhat, R. V., Preston, G. M., Guggino, W. B. & Agre, P. (1994) Proc. Natl. Acad. Sci. USA 91, 13052-13056]. Similar to other aquaporins, the AQP4 gene is composed of four exons encoding 127, 55, 27, and 92 amino acids separated by introns of 0.8, 0.3, and 5.2 kb. Unlike other aquaporins, an alternative coding initiation sequence (designated exon 0) was located 2.7 kb upstream of exon 1. When spliced together, M1 and the subsequent 10 amino acids are encoded by exon 0; the next 11 amino acids and M23 are encoded by exon 1. Transcription initiation sites have been mapped in the proximal promoters of exons 0 and 1. RNase protection revealed distinct transcripts corresponding to M1 and M23 mRNAs, and AQP4 immunoblots of cerebellum demonstrated reactive polypeptides of 31 and 34 kDa. Using a P1 and a λEMBL subclone, the chromosomal site of the human AQP4 gene was mapped to chromosome 18 at the junction of q11.2 and q12.1 by fluorescence in situ hybridization. These studies may now permit molecular characterization of AQP4 during human development and in clinical disorders.

Recognition of the aquaporin family of water transport proteins has provided a molecular explanation for the high water permeability of membranes of certain tissues. Characterization of the organization of aquaporin genes and identification of their sites within the human and mouse genomes have established a primary role for some aquaporins in specific clinical disorders (for review, see ref. 1). Recognition that Colton blood group antigens are polymorphisms on red cell AQP1 (2) led to the recognition that the AQP1 null phenotype is not clinically severe and the prediction that backup water transport mechanisms exist (3). In contrast, mutations in AQP2 cause a severe form of nephrogenic diabetes insipidus (4), and mutations in the MIP gene (encoding a lens aquaporin) cause congenital cataracts in mice (5).

AQP4 is known to be expressed predominantly in brain (6, 7). Although AQP4 may contribute to the pathophysiology of normal pressure hydrocephalus, pseudotumor cerebri, or postinscnic edema, such linkages have not yet been established. To facilitate such studies, isolation and characterization of human AQP4 cDNAs and genomic DNAs were pursued. While these studies were underway, another group reported on the organization and localization of the gene encoding MIWC (an alternate name for AQP4; ref. 8). Although several observations were in agreement with ours, significant incompatibilities were noted, including unrelated N-terminal sequences, multiple discrepancies in the coding region, dissimilarities in the genomic organization, and conflicting localizations of the gene to different sites on human chromosome 18. Additional studies were therefore undertaken to reconcile these differences.

MATERIALS AND METHODS
cDNA and Genomic DNA Cloning. A 611-bp Perl fragment from the coding region of rat AQP4 cDNA (6) was random-labeled with [α-32P]dCTP (Boehringer Mannheim, Amersham); this was used to probe 5 × 105 plaques from a human fetal brain cDNA library (AZAP; Stratagene) and an adult human retina cDNA library (AMaxi; CLONTECH). Purified clones were sequenced with dyeoxyxynucleotide chain termination (United States Biochemical). A human genomic DNA library in λEMBL-3 (9) was screened with the rat AQP4 cDNA fragment, and positives were restriction-digested and subcloned for Southern analysis and sequencing (10). Intron sizes were estimated by restriction mapping and PCR amplification of genomic DNA. A human P1 genomic DNA library (11, 12) was screened similarly, and presence of AQP4 in clone ICRF P700M1638 (11) was confirmed by cycle sequencing for exons 1, 2, 3, and 4 (4).

RNA Studies. To determine the site of transcription initiation of exon 0, an antisense oligonucleotide primer 5'-AGCCAGAGTGCAGCTCTCAT-3' was end-labeled with [α-32P]ATP; 1 × 106 cpm was hybridized to 10 μg of total human brain RNA and extended (10). Transcription initiation for exon 1 was undertaken similarly with the antisense primer 5'-AAGATCAGCTTCTGACGAA-3'. The template plasmid for RNase protection assays (6) was constructed with a 200-bp EcoRI–BamHI fragment of M1AQP4 cDNA (corresponding to 70 bp of exon 0 and 130 bp encoding exon 1) inserted into pBluescript II KS(−) and linearized with XbaI or KpnI to generate sense or antisense RNAs. A uniformly labeled 311-nt antisense probe was synthesized from the T3 promoter with [α-32P]UTP (800 Ci/mmol (1 Ci = 37 GBq); Amersham); 1010 cpm was hybridized at 42°C to 10 μg of total RNA from adult human brain or 60 μg of RNA from lung or salivary gland (CLONTECH). After digestion with RNase A (20 μg/ml) and RNase T1 (200 units/ml), protected fragments (3 × 105 cpm) were analyzed on 10% polyacrylamide sequencing gels next to sequencing standards. Sense RNAs generated from the T7 promoter (5 ng) or yeast tRNA (10 μg) were hybridized and digested similarly.

Data deposition: The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U63611–U63623).

To whom reprint requests should be addressed at: Department of Biological Chemistry, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205-2185. e-mail: peter-agre@jmail.jhu.edu.

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.
RESULTS

Isolation and Characterization of Human AQP4 cDNAs. Nine unique clones were isolated from a human fetal brain cDNA library, and two were isolated from an adult human retina cDNA library by probing with the rat AQP4 cDNA (6). The longest brain clone contained 40 bp of 5'-untranslated sequence, an open reading frame encoding a 323-aa polypeptide with two potential initiating methionines (corresponding to M1 and M23 from rat), followed by ~3 kb of 3'-untranslated DNA with a polyadenylation consensus. Neither of the retina clones extended to an initiation site, but a cDNA was amplified from retina RNA with a sense primer designed to the M23 conservative substitutions (Fig. 1). The nucleotide sequences of both DNA strands of cDNA library, and two were isolated from an adult human brain and retina cDNAs were identical from M23 to the initiation site at nucleotide +42, whereas the farther TATA boxes were identical to the human brain and retina cDNAs. Intron splicing was determined by restriction mapping and PCR using sense and antisense primers designed from the following coding sequences: a 0.8-kb intron between exons 1 and 2; 278 bp between exons 2 and 3; and 5.2 kb between exons 3 and 4 (Fig. 2). Thus splicing of exons 1–4 would correspond to the M23 cDNA from retina (Fig. 2).

The existence of another exon (here designated exon 0) was inferred because the cDNA sequence at the 5' end of the M1 cDNA was not present in exons 1 and 2. The first 34 bp in the 5' DNA flanking exon 1 corresponded to the sequence in the M1 cDNA and was preceded by an intron splice acceptor site. By primer extension analysis and sequencing with oligonucleotides corresponding to M1 and M23 mRNAs were seen in all three human tissues, although 6-fold higher concentrations of lung and salivary gland RNA were needed to create signals of intensity equal to that of brain. When brain RNA was present, the 200-nt fragment of intron 0 was used to determine the sites of transcription initiation for the M1 cDNA (Fig. 3). The resulting 10 and 2/3 amino acids in the M1 transcript followed by a class II splice donor site. The next 11 and 1/3 amino acids preceding M23 were encoded by the 5' DNA flanking exon 1 (Figs. 2 and 3).

Transcription and Translation of AQP4. Primer extension was used to determine the sites of transcription initiation for exons 0 and 1 using human brain RNA as a template. Two potential transcription initiation sites were identified upstream from M1 in exon 0 (Fig. 4A). Although two TATA motifs were also identified, only the closer may be functional, because it resides just 28 bp upstream from the second transcription initiation site at nucleotide +42, whereas the farther TATA motif lies 173 bp from the first initiation site (Figs. 3 and 4A). Mapping the transcription initiation site of the M23 mRNA was undertaken similarly and revealed a start site 36 bp downstream from the single TATA element (Fig. 4A), which differs from a recent report (8).

To confirm that two AQP4 transcripts exist, RNase protection was performed with total RNA from human brain, submandibular salivary gland, and lung using a 32P-labeled antisense cRNA probe overlapping sequences in exons 0 and 1 (Fig. 4B). Two distinct protection fragments corresponding to M1 and M23 mRNAs were seen in all three human tissues, although 6-fold higher concentrations of lung and salivary gland RNA were needed to create signals of intensity equal to that of brain. When brain RNA was present, the 200-nt fragment (corresponding to M1 mRNA) was of higher intensity than the 130-nt fragment (corresponding to M23 mRNA), but the fragments were equivalently protected by RNA from the other two human tissues.

The potential functional significance of the two AQP4 transcripts was analyzed by immunoblotting brain membranes. Because of proteolysis in postmortem human tissues, rat cerebellum was evaluated with anti-AQP4 (Fig. 4C). Reactive bands were observed with electrophoretic mobilities of 31, 34, 39, and 64 kDa. The 31- and 34-kDa bands are in agreement with the predicted sizes of AQP4 protein translated from M1 and M23 cDNA.
Fig. 2. Exon-intron organization of the human AQP4 gene. A segment of the 20-kb genomic AQP4 clone is represented at the top of the figure. Shown are untranslated regions of exons 0 and 1, and the first polyadenylation consensus of exon 4 (open rectangles), and coding sequences (filled rectangles). The sequence that encodes residues 12-22 in M1 mRNA but is not translated in M23 mRNA is represented by a hatched rectangle. Restriction sites are also denoted. B, BamHI; S, Sall; E, EcoRI). Spliced forms corresponding to M1 mRNA and M23 mRNA are depicted at the bottom of the figure with translation initiation sites (arrows).

Fig. 3. Genomic sequences and exon-intron boundaries of the human AQP4 gene. Shown are nucleotide sequences of the 5' flanking regions and 5' and 3' intron splices (lower case letters) and 5' untranslated and coding sequences (upper case letters). Deduced amino acids are listed below nucleotide sequences; residues unique to M1 mRNA are enclosed in parentheses. The methionines for M1 and M23 mRNAs conform to the translation initiation consensus (21) and are indicated by an asterisk. The 5'- and 3'-terminal amino acids of each exon are numbered relative to M1 mRNA; numbers corresponding to interior nucleotide and deduced amino acid sequences for exons 2, 3, and 4 are enclosed in brackets. Potential transcription initiation sites of exon 0 and exon 1 identified by primer extension (see Fig. 4) are indicated by arrows, and characteristic tatata motifs are underlined.

and M23 (32.3 and 34.8 kDa), although the abundance of the 31-kDa band was greater than the 34-kDa band. The 59- and 64-kDa bands may represent dimers but do not represent N-glycosylated subunits, because their electrophoretic mobilities were not altered by digestion with Peptide:N-glycosidase F (New England Biolabs; data not shown).

Chromosomal Localization of the AQP4 Gene. Determination of the chromosomal site for AQP4 was established by fluorescence in situ hybridizations. The 5.8-kb BamHI genomic clone containing exon 0 and part of exon 1 yielded fluorescence signals localized at chromosome 18 q11-q12 (data not shown). To establish the chromosomal location with highest
The homeodomain of the drosophila protein Is upregulated in response to mechanical stimuli 
and plays a role in the regulation of the Notch signaling pathway. This protein is known to be involved in 
the development of the nervous system and is expressed in the embryonic CNS. The homeodomain 
consists of a highly conserved DNA binding motif that is responsible for recognizing specific DNA 
sequences. In this study, we investigated the expression pattern of the drosophila Is homeodomain 
protein in response to mechanical stimuli. We found that the expression of the Is homeodomain 
protein is upregulated in response to mechanical stimuli, suggesting a role in the regulation of 
Notch signaling pathway.

DISCUSSION

The results presented in this study indicate that the expression of the Is homeodomain protein is 
upregulated in response to mechanical stimuli. This is consistent with previous studies that have 
shown a role for the Is homeodomain protein in the regulation of the Notch signaling pathway. 
Furthermore, the results presented in this study provide new insights into the role of mechanical 
stimuli in the regulation of the Notch signaling pathway.