Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 13052–13056, December 1994
Physiology

Molecular characterization of an aquaporin cDNA from brain: Candidate osmoreceptor and regulator of water balance

(water channel/vasopressin-secretory neurons/Purkinje cells/ependymal cells/cerebrospinal fluid)

JIN SUP JUNG*, RATAN V. BHAT, GREGORY M. PRESTON, WILLIAM B. GUGGINO, JAY M. BARABAN, AND PETER AGRE†

Departments of Biological Chemistry, Medicine, Neuroscience, and Physiology, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, MD 21205

Communicated by Maurice B. Burg, August 29, 1994 (received for review July 14, 1994)

ABSTRACT **The aquaporins transport water through membranes of numerous tissues, but the molecular mechanisms for sensing changes in extracellular osmolality and regulating water balance in brain are unknown. We have isolated a brain aquaporin by homology cloning. Like aquaporin 1 (AQP1, also known as CHIP, channel-forming integral membrane protein of 28 kDa), the deduced polypeptide has six putative transmembrane domains but lacks cysteines at the known mercury-sensitive sites. Two initiation sites were identified encoding polypeptides of 301 and 323 amino acids; expression of each in *Xenopus* oocytes conferred a 20-fold increase in osmotic water permeability not blocked by 1 mM $HgCl_2$, even after substitution of cysteine at the predicted mercury-sensitive site. Northern analysis and RNase protection demonstrated the mRNA to be abundant in mature rat brain but only weakly detectable in eye, kidney, intestine, and lung. *In situ* hybridization of brain localized the mRNA to ependymal cells lining the aqueduct, glial cells forming the edge of the cerebral cortex and brainstem, vasopressin-secretory neurons in supraoptic and paraventricular nuclei of hypothalamus, and Purkinje cells of cerebellum. Its distinctive expression pattern implicates this fourth mammalian member of the aquaporin water channel family (designated gene symbol, *AQP4*) as the osmoreceptor which regulates body water balance and mediates water flow within the central nervous system.**

The aquaporins are a family of water-selective membrane channels found in animals, plants, and microorganisms (reviewed in refs. 1 and 2). Aquaporin 1 (AQP1, also known as CHIP, channel-forming integral membrane protein of 28 kDa) was the first protein shown to function as a molecular water channel (3) and is naturally expressed in mammalian red cells, renal proximal tubules (4–6), and other water-permeable epithelia (7). AQP2 is the vasopressin-regulated water channel in renal collecting ducts (8, 9) and is the site of mutations in some forms of nephrogenic diabetes insipidus (10). AQP3 is the water channel in basolateral membranes of renal medullary collecting duct (11).

Because of restricted space within the cranium, regulation of salt and water balance is essential for normal functions of the mammalian brain (reviewed in ref. 12). Moreover, vasopressin is released by the neurohypophysis in response to small changes in osmolality around the supraoptic and paraventricular nuclei (13). AQP1 (CHIP) is abundant in the choroid plexus, but not elsewhere in brain where molecular mechanisms responsible for transmembrane water movements and osmoreception are still unknown (7). By homology cloning we have isolated a cDNA from a rat brain library and have established its function and distribution. As this work was nearing completion, a highly related cDNA was cloned from rat lung (14); however, discrepancies with our studies were observed in the initiation site, in the coding and 3′ untranslated sequences, and in the sites of expression and relative abundance. Thus, additional studies were undertaken to resolve these differences and further define this aquaporin.‡

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.

METHODS

cDNA Cloning. Nested, degenerate oligonucleotide primers were designed to correspond to highly conserved sequences surrounding the Asn-Pro-Ala (NPA) motifs in the aquaporins (15): sense primers were MDU-1 (5′-STBGGNCAYRTBAGYGGNGCNCA-3′) and MDU-2 (5′GGGATCCGCHCAYNTNAAYCCHGYNGTNAC-3′); antisense primers were MDD-1 (5′-GCDGRNSCVARDGANCGNGCNGG-3′) and MDD-2 (5′-CGGAATTCGDGCDGGRTTNATNSHNSMNCC-3′). Rat brain mRNA (1 μg) was reverse transcribed and amplified by 30 cycles (94°C, 1 min; 52°C, 1 min; 72°C, 1 min) of polymerase chain reaction (PCR) using 100 pmol of MDU-1 and MDD-1; 10 μl of products was reamplified with 100 pmol of MDU-2 and MDD-2. PCR products of ≈360 bp were subcloned and sequenced by double-strand dideoxynucleotide termination (United States Biochemical). A ^{32}P-labeled cDNA probe was prepared by the random DNA labeling system (Boehringer Mannheim) and used to screen 4.5×10^5 plaques from a rat brain cDNA library (λZAP; Stratagene).

Determination of the nucleotide sequence corresponding to the amino terminus was accomplished by rapid amplification of cDNA 5′ ends (5′-RACE) by PCR. Rat brain poly(A)$^+$ RNA (1 μg) was reverse transcribed into cDNA by using an oligonucleotide primer derived from the 5′-end sequence of the 1.6-kb clone and a gene-specific primer 1 (GSP-1, 5′-GGCTTGAGTCCAGACGCCTTTGAAAGCCACC-3′). Poly(dA) was added to the 5′ termini of the cDNAs by terminal deoxynucleotidyltransferase (BRL). The (dA)-tailed cDNA (5 μl) was amplified by 40 cycles of PCR using *Taq* polymerase with an annealing temperature of 45°C, an oligo(dT) primer (5′GATGGATCCTGCAGAAGCTTTTTTTTTTTTTTTT-3′), and the GSP-1 primer. The product was reamplified with a 5′ anchor primer (5′-GATGGATCCTGCAGAAGC-3′) and a second gene-specific primer (GSP-2, 5′-CGGAATTCCAGGGAGGTCCACACTTAC-3′) located just upstream of GSP-1. The coding regions for M1 (bp −9 to +1310) and M23 (bp +7 to +1310) were cloned into the *Xenopus* oocyte expression construct pXβG,

Abbreviations: AQP*n*, aquaporin *n*; 5′-RACE, rapid amplification of cDNA 5′ ends.

*Present address: Department of Physiology, Pusan National University, Korea.
†To whom reprint requests should be addressed.

‡The sequence reported in this paper has been deposited in the GenBank data base (accession no. U14007).

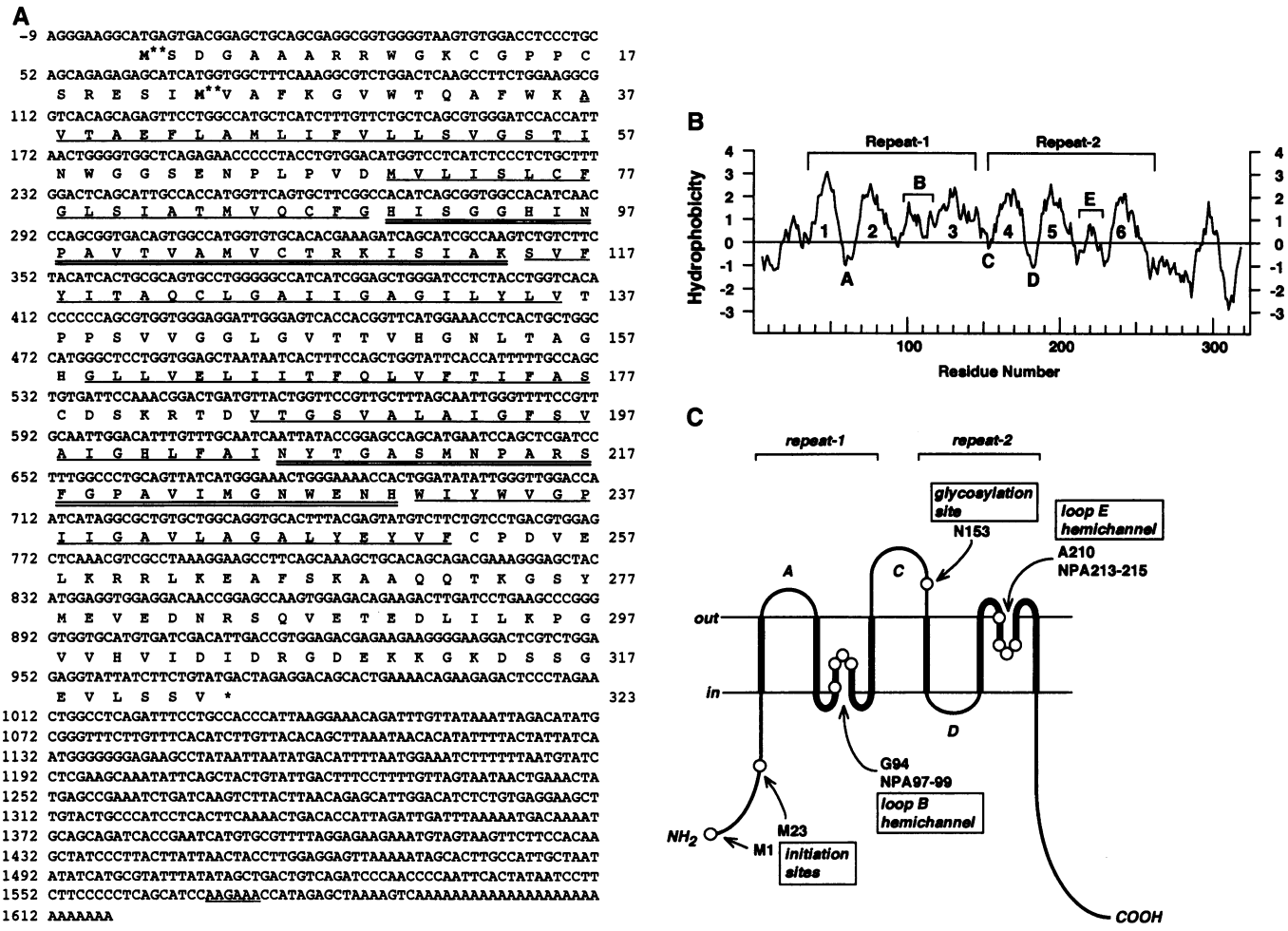

FIG. 1. Sequence analysis and predicted topology of *AQP4* cDNA. (*A*) Nucleotide sequence and deduced amino acid sequence of the isolated clone and determined by 5′-RACE by PCR. Identified are two potential translation initiation sites (double asterisks), presumed membrane bilayer-spanning domains (underlined), and presumed aqueous hemichannel domains (doubly underlined). (*B*) Hydropathy analysis of the deduced amino acid sequence, using a 7-aa window (18). (*C*) Proposed membrane topology based upon the hourglass model for aquaporins (19, 20), comprising six presumed bilayer-spanning domains and five connecting loops (A–E). Domains with >50% of residues identical to AQP1 are heavy lines (six bilayer-spanning domains; loops B and E); domains and loops with <25% identity are thin lines (amino and carboxyl termini; loops A, C, and D).

which contains the *Hin*dIII–*Pst* I insert of pSP64T (3) in pBluescript II KS (Stratagene). The Muta-Gene phagemid *in vitro* mutagenesis kit (Bio-Rad) was used with primers A210C (5′-CAATTATACCGGATGCAGCATGAATCCA-3′) and H201P (5′-GTTGCAATTGGACCTTTGTTTGCAATC-3′), and capped RNA transcripts were synthesized (16).

Expression and Functional Studies. Protein synthesis was performed in a cell-free translation system (Promega) with [^{35}S]methionine and canine pancreatic microsomes. Defolliculated stage V–VI oocytes from *Xenopus laevis* were injected with 50 nl of water or up to 5 ng of sample cRNAs, and a swelling assay was performed (3, 16).

RNA Analyses. RNA from various tissues was size-separated in formaldehyde/1.2% agarose gels and transferred to nylon membranes; these Northern blots were hybridized at high stringency with a ^{32}P-labeled probe (5×10^5 cpm/ml) corresponding to the coding region of the cDNA. RNase protection assays were performed with a ^{32}P-labeled PCR-amplified fragment, corresponding to bp +278 to +626, hybridized overnight at 45°C with 10 μg of total tissue RNA and incubated for 45 min at 30°C in digestion buffer containing RNase A (20 μg/ml) and RNase T_1 (2 μg/ml). The 348-bp protected fragments in denaturing 6% polyacrylamide gels were visualized by autoradiography.

***In Situ* Hybridization of Brain.** Cryosections were cut 12 μm thick from rat brains, fixed with 4% paraformaldehyde in phosphate-buffered saline, and treated with 0.25% acetic anhydride/0.1 M triethanolamine for 10 min. [α-[^{35}S]thio]-UTP-labeled antisense and sense RNA probes were made with T7 or T3 RNA polymerase from a linearized plasmid containing the *AQP4* cDNA. Sections were hybridized overnight at 56°C with probe (10^6 cpm) (17). After RNase treatment, the sections were exposed to autoradiographic film for 1 day and then to photographic emulsion for 4 days.

RESULTS AND DISCUSSION

Brain cDNA. A product obtained by degenerate PCR was used to isolate plaques from a rat brain cDNA library (see *Methods*). The longest insert, 1.6 kb, encodes a methionine (M23; Fig. 1*A*) which conforms to the initiation consensus (21). This precedes an open reading frame encoding 301 aa and 622 bp of 3′ untranslated sequence ending just after a polyadenylylation consensus sequence. Analysis of the GenBank database revealed 95.2% identity with the amino acid sequence encoded by EST02306, a partial cDNA isolated from human fetal brain (22), which may be the human homolog of big brain, a *Drosophila* neural development gene (23). Surprisingly, the sequences 5′ to the apparent initiating

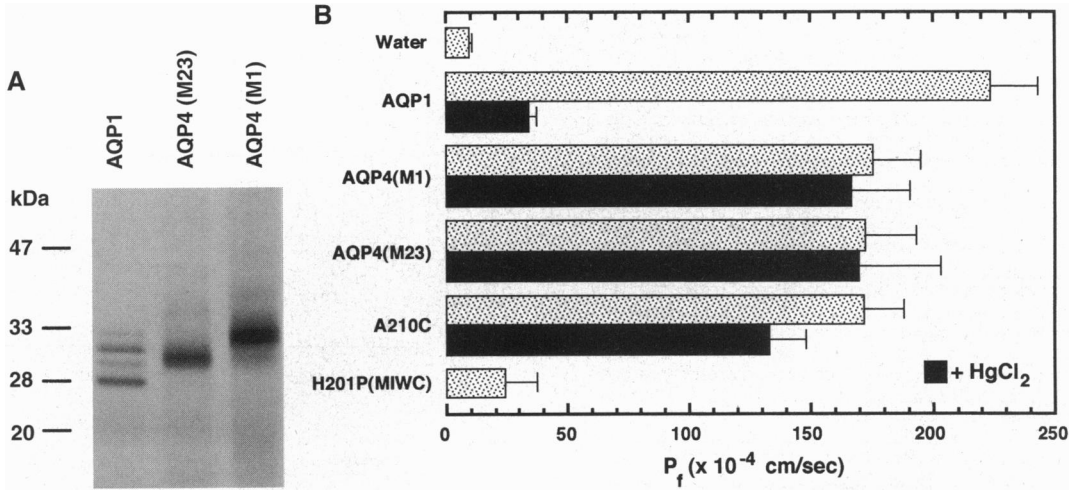

FIG. 2. Expression and function of AQP4. (*A*) Cell-free expression in the presence of microsomes of cRNAs encoding AQP1 or AQP4 with both (M1) or one (M23) translation initiation sites. (*B*) Osmotic water permeability (P_f) of oocytes injected with 50 nl of water or 5 ng of the indicated cRNAs. Bars show mean and SD of three or four determinations for oocytes receiving no further treatment (stippled bars) or oocytes after 5 min incubation in 1 mM $HgCl_2$ (black bars).

methionine of the rat brain cDNA and EST02306 were nearly identical, suggesting that they are both coding sequences (Fig. 1*A*). Therefore, rat brain RNA was used as the template for 5′-RACE, which identified 15 additional nucleotides and an upstream methionine (M1) in frame with the second site (M23), which also conforms to the initiation consensus (21). Existence of additional 5′ coding sequence similar to that of big brain (23) cannot be excluded.

Deduced Structure of the Brain Polypeptide. Overlapping polypeptides of 323 or 301 aa are encoded by the cDNA beginning at the upstream (M1) or downstream (M23) initiation site. Similar to other aquaporins, hydropathy analysis of the deduced amino acid sequence revealed six potential membrane bilayer-spanning domains and five connecting loops (A–E), of which loops B and E are significantly hydrophobic (Fig. 1*B*). Based upon the membrane topology of AQP1 (19, 20), the brain polypeptide was predicted to have a cytoplasmic amino terminus containing both potential initiation sites and a cytoplasmic carboxyl terminus of ≈70 aa (Fig. 1*C*). Loops B and E both contain the NPA amino acid sequence motif which exists in all aquaporins and homologous proteins (1, 2). Cysteines were not present at sites G94 and A210, corresponding to the known mercury inhibitory sites of AQP1 (16, 20), predicting lack of mercurial inhibition of the aqueous pathway. Three potential N-glycosylation sites were identified. The first (N153) is in loop C at a site corresponding to the N-glycosylation sites in renal collecting-duct homologs AQP2 (8) and AQP3 (11). This location is consistent with the extracellular location of loop C as predicted in the six bilayer-spanning topology established for AQP1 (19). Consistent with this prediction, only the first consensus site is occupied in other polytopic integral proteins with multiple potential N-glycosylation sites (24).

Protein synthesis occurs at both initiation sites by cell-free translation of cRNAs. In the presence of microsomes, cRNA containing the downstream site (M23) yielded a single product of 30 kDa, corresponding to the 301-aa polypeptide, and cRNA containing both sites (M1 and M23) yielded a major product of 32 kDa, corresponding to the 323-aa polypeptide, and a minor product of 30 kDa (Fig. 2*A*).

Transport Studies. Functional properties of the brain cRNAs were evaluated by expression in *Xenopus* oocytes. Swelling was monitored after the oocytes were transferred from 200 mOsm to 70 mOsm modified Barth's solution, and coefficients of osmotic water permeability (P_f) were calculated. Similar to AQP1 (3), P_f values of oocytes injected with cRNAs encoding the 323- or 301-aa polypeptide increased by ≈20-fold (Fig. 2*B*). Therefore, polypeptides encoded by the brain-derived cDNAs represent the fourth known mammalian member of the aquaporin family of water channels (1), and the Genome Data Base has designated the gene as *AQP4*.

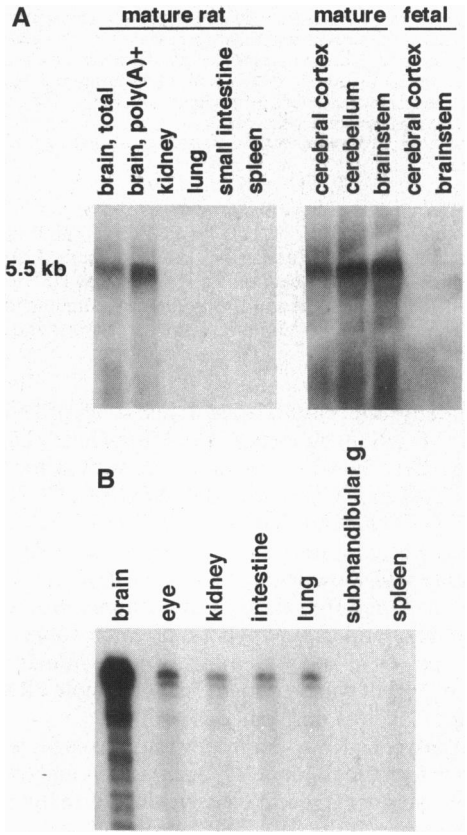

FIG. 3. Analysis of *AQP4* mRNA. (A) Northern blot of brain poly(A)$^+$ RNA (1 μg) or total RNA (20 μg) from various rat tissues hybridized with ^{32}P-labeled probe corresponding to nt +7 to +886 of the *AQP4* cDNA. Equivalent amounts of RNA were loaded as assessed from abundance of 18S and 28S RNAs and Northern analysis using a chicken β-actin probe (data not shown). (*B*) Protection of 10 μg of total RNA isolated from selected tissues incubated with ^{32}P-labeled RNA probe prior to digestion with RNase (see *Methods*). g., Gland.

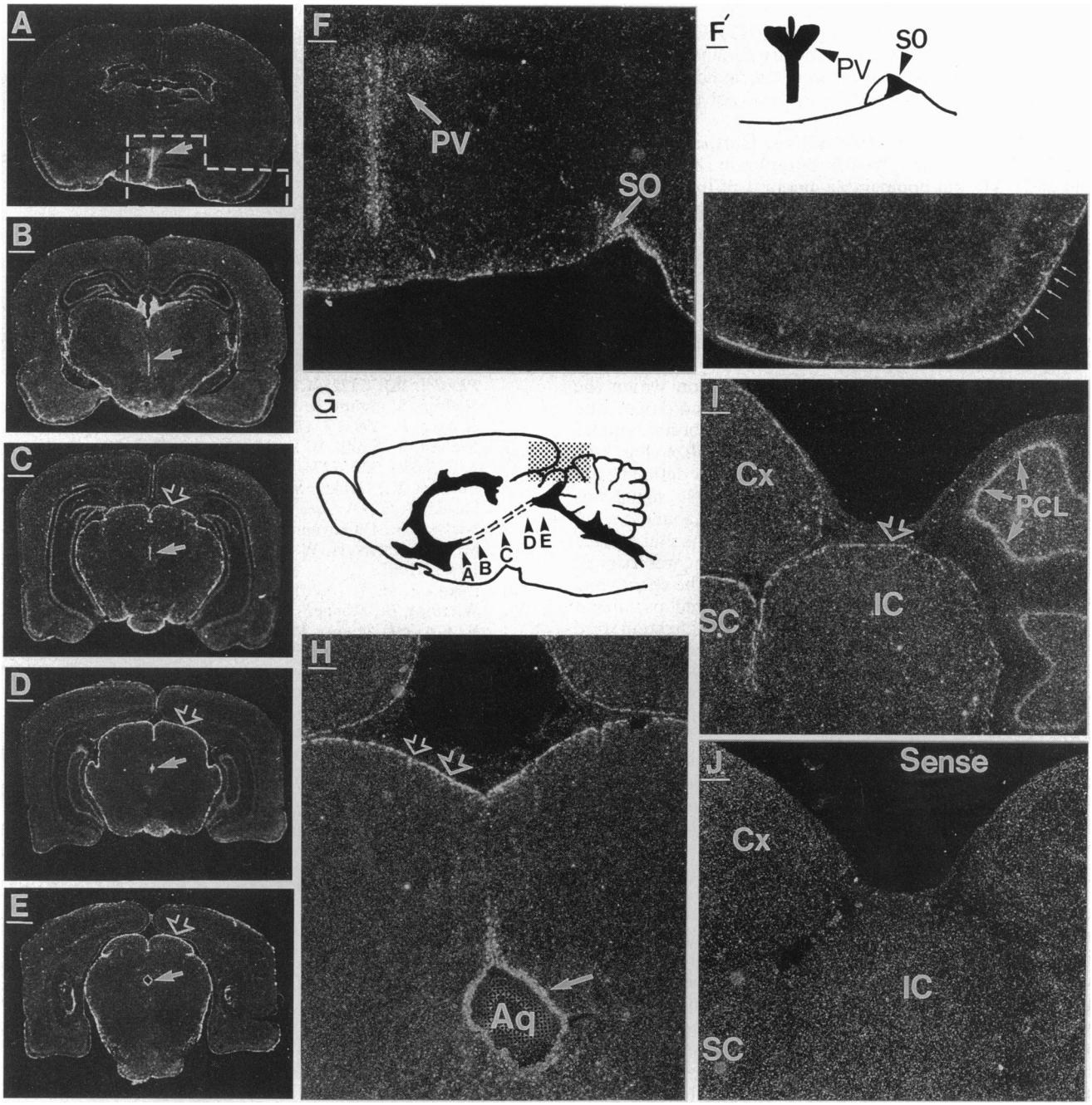

FIG. 4. *In situ* localization of *AQP4* mRNA in rat brain. (*A–E*) Autoradiograms of coronal sections hybridized with antisense probe reveal intense signal on the ependymal lining of the aqueduct system at the midline (white arrows in *B–E*) and on the surface of the brainstem (open arrowheads in *C, D, E, H,* and *I*). (*F* and *F'*) High-magnification darkfield image and schematic diagram of emulsion-dipped coronal section corresponding to box in *A; F* shows hybridization signal over the paraventricular (PV) and supraoptic (SO) nuclei located near the base of the brain and over a layer of glial cells at the edge of the brain parenchyma (multiple small arrows). (*G*) Schematic diagram of a sagittal brain section representing the aqueduct system and levels of the coronal sections (black arrowheads). (*H*) High magnification of aqueduct (Aq) in a coronal section corresponding to *E*. (*I*) High magnification of sagittal section corresponding to stippled area in *G*, representing the junction of cerebellum, superior colliculus (SC), inferior colliculus (IC), and cerebral cortex (Cx). Intense hybridization is seen over the Purkinje cell layer (PCL) of cerebellum. (*J*) Hybridization of the section adjacent to *I* with the sense probe shows only negligible signal.

The AQP1-mediated increase in P_f is blocked by treatment with 1 mM $HgCl_2$ (3). Similar treatment of oocytes expressing *AQP4* failed to significantly inhibit P_f (Fig. 2*B*). C189 in loop E is the mercury-sensitive site in AQP1 (16). Substitution of cysteine for the alanine at the corresponding position of AQP4 (aa 210) did not affect the baseline P_f, nor did it confer mercury sensitivity, suggesting that the structures of the channel apertures in AQP4 and AQP1 are not identical (16, 20). As with *AQP1*, oocytes expressing *AQP4* exhibited no increase in membrane transport of [^{14}C]urea or [^{14}C]glycerol over that of water-injected oocytes, and electrophysiological studies failed to reveal increased membrane conductance (data not shown).

While these studies were nearing completion, a report appeared describing an aquaporin cDNA from rat lung referred to as WCH4 or MIWC (mercury-insensitive water channel; ref. 14). The coding sequence of MIWC corresponds to that of the 301-aa AQP4 polypeptide with the exception of aa 201, which is a proline in MIWC but is a histidine in AQP4 (Fig. 1*A*), as well as AQP1, AQP2, and MIP. To determine whether this discrepancy was due to a natural polymorphism,

the cDNA of the 301-aa AQP4 was mutagenized at this site to correspond to the reported sequence (H201P). Oocytes injected with the H201P cRNA exhibited a P_f only slightly above that of water-injected control oocytes, indicating that the sequence reported for MIWC (14) does not encode a fully active water channel.

Tissue Distribution of *AQP4* mRNA. Northern analysis of total RNA or poly(A)$^+$ RNA from rat brain showed a single band of 5.5 kb, but comparable amounts of total RNA from mature rat kidney, lung, small intestine, and spleen failed to react with the *AQP4* probe (Fig. 3*A*). RNase protection performed with a probe corresponding to nt +278 to +626 provided a more sensitive and quantitative method to detect RNA in tissues (Fig. 3B). Protection by brain RNA was >>10-fold above that by RNAs from eye, kidney, intestine, or lung; no protection signal was even detected with RNA from spleen or salivary gland.

AQP4 mRNA has a widespread distribution within the brain. Cerebral cortex, cerebellum, and spinal cord of mature rats exhibited comparable signals, whereas brain samples obtained from day 17 fetal rats failed to hybridize (Fig. 3*A*). The localization in mature rat brain was further defined by *in situ* hybridization with *AQP4* RNA probes. No signal was identified in the choroid plexus, where AQP1 is abundant, but high levels of *AQP4* mRNA were located at the pial surface of the brain, where the brain is in contact with cerebrospinal fluid in the subarachnoid space (Fig. 4 *A*–*E*). The ependymal lining of the aqueductal system also displayed prominent staining (Fig. 4 *B*–*E* and *H*). High levels of hybridization were also detected in the paraventricular and supraoptic nuclei, and specific hybridization was apparent in several neuronal cell layers, including the granule cell layer of the dentate gyrus and the Purkinje cell layer in the cerebellum (Fig. 4 *F* and *I*).

Taken together, these studies indicate that AQP4 mediates transmembrane water movement in adult brain. Unlike the homologous gene big brain, which is essential for embryonic neuronal development in *Drosophila* (23), *AQP4* is not expressed during fetal brain development in rat (Fig. 3*A*). In mature rat, *AQP4* is expressed prominently in the ependymal cells lining the aqueductal system and over the surface of the brain in contact with the subarachnoid space. As cerebrospinal fluid fills these structures, AQP4 may facilitate water balance between brain parenchyma and the fluid compartment. *AQP4* is also expressed abundantly in neurons of the paraventricular and supraoptic nuclei which synthesize vasopressin and project axons into the neurohypophysis. Patch-clamp studies have demonstrated that these magnocellular neurosecretory cells possess mechanosensitive ion channels that rapidly respond to swelling or shrinking (25). AQP4 in the plasma membranes of these neurons could mediate rapid changes in cell volume in response to local shifts in extracellular osmolality. Lower levels of AQP4 or other, as yet unidentified aquaporins may contribute to osmotic equilibrium throughout the central nervous system. Further investigation of AQP4 is warranted, as abnormalities in brain osmotic and volume homeostasis can have severe consequences in a wide variety of pathological conditions such as hyperglycemia, uremia, and postischemic swelling.

We thank Hyug Moo Kwon, Peter Aronson, and Tiziana P. Carroll for their help. This work was supported in part by grants from the National Institutes of Health.

1. Agre, P., Preston, G. M., Smith, B. L., Jung, J. S., Raina, S., Moon, C., Guggino, W. B. & Nielsen, S. (1993) *Am. J. Physiol.* **265,** F463–F476.
2. Knepper, M. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 6255–6258.
3. Preston, G. M., Carroll, T. P., Guggino, W. B. & Agre, P. (1992) *Science* **256,** 385–387.
4. Denker, B. M., Smith, B. L., Kuhajda, F. P. & Agre, P. (1988) *J. Biol. Chem.* **263,** 15634–15642.
5. Sabolic, I., Valenti, G., Verbavatz, J. M., van Hoek, A. N., Verkman, A. S., Ausiello, D. A. & Brown, D. (1992) *Am. J. Physiol.* **263,** C1225–C1233.
6. Nielsen, S., Smith, B. L., Christensen, E. I., Knepper, M. A. & Agre, P. (1993) *J. Cell Biol.* **120,** 371–383.
7. Nielsen, S., Smith, B. L., Christensen, E. I. & Agre, P. (1993) *J. Cell Biol.* **120,** 7275–7279.
8. Fushimi, K., Uchida, S., Hara, Y., Hirata, Y., Marumo, F. & Sasaki, S. (1993) *Nature (London)* **361,** 549–552.
9. Nielsen, S., DiGiovanni, S. R., Christensen, E. I., Knepper, M. A. & Harris, H. W. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 11663–11667.
10. Deen, P. M. T., Verdijk, M. A. J., Knoers, N. V. A. M., Wieringa, B., Monnens, L. A. H., van Os, C. H. & van Oost, B. A. (1994) *Science* **264,** 92–95.
11. Ishibashi, K., Sasaki, S., Fushimi, K., Uchida, S., Kuwahara, M., Saito, H., Furukawa, T., Nakajima, K., Yamaguchi, Y., Gojobori, T. & Marumo, F. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 6269–6273.
12. Strange, K. (1992) *J. Am. Soc. Nephrol.* **3,** 13–27.
13. Dunn, F. L., Brennan, T. J., Nelson, A. E. & Robertson, G. L. (1973) *J. Clin. Invest.* **52,** 3212–3219.
14. Hasegawa, H., Ma, T., Skach, W., Matthay, M. A. & Verkman, A. S. (1994) *J. Biol. Chem.* **269,** 5497–5500.
15. Preston, G. M. (1993) in *Methods in Molecular Biology: PCR Protocols: Current Methods and Applications,* ed. White, B. A. (Humana, Totowa, NJ), Vol. 15, pp 317–337.
16. Preston, G. M., Jung, J. S., Guggino, W. B. & Agre, P. (1993) *J. Biol. Chem.* **268,** 17–20.
17. Bhat, R. V., Baraban, J. M., Johnson, R. C., Eipper, B. A. & Mains, R. E. (1994) *J. Neurosci.* **14,** 3059–3071.
18. Kyte, J. & Doolittle, R. F. (1982) *J. Mol. Biol.* **105,** 105–132.
19. Preston, G. M., Jung, J. S., Guggino, W. B. & Agre, P. (1994) *J. Biol. Chem.* **269,** 1668–1673.
20. Jung, J. S., Preston, G. M., Smith, B. L., Guggino, W. B. & Agre, P. (1994) *J. Biol. Chem.* **269,** 14648–14654.
21. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8132.
22. Adams, M., Dubnick, M., Kerlavage, A. R., Moreno, R., Kelley, J. M., Utterback, T. R., Nagle, J. W., Fields, C. & Venter, J. C. (1992) *Nature (London)* **365,** 632–634.
23. Rao, Y., Jan, L. Y. & Jan, Y. N. (1990) *Nature (London)* **345,** 163–167.
24. Landolt-Marticorena, C. & Reithmeier, R. A. F. (1994) *Biochem. J.* **302,** 253–260.
25. Oliet, S. H. R. & Bourque, C. W. (1993) *Nature (London)* **364,** 341–343.