Available online at www.sciencedirect.com

ScienceDirect

Genomics 91 (2008) 367–377

GENOMICS

www.elsevier.com/locate/ygeno

New isoforms of rat Aquaporin-4

Svein Erik Moe [a,b,1], Jan Gunnar Sorbo [a,b,1], Rikke Sogaard [b,c], Thomas Zeuthen [b,c], Ole Petter Ottersen [a,b], Torgeir Holen [a,b,*]

[a] *Center for Molecular Biology and Neuroscience, University of Oslo, Oslo, Norway*
[b] *Nordic Center of Excellence for Research in Water Imbalance Related Disorders, University of Oslo, Oslo, Norway*
[c] *Panum Institute, University of Copenhagen, Copenhagen, Denmark*

Received 11 September 2007; accepted 5 December 2007
Available online 5 February 2008

Abstract

Aquaporin-4 (AQP4) is a brain aquaporin implicated in the pathophysiology of numerous clinical conditions including brain edema. Here we show that rat AQP4 has six cDNA isoforms, formed by alternative splicing. These are named AQP4a–f, where AQP4a and AQP4c correspond to the two classical M1 and M23 isoforms, respectively. The various isoforms are differentially expressed in kidney and brain, and their prevalence does not correspond to the level of the respective mRNAs, pointing to posttranscriptional regulation. The three isoforms lacking exon 2, AQP4b, AQP4d, and AQP4f, have an intracellular localization when expressed in cell lines and do not transport water when expressed in *Xenopus* oocytes. In contrast, the largest of the new isoforms, AQP4e, which contains a novel N-terminal domain, is localized at the plasma membrane in cell lines and functions as a water transporter in *Xenopus* oocytes.

Keywords: AQP4; Aquaporin; Alternative splicing; Isoforms

Aquaporin-4 (AQP4), originally isolated from rat fetal lung and rat brain [1,2], is the aquaporin most strongly expressed in brain. It is localized mainly in astrocytic endfeet around blood vessels and bordering the subarachnoidal space, in ependymal cells lining the ventricles, and in glial lamellae in the supraoptic nucleus and other circumventricular organs [3]. This pattern of expression suggests that AQP4 is involved in transport of water at the brain/blood and brain/cerebrospinal fluid interfaces and indicates a role in osmotransduction [3,4]. In agreement, studies of AQP4-/- mice and mice with a disruption of AQP4 anchoring (mdx and alpha-syntrophin-/- mice) have implicated AQP4 in a long list of physiological processes [5,6]. The most striking effect of AQP4 knock-out is reduced brain edema after acute water intoxication and ischemic stroke [7].

AQP4 has two known isoforms, M1 and M23 [2], which differ in their ability to form square arrays in cell culture [8] and in their ability to form 2D-crystals when reconstituted in lipid bilayers [9]. The water transport capacity of the two isoforms has been found to be similar in *Xenopus* oocytes [10] but in epithelial cells the M23 variant has been reported to have increased single-channel osmotic water permeability by one order of magnitude [11].

The aim of this study was to provide a thorough mapping of the rat AQP4 gene. By rapid amplification of cDNA ends (RACE) analysis we find six cDNA isoforms: here named AQP4a, AQP4b, AQP4c, AQP4d, AQP4e, and AQP4f. AQP4a and AQP4c correspond to the two classic isoforms, M1 and M23. We find no evidence for M23 having a transcription start in intron 1. By designing TaqMan assays against the new exon–exon boundaries, we assessed the abundance of the new isoforms in different tissues and demonstrated a mismatch on the mRNA and protein levels, a mismatch that indicates significant posttranscriptional regulation of AQP4 gene expression.

To investigate protein localization, we expressed the new isoforms in cell lines and found that three of the new isoforms were intracellular, whereas AQP4e, the largest new isoform, was transported to the cell membrane. In *Xenopus* oocytes we demonstrate that AQP4e transports water, whereas the intracellular

* Corresponding author.
E-mail address: torgeir.holen@medisin.uio.no (T. Holen).
[1] These authors contributed equally to this study.

0888-7543/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2007.12.003

isoforms do not. Since the rat has been the main animal model for water transport studies of the brain, the newfound complexity of the rat AQP4 gene must be taken into account in further studies of the physiological and pathophysiological roles of AQP4.

Results

We have recharacterized the rat AQP4 gene and found six cDNA isoforms instead of the two classical isoforms M1 and M23. The increased complexity necessitates a simple nomenclature; thus we have named the isoforms AQP4a, AQP4b, AQP4c, AQP4d, AQP4e, and AQP4f, where AQP4a corresponds to the classical M1 isoform and AQP4c corresponds to M23. Fig. 1A. summarizes our results, with the novel exons, mRNAs, and predicted proteins and protein sizes indicated.

Our new model implies three transcription starts, at the 5' end of exon Z, exon 0, and exon X. Exon Z has a new open reading frame coding for the new, large AQP4e isoform with a novel N-terminal domain (Fig. 1B). The alternative splicing of exon 2, from the basic forms AQP4a, AQP4c, and AQP4e, produces AQP4b, AQP4d, and AQP4f (Fig. 1B), resulting in a total of six cDNA isoforms in the new model.

Our data (see below) are consistent with the initial characterization of AQP4 by Verkman and co-workers [1] and with a later revision [2]. The exception is that we find a T-to-C polymorphism in exon 1 (Fig. 1B, purple C, DNA position 468). This polymorphism, not previously described, occurs in several clones from three different tissues, indicating a genome polymorphism, but does not cause an amino acid substitution in the AQP4 protein.

No PCR evidence for classical M23 transcription start in intron 1

The classical structure of the AQP4 locus involved two transcription starts [1,2,17]. The larger isoform is transcribed from exon 0 (Fig. 2A, open arrow), whereas the shorter M23 isoform is supposed to have a transcription start in intron 1 (Fig. 2A, closed arrow).

This model was supported by an analysis of the 5' end of the M23 mRNA in the rat [18]. Furthermore, a cDNA clone released from Amgen also contained sequences upstream of exon 1 (Accession No. CB614668). In mouse, a similar model was proposed by Zelenin et al. [19], although the confounding effect of genomic DNA contamination could not be excluded.

A careful bioinformatic analysis revealed several inconsistencies in the sequence databases in regard to the 5' end of the rat AQP4 mRNA (data not shown, and Supplementary Fig. 1). To test the model, we synthesized a series of PCR primers located just upstream of exon 1 (Fig. 2B, red primers), but no sign of exon 1–extended transcripts were found by PCR (Fig. 2C). We synthesized a second series of primers (Fig. 2B, light blue primers) and still found no sign of upstream extension of exon 1 (Fig. 2C, lanes 7–12).

On the other hand, successful amplification of the M1 transcript was easily obtained by designing a single primer designed against exon 0 (Fig. 2C, lane 13). Also a control primer in exon 1 was successfully designed and tested (Fig. 2C lanes 5 and 6). Thus our PCR experiments did not support the hypothesis that M23 is coded from an extension of exon 1.

RACE characterization of the 5'end of rat AQP4 mRNA finds new exons

In their revision of the rat AQP4 characterization, Agre and co-workers [2] performed a limited RACE screen, finding exon 0 and noting that there could be additional exons. To map *de novo* the 5' end of the rat AQP4 mRNA, we used a PCR – based 5' RACE technique combined with Topo-cloning. We aimed to isolate multiple transcript clones from multiple tissues, to test the classical model of the AQP4 locus, to identify possible new exons, and to obtain a semiquantitative measure of the ratio of M1 and M23 transcripts in the rat.

A pair of RACE primers was designed in exon 2. Of 38 unique AQP4 clones isolated from kidney, hippocampus, and cerebellum RNA, none demonstrated an extension of exon 1 (Supplementary Fig. 2; cDNA sequences submitted to GenBank). This again failed to support the idea that the classical M23 isoform is encoded by a transcription start in intron 1.

However, we did find evidence of two new exons. The RACE transcripts fell into three distinct groups (Supplementary Fig. 2). One group of transcripts aligned perfectly with exon 2 and exon 1 and then skipped to a region of the genome between the previously mapped exon 1 and exon 0. We have named this new region exon X, because this exon has sequence similarity to exon X found by Zelenin and co-workers [19] in mouse. We consider exon X a strong candidate for coding the 5' UTR of the AQP4c mRNA isoform (corresponding to the classical isoform M23).

The majority of transcripts aligned with exon 2 and exon 1 and then skipped to the classical exon 0 (Supplementary Fig. 2, middle group). Here the transcripts aligned perfectly with the reference sequence of the M1 form of AQP4 (NM_012825). However, exon 0 was extended in almost all transcripts another 19 to 20 nucleotides. In this extension, 9 nucleotides were known from a previous study. [2]. The revised exon 0 is indicated in blue in Fig. 1B. In total, 21 independent clones of ~600 bp sequence came to an end within 1 or 2 nucleotides, thus constituting a probable transcription start point of AQP4a (the classical isoform M1).

The preponderance of exon 0 clones over exon X clones should imply a stronger expression of the larger, classical M1 protein isoform. However, on immunoblots, the shorter classical M23 isoform is the predominant isoform [10]. This discrepancy may indicate significant regulation at the level of translation. Another possible explanation is that the M1 and the M23 protein isoforms differ in regard to their stability.

A third group of RACE transcripts progressed even further upstream from the extended exon 0 (Supplementary Fig. 2, lower group). The longest clone isolated, R90_185 (GenBank Accession No EV243422), stretched 185 bp upstream of the revised exon 0 and contained a new start codon (Fig. 1B). This start codon was in reading frame with the rest of AQP4, making this a candidate for encoding the faint, minor AQP4 protein isoforms seen on immunoblots [10]. We named the longer exon 0-extension exon Z and named the new putative protein isoform

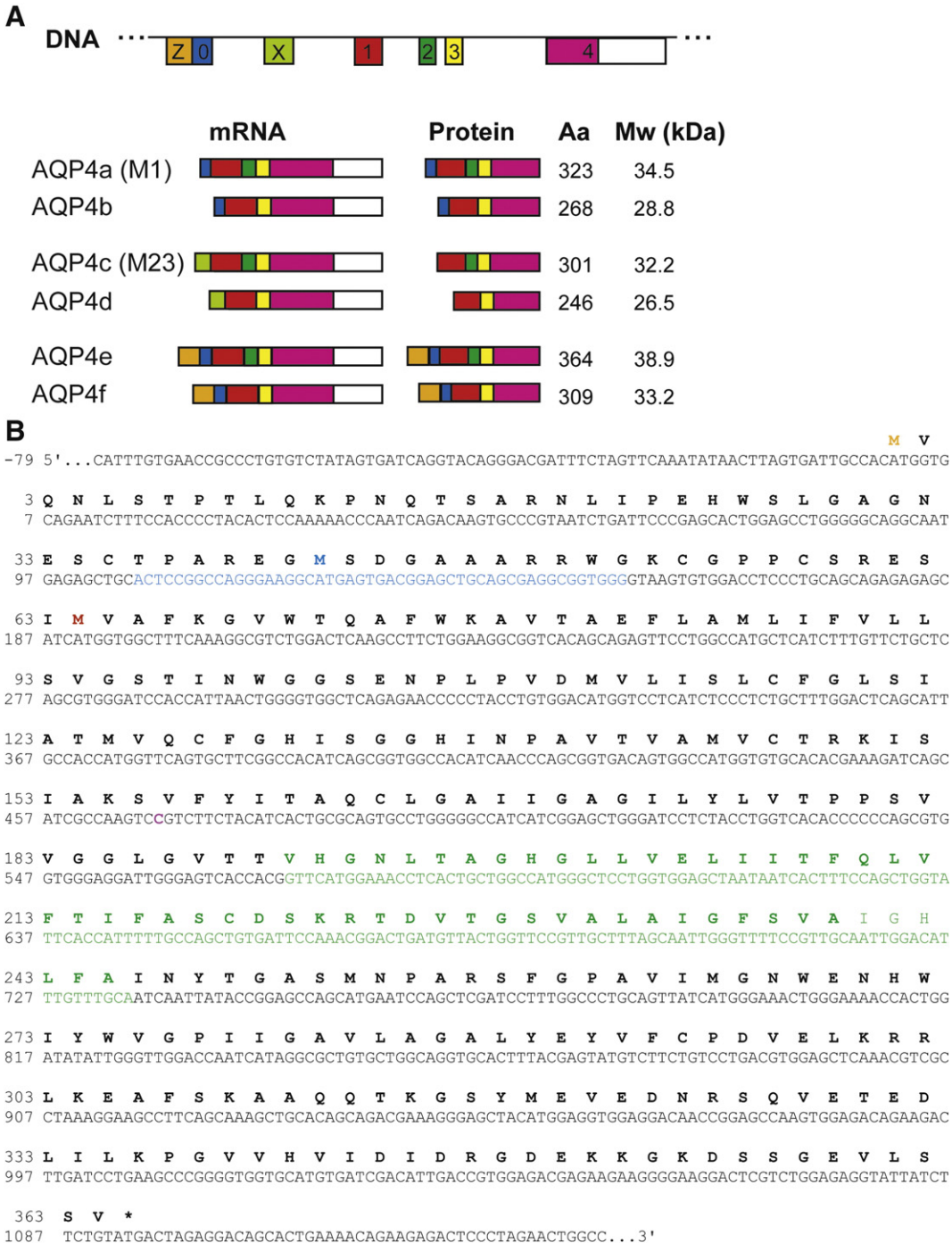


Fig. 1. (A) Model of six isoforms of rat AQP4 isoforms. A new start codon was found in the new exon Z, giving rise to AQP4e, the largest cDNA and protein. AQP4b, AQP4d, and AQP4f are permutations of AQP4a, AQP4c, and AQP4e, respectively, lacking exon 2. (B) DNA and protein sequence of the longest clone isolated (R90_185), showing three start codons (for AQP4e (orange M), for AQP4a (blue M), and for AQP4c (red M)). The alternatively spliced exon 2 is shown in green, and the revised exon 0 is shown in blue. DNA base 468 (purple) deviates from reference AQP4 sequence (Accession No. NM_012825).

AQP4e (Fig. 1A). No homologues of AQP4e were found by BLAST search of databases.

Validation of exon Z and exon X by PCR

To validate by PCR the existence of the AQP4e transcript and to clone the full-length transcript, we designed new primers upstream of the start codon and downstream of the stop codon.

Strong PCR bands validated the existence of exon Z with primers in exon Z and exon 2, in kidney cDNA (Fig. 3A, lane 1), and in cortex cDNA (Fig. 3A, lane 5). Exon X was also validated (Fig. 3A, lane 8). A second series of PCR primers in exon Z and exon X gave identical results (data not shown).

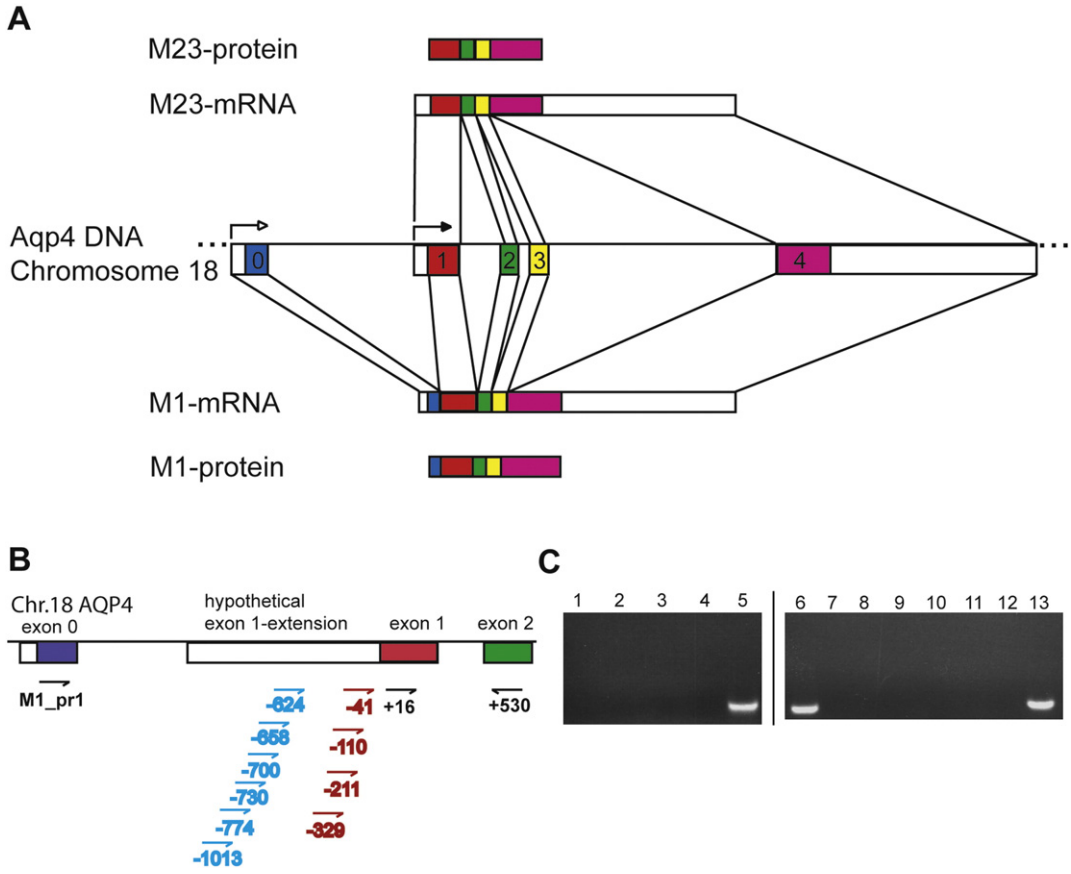


Fig. 2. (A) Classical genomic structure of AQP4 on DNA, mRNA, and protein levels, showing two isoforms M1 and M23, with putative transcription starts shown as open arrow and black arrow, respectively. (B) Diagram over PCR primers for validation of exon 0 and investigation of M23-exon 1-extension. (C) PCR investigation of hypothetical exon 1-extension on hippocampus cDNA. Lanes 1–4 are first series (primers in red, B, with primer +530). Lanes 5 and 6 are positive control with primers +16 and +530. Lanes 7–12 are with second series (primers in blue, B, with primer +530). Lane 13 is positive control and validation of exon 0 (primers M1_pr1 and + 530).

Cloning of AQP4e by PCR shows alternative splicing of exon 2

Interestingly, when amplifying full-length AQP4e from cortex cDNA, a secondary band appeared below the major band (Fig. 3A, lanes 2 and 3, open arrowhead). We isolated this band because we suspected that this could be an alternative splicing variant of the AQP4e isoform. TOPO-cloning and sequencing confirmed that this was AQP4e but with exon 2 missing. We named this cDNA isoform AQP4f (Fig. 1A).

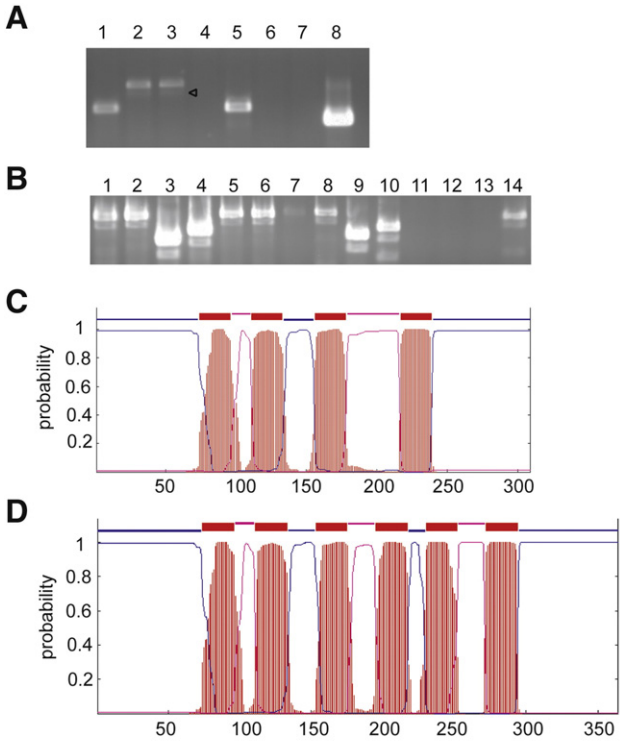
To validate the exon 2-less transcripts we used PCR with sets of primers in exon 0, exon Z, and exon X, combined with primers in exon 3 and exon 4. In all cases, we obtained clear double bands, indicating alternative splicing, with all primer pairs (Fig. 3B, lanes 1–10, lane 14, lanes 11–13 negative controls). Isolation, cloning, and sequencing of these cDNA bands confirmed their identity as exon 2-less AQP4 transcripts (data not shown). The faint third bands seen in lanes 10 and 14 were also cloned, sequenced, and found to be PCR artifacts.

The existence of exon 2-less transcripts, by permutation, doubled the possible number of AQP4 protein isoforms from three to six, named AQP4a, AQP4b, AQP4c, AQP4d, AQP4e, and AQP4f (Fig. 1A). A similar exon 2-less clone has been demonstrated, but not submitted to the sequence databases, in a previous paper [1]. The existence of AQP4b is so far theoretical, with no direct evidence obtained, or obtainable, by PCR, because the double band could arise from the AQP4f transcript.

Lack of exon 2 results in AQP4 isoforms with four transmembrane helices

Protein hydrophobicity plots of the six transcripts demonstrated that loss of exon 2 would leave the general transmembrane helix structure intact and with no frame shift but, interestingly, with only four transmembrane helices (Fig. 3C). All other aquaporins, including the novel isoform AQP4e, have six transmembrane helices predicted by the TMHMM server (Fig. 3D and Supplementary Fig. 3). In light of the known properties of the aquaporin fold from structural studies [9,20], it would be unlikely that the four-helix isoforms are able to transport water.

The loss of exon 2 removes helix 4, loop D, and helix 5 (Figs. 3C, and D) and leaves the two crucial NPA motifs intact (Fig. 1B, pos. 138–140 and pos. 254–256). We note that the AQP4 pore structure recently published indicates that loop D, connecting helix 4 and helix 5, stabilizes the tetramer structure [9]. AQP4d expressed in HeLa cells is unable to form tetramers, as demonstrated by blue-native PAGE gels, unlike six-helix isoforms AQP4a, AQP4c, and AQP4e, which do form tetramers (Sorbo et al., unpublished data).



inside	1 - 73	N-domain	TMhelix	195 - 217	Helix 4
TMhelix	74 - 96	Helix 1	inside	218 - 229	Loop D
outside	97 - 110	Loop A	TMhelix	230 - 252	Helix 5
TMhelix	111 - 133	Helix 2	outside	253 - 271	Loop E
inside	134 - 152	Loop B	TMhelix	272 - 294	Helix 6
TMhelix	153 - 175	Helix 3	inside	295 - 364	C-domain
outside	176 - 194	Loop C			

Fig. 3. (A) PCR cloning of AQP4e in cortex cDNA. Lane 2 used primers zAQP4_10 and zAQP4_1181. Lane 3 used zAQP4_10 and zAQP4_1201. Arrowhead show faint lower band. Lanes 1 and 5 validate exon Z in kidney and brain, respectively, using primers zAQP4_10 and +530. Lanes 4, 6 and 7 are negative controls. Lane 8 validates exon X in brain (primers yAQP4_41 and +530). (B) PCR validation of exon 2-less transcripts on rat cortex cDNA in exon 0, exon Z, and exon X. Lanes 1–4 are primer in exon 1 (+16) combined with zAQP4_1181, zAQP4_1201, M23_637, and M23_757 (in exon 4, 4, 3, and 4, respectively). Lanes 5 and 6 are M1_pr1 in exon 0 with zAQP4_1181 and zAQP4_1201 in exon 4. Lanes 7–10 are yAQP4_41 in exon X with zAQP4_1181, zAQP4_1201, M23_637, and M23_757. Lanes 11–13 negative controls. Lane 14 yAQP4_75+ zAQP4_1201. (C) Transmembrane helix prediction of AQP4f. Shown are amino acids 1–309, with predicted transmembrane regions (red boxes and probability graph) (by TMHMM server at http://www.cbs.dtu.dk/services/TMHMM-2.0/). (D) Transmembrane helix prediction of AQP4e (top) with helices, loops, and N- and C-terminal domains limited to amino acid positions (table below).

Quantitative TaqMan mRNA isoform expression analysis of mRNA isoforms in different tissues

The importance of the new AQP4 cDNA isoforms might be indicated by their expression levels in various tissues. Due to the large size of the AQP4 mRNA, of over 5.5 kb [1,2], Northern blots could not resolve AQP4 mRNA isoforms directly (data not shown). To obtain quantitative data on the expression level of the new cDNA isoforms, TaqMan probes were designed against the exon 1–exon 3 boundary (Fig. 4A, T_4) because the standard AQP4 TaqMan probe from Applied Biosystems (Fig. 4A, T_1), is based on the exon 1–exon 2 boundary and therefore does not detect the exon 2–less transcripts. Further, we designed one

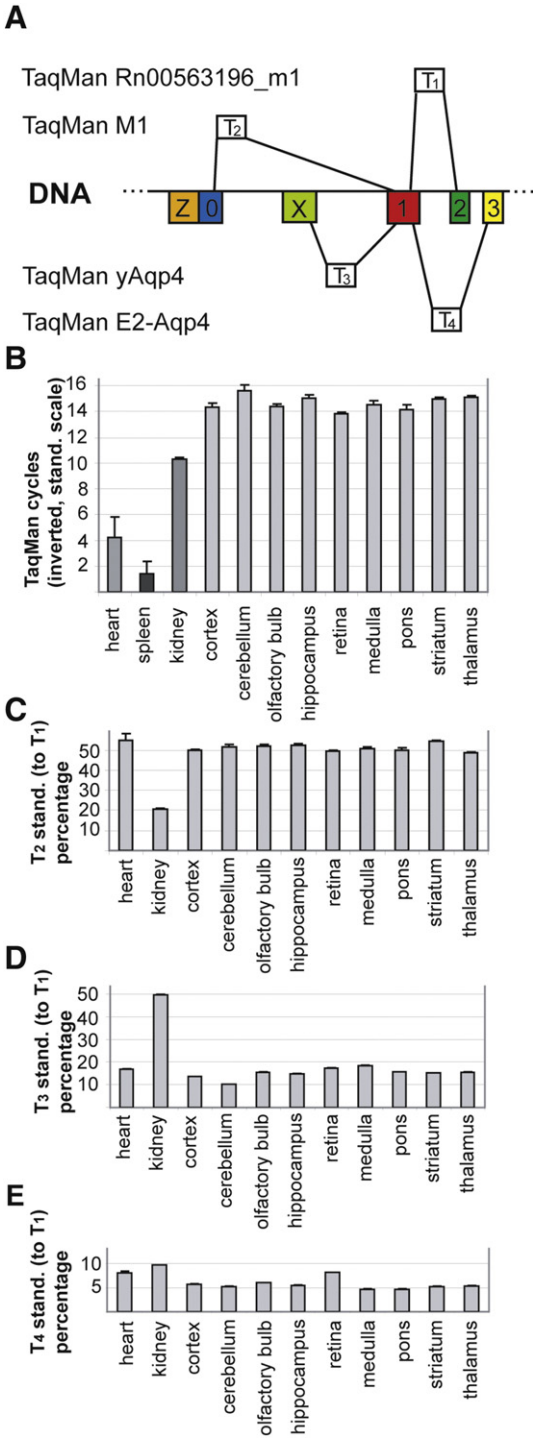
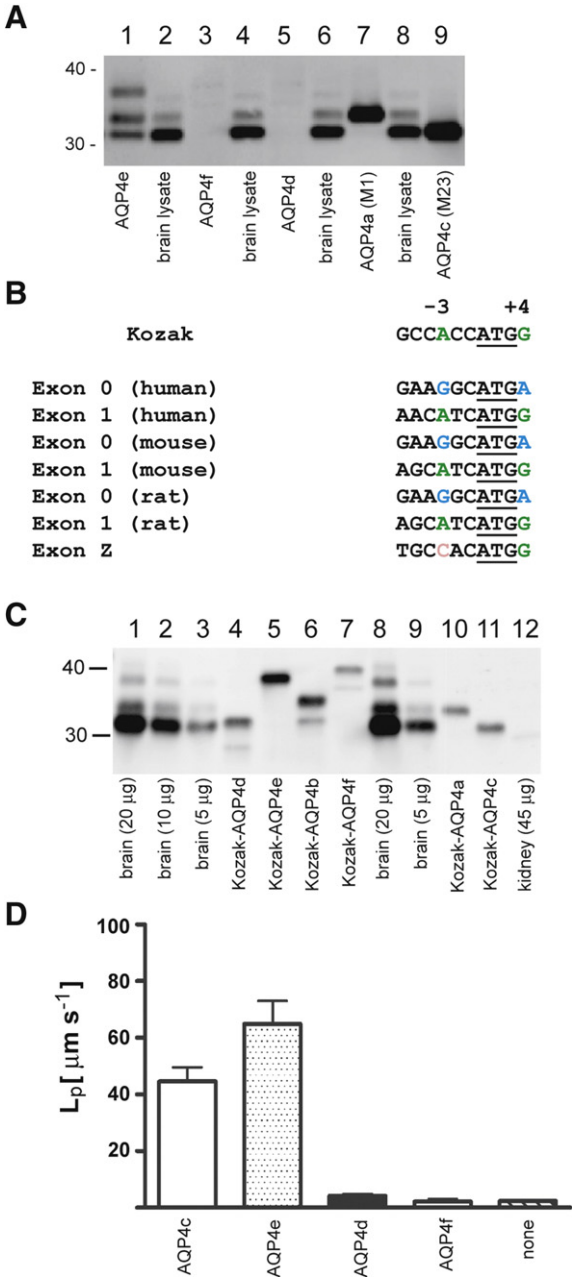


Fig. 4. Taqman real-time PCR analysis in various tissues from female Wistar rats. (A) Diagram of four TaqMan probes against different exon–exon junctions in AQP4 mRNA. (B) TaqMan results (absolute level, standardized to RNA load) using TaqMan probe Rn00563196_m1 (T_1) (inverted, standardized scale). (C) Relative expression of exon 0–exon 1 junction mRNAs (probe T_2) relative to the T_1 probe. (D) Relative expression of exon X–exon1 junction mRNAs (probe T_3), relative to the T_1 probe. (E) Relative expression of exon1–exon3 junction mRNAs (probe T_4) relative to the T_1 probe. All data from at least three independent experiments, with SD.

TaqMan probe against the exon 0–exon 1 boundary (Fig. 4A, T_2) and another against the exon X–exon 1 boundary (Fig. 4A, T_3).

The transcript signal obtained with the standard AQP4 TaqMan probe (Rn_00563196_m1); T_1, Fig. 4A) was ~16-fold stronger in the brain than in the kidney (four cycles, e.g., $2^4=16$; Fig. 4B). Within the brain, the expression level did not vary much, with the highest expression level found in cerebellum (Fig. 4B), consistent with a previous report [21]. Interestingly, the TaqMan assay revealed a faint expression of AQP4 in heart, while the lowest expression was found in the spleen (Fig. 4B).

The transcript signals were standardized to the standard AQP4 TaqMan probe (Rn_00563196_m1; T_1, Fig. 4A). To validate the quantification, TaqMan probes were calibrated and compared against the various AQP4 isoforms cloned into pcDNA3 plasmids (data not shown).



The T_2 probe showed that the mRNA with the exon 0–exon 1 boundary were more highly expressed than the MRNA with the exon X–exon 1 boundary (probe T_3) in most tissues (Figs. 4C and D). Again, this contrasts with the ~10:1 ratio between M23 and M1 bands at the protein level [10]. The main exception was in kidney, where the exon X transcripts were about twice as abundant as those in the brain, and the exon 0/exon Z transcripts were expressed at only one-third of the level of the brain (Figs. 4C, and 4D).

The transcripts detected by the E2-AQP4 (T_4) probe showed that exon 2 - less transcripts (whether AQP4b, AQP4d, or AQP4f) are expressed at approximately 5% of the standard (T_1) transcript in most tissues (Fig. 4E). The interesting exceptions are heart, kidney, and retina, where the T_4 probe signal is 10% of the standard.

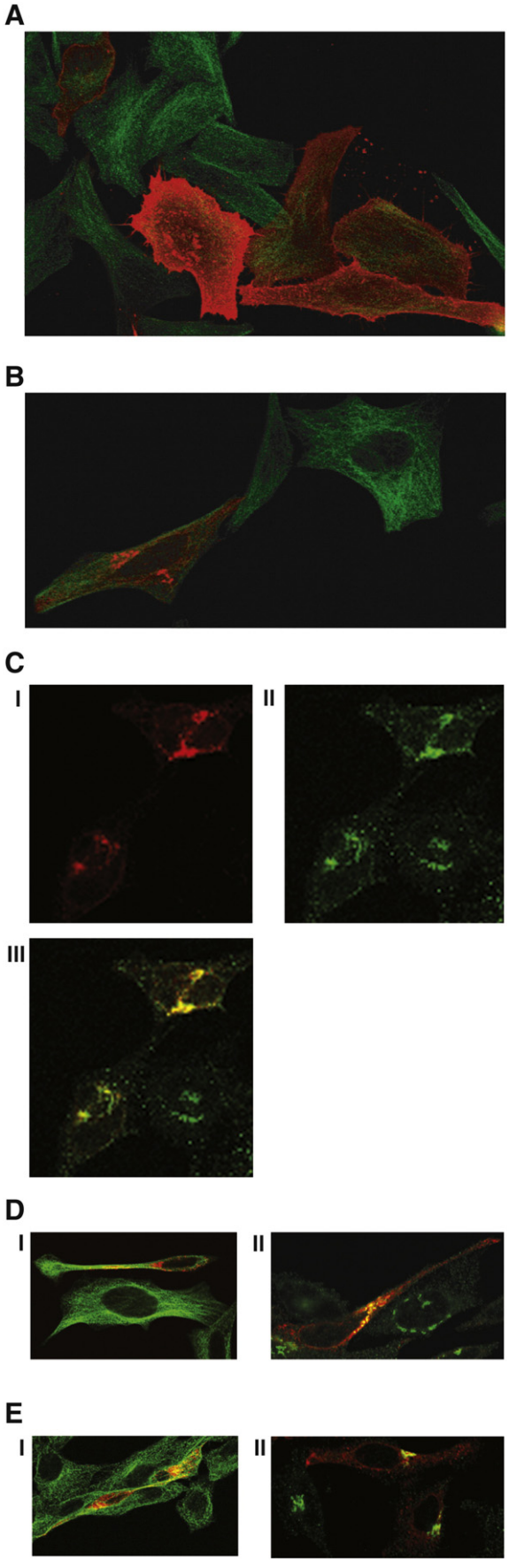
Comparing isoform protein size by immunoblot of plasmid-transfected HeLa cells

We cloned the new, largest AQP4e isoform cDNA into the pcDNA3.1/Zeo expression vector and transfected the plasmids into HeLa cells for immunoblot analysis. The AQP4e construct expressed three bands (Fig. 5A, lane 1). The highest band might correspond to the faint third isoform of rat brain seen by Neely and co-workers [10] (Fig. 5A, lane 2): however, the largest AQP4e band seemed slightly larger.

Two lower bands of the AQP4e expression construct were observed at the right size to correspond to the classical M1 and M23 bands (Fig. 5A, lanes 7 and lane 9, respectively). The AQP4e wild-type start codon therefore seemed to leak in translation.

Analysis of the Kozak sequence surrounding the AQP4 start codons in exon 0 and exon 1 (corresponding to the M1 and M23 isoforms, respectively) revealed that both were stronger (conformed more closely to the Kozak consensus sequence) than the new start codon in exon Z in rat, mouse, and man (Fig. 5B). Furthermore, the M23 start codon in exon 1 was stronger than the M1 start codon in exon 0 in rats, mouse, and humans, supporting the idea of leaky translation.

Fig. 5. Expression of various rat AQP4-isoforms in HeLa cells. (A) Immunoblot of plasmid transfected HeLa cells. Lane 1 is pcDNA3 -AQP4e, lane 3 is pcDNA3-AQP4f, and lane 5 is pcDNA3-wt-AQP4d. Identical-sized control rat brain lysates in lanes 2, 4, 6, and 8. Size controls with transfected plasmids of AQP4a (M1) and AQP4c (M23) in lanes 7 and 9, respectively. Size markers indicated to the left (in kDa). (B) Kozak analysis of start codon sequences in exon 0, exon 1, and exon Z in rat, mouse, and human. Optimal Kozak sequence shown above, with best bases in the crucial positions -3 and +4 shown in green, and second best base shown in blue [31]. (C) Expression of Kozak-optimized AQP4 isoforms compared with wild-type rat brain forms compared with immunoblot. Lanes 1–3 are rat brain membranes (20, 10, and 5 µg protein loaded). Lane 4 is pcDNA3-K-Aqp4d. Lane 5 is pXoom-AQP4e. Lane 6 is pcDNA3.1/Zeo(+)-K-Aqp4b. Lane 7 is pcDNA3.1/Zeo(+)-K-Aqp4f. Lanes 8 and 9 are rat brain membranes (20 and 5 µg). Lane 10 is Aqp4a (M1). Lane 11 is Aqp4c (M23). Lane 12 is 45 µg of rat kidney. The 30-kDa and 40-kDa markers are indicated on the left. (D) Water permeability (Lp) of oocytes expressing various AQP4 isoforms. Lp values were derived from swelling experiments where the removal of 20 mosM mannitol from the bathing solution induced oocyte volume increases. Values are expressed as mean±SEM (n=3–14). Column 1 is AQP4c (n=10), column 2 is AQP4e (n=14), column 3 is AQP4d (n=3), column 4 is AQP4f (n=3), and column 5 is uninjected, negative control (n=3).



The discrepancy between protein and mRNA expression levels of AQP4 might thus be partly explained by the mechanism of leaky scanning during translation initiation from the start codons of the mRNA [22]. It should be noted, however, that a wild-type AQP4a (M1) start codon showed only modest signs of leakage upon cell transfection and immunoblotting (data not shown). Why this start codon should leak more in the context of the AQP4e transcript is unclear, but it might be connected to the general translation preference of the first codon in an mRNA [22].

The wild-type start codons for the AQP4f and AQP4d isoforms gave only weak or nondetectable expression (Fig. 5A, lanes 3 and 5). To force the expression of single proteins we modified the wild-type Kozak sequences of AQP4e, AQP4f, and AQP4d plasmids to the optimized Kozak sequences of AQP4b, AQP4d, AQP4e and AQP4f plasmids to the optimized (Fig. 5B) and expressed these plasmids in HeLa cells. As a result mostly single bands were seen (Fig. 5C).

A comparison of the sizes of the various isoforms with AQP4 from brain (Fig. 5C, lanes 1–3, 8 and 9) again indicated that AQP4e might correspond to the faint, higher bands found in the rat brain (Fig. 5C, lane 5). Even with very high protein load, only a very faint band could be observed in rat kidney (Fig. 5C, lane 12).

Surprisingly, all isoforms lacking exon 2 ran at higher apparent molecular weights than expected. In fact, AQP4f appeared larger than AQP4e (Fig. 5C, lane 7 versus lane 5), AQP4b appeared larger than AQP4a (Fig. 5C, lane 6 versus lane 10), and AQP4d appeared larger than AQP4c (Fig. 5C, lane 4 versus lane 11), despite being smaller proteins (Fig. 1A). Thus the AQP4 protein isoform sizes could not be correlated with bands seen in immunoblots of rat brain. Unusual migration and the appearance of dimers of AQP4, even in denaturing gels, is a known problem in the field [17,23] Our analysis of exon 2-less isoforms illustrates that precise size determination of this kind of membrane protein is not easily obtained by SDS–PAGE immunoblot.

Water transport in Xenopus oocytes

The AQP4 isoforms were cloned into the *Xenopus* expression vector pXoom [13] and shown to express a single predominant protein on immunoblots (data not shown). Functional analysis of water transport in *Xenopus* oocytes showed that the AQP4e form is able to transport water (Fig. 5D). Oocytes injected with AQP4d did not transport water more efficiently than noninjected oocytes nor did AQP4f transport water above background levels (Fig. 5D). AQP4b was not cloned to pXoom because we so far lack direct empirical evidence for this isoform.

Fig. 6. (A) Confocal images of immunohistochemistry on astrocyte cell line CRL-2006 cells transfected with Aqp4e (AQP4 signal in red, tubulin in green). (B) HeLa cells transfected with AQP4d. (C) Colocalization of AQP4d (I) with *cis*-Golgi-marker GM130 (II) in HeLa cells. III has overlay of signals, with yellow signal showing colocalization. (D) HeLa cells transfected with AQP4b. I has antibody staining for AQP4 (red) and tubulin (green). II has antibody staining for AQP4 (red) and GM130 (green). (E) HeLa cells transfected with AQP4f. I has antibody staining for AQP4 (red) and tubulin (green). II has antibody staining for AQP4 (red) and GM130 (green). Yellow signal indicates colocalization.

Subcellular localization of AQP4 isoforms

Confocal microscopy of transfected HeLa cells showed that the AQP4e isoform was transported to the plasma membrane, with labeling of filopodia observable in the most highly transfected cells (Fig. 6A). Internal vesicle labeling was also seen, consistent with the proposed clathrin-dependent recycling of AQP4 [24]. The AQP4d isoform, on the other hand, seemed confined to intracellular locations (Fig. 6B), as did AQP4b and AQP4f (Figs. 6D and 6E). The AQP4d signal colocalizes with the *cis*-Golgi-marker GM130 (Fig. 6C) but not the ER marker PDI (data not shown). Also AQP4b and AQP4f are localized intracellularly but do not colocalize fully with GM130 (Figs. 6D and 6E). Furthermore, the intracellular localization of AQP4d could partly explain why this isoform does not transport water in the *Xenopus* oocyte assay [1] (Fig. 5D).

Discussion

In this study, we have identified several new isoforms of AQP4 in the rat. This finding has important implications, given that the rat has been the species of choice for most lines of investigations on brain aquaporins. Cloning and expression studies [1,2,25], freeze fracture analyses of square arrays in transfected cell lines [8,26,27], and immunogold localization studies [3] have all been performed in rat. Rat AQP4 was also used for the determination of the AQP4 pore structure [9].

Our study has revealed two new exons (exon Z and exon X) that code for two new types of mRNA transcript. Combined with the M1 transcript and exon 2-less transcripts by differential splicing, this brings the total number of AQP4 isoforms to six.

In this model, we have named these six cDNA isoforms AQP4a, AQP4b, AQP4c, AQP4d, AQP4e, and AQP4f, where AQP4a corresponds to the M1 cDNA and AQP4c corresponds to the M23 cDNA in the classical model of two AQP4 transcripts (Fig. 1A). We have shown by PCR and by analysis of RACE clones that the M23 mRNA transcript is unlikely to be coded by a transcription start at an extension of exon 1.

Significant regulation at the translational level is implied by quantitative analyses of RNA using TaqMan probes designed against the various exon–exon junctions and by analysis of the relative abundance of the RACE clones. The finding of a ~3:1 ratio of exon 0 versus exon X transcripts contrasts with the ~10-fold overabundance of M23 at the protein level [10]. This mismatch may be explained if the isoforms differ in their stability or if there is a differential regulation of mRNA translation.

Regulation of AQP4 mRNA translation has the important implication that the RNA level cannot be used as a proxy for protein measurements. In fact, because AQP4d overlaps AQP4c in the classical M23 band in Western blots and RNA levels do not reflect protein level, AQP4d could constitute a significant proportion of the protein in the M23 band in rat brain, despite the fact that TaqMan analysis showed the exon 2-less isoforms to constitute only 5–10% of the total mRNA pool.

A partial explanation for the isoform ratios differing between the RNA and the protein levels could be that some mRNAs are subjected to leaky scanning [22], a hypothesis that is supported by analysis of the Kozak sequences in AQP4 translation initiation sites. Leaky scanning would constitute a novel mechanism of AQP4 gene regulation.

We also demonstrate weak but consistent expression of AQP4 mRNA in the heart. A similar observation was made recently in mouse [28,29]. The function of AQP4 in the heart is unclear.

By expression studies in HeLa cells, we have shown that exon 2-less cDNA isoforms obtained from transfected cell lines migrate shorter on SDS–PAGE gels than predicted. As a consequence, the AQP4d protein isoform is detected at the same apparent molecular weight as AQP4c, making it technically difficult to assess by immunoblotting the expression level of AQP4d *in vivo*. Thus, the major endogenous AQP4 protein band, termed M23 in previous publications, is probably heterogenic in nature.

The novel cDNA isoforms AQP4e and AQP4f were initially considered natural candidates for the faint, higher protein bands in AQP4 found by Neely and co-workers [10]. As AQP4 exon 2-less isoforms migrate shorter than their exon 2-containing counterparts, AQP4b is an additional candidate for the higher bands. However, expression of AQP4b and AQP4e in HeLa cells gave rise to slightly differing apparent molecular weights on immunoblots, while AQP4f matched the highest immunoblot band found in rat brain. Differing posttranslational modifications between rat brain and HeLa cells may explain some of the migration discrepancies. Our study illustrates that immunoblot analysis has significant limitations as a method for determining the molecular weight of AQP4 isoforms.

Expression studies in HeLa cells showed that AQP4d is retained in the *cis*-Golgi instead of being transported to the plasma membrane. This localization could mirror a role of exon 2 in the cellular sorting of AQP4. The lack of exon 2 probably causes a loss of transport capacity, explaining why the AQP4d isoform was unable to transport water when expressed in *Xenopus* oocytes. Obviously, if oocytes (like HeLa cells) fail to translocate AQP4 to the plasma membrane we would not be able to record any water transport, irrespective of the inherent water transport capacity of this isoform. Interestingly, the new AQP4e isoform, which contains a large N-terminal domain and is transported to the membrane, displayed water transport. Further studies are needed to resolve the function of the new N-terminal domain.

Like the three novel isoforms AQP4b, AQP4d, and AQP4f, several other aquaporins are associated with intracellular membrane compartments. This is true for AQP2, AQP6, AQP8, AQP9, and AQP10 [4], and for AQP11 [30]. So far, information is scarce about the functional roles of intracellular aquaporins, including those described here.

The studies described here represent a contribution to a more detailed elucidation of the molecular aspects of AQP4. As rodents are the principal animals used for studies of water transport in mammals, the complexity of the AQP4 gene presently revealed will have implications for the design and interpretation of future experiments aimed at revealing the precise functions of AQP4 in brain.

Materials and methods

RNA isolation, Northerns, RACE, topo-cloning, and sequence analysis

RNA was isolated from 50-mg tissue samples using RNeasy (Qiagen). The RNA concentration was determined using NanoDrop (NanoDrop Technologies, Wilmington, USA). For first-strand cDNA synthesis in RACE (5' RACE System for Rapid Amplification of cDNA Ends, Version 2.0; Invitrogen), we used 1.5 μg of RNA obtained from hippocampus, cerebellum, and kidney of female Wistar rats. First-strand primer used was +530, and M23_515 and AAP (Invitrogen) primers were used for RACE PCR.

PCR products were isolated from agarose gels using GFX PCR DNA and Gel Band Purification Kit (Amersham), cloned into plasmids using Topo TA Cloning (Invitrogen), and sequenced. The sequenced cDNA clones are shown in Supplementary Fig. 2.

Analysis and alignment of sequences were performed using Vector NTI (Invitrogen) and BLAST (at NCBI online). EST databases were accessed and analyzed through the PubMed Genome Entrez system. Northern tissue blots were purchased from BD Biosciences (Amersham) and probed using AQP4 CDS region essentially as described [12].

PCR and PCR primers

PCR and RACE PCR were performed using BioTaq (Q-BIOgene) buffers and enzyme. DNA primers were designed using Oligo 6.68 (Molecular Biology Insights Inc.) and purchased from Eurogentec (Belgium). The sequences of the primers used in this study were as folows.

Primer name	Sequence	Primer name	Sequence
-1013	TGCCTGTCCCTAAGGTCCAC	yAQP4_41	CCCAGAAGACAGCACCTGTG
-774	GGTGTGTGCTTGCATAATGA	yAQP4_75	CTCTGCCACCTGGCTACAAC
-730	GTCCCCCTGTCCAACCATCT	yAQP4_269-Kz	AAGCTTGCCACCATGGTGGCTTTCAAAGG
-700	CACTCCAATCCCTGGGCTAC	yAQP4_269-pX	GAATTCGCCACCATGGTGGCTTTCAAAGG
-658	TCTCCCTGCAGTGGACATCA	AQP4-M1	AAGCTTGCCACCATGAGTGACGGAGCTGC
-624	GCACAACAGATGCTGGGAAT	M1_pr	AGGGAAGGCATGAGTGACGG
-329	GCTCAAAACTACACGCGAAA	M23_515	ACAGCTGGCAAAAATGGTGA
-211	TTACCGTCCAGCGAATAGGA	zAQP4_10	AACCGCCCTGTGTCTATAGT
-110	CGTTCAGAAAACCATCTCCA	zAQP4_79-pX	GAATTCGCCACCATGGTGCAGAATCTTTCC
-41	TTCCACTGGAGGCTTCTCAT	zAQP4_79	AAGCTTGCCACCATGGTGCAGAATCTTTCC
+16	CCTGCAGCAGAGAGAGCATC	zAQP4_1181	TTCTGTTTTCAGTGCTGTCC
+530	CCAATTGCTAAAGCAACGGA	zAQP4_1201	GGCCAGTTCTAGGGAGTCTC
M23_637	CCAAAGGATCGAGCTGGATT	zAQP4-1201-Kz	CTCGAGGCCAGTTCTAGGGAGTCTC
M23_757	TTGAGCTCCACGTCAGGACA	zAQP4-1201-pX	GCGGCCGCCAGTTCTAGGGAGTCTC

TaqMan Assays

RNA (2 μg) was converted into cDNA using High Capacity cDNA Archive Kit (Applied Biosystems). The cDNA level was measured using TaqMan assays (Applied Biosystems) as indicated in figures, standardized to RNA levels or a GAPDH control assay, on a 7900HT Fast Real-Time PCR System (Applied Biosystems).

Expression plasmid construction and GenBank submission

The cDNAs of interest were subcloned directly from pCR-II-Topo plasmids (see Supplementary Fig. 2) or indirectly using modified PCR primers into pcDNA3, pcDNA3.1/Zeo(+), or pXoom. The sequences of the following nine expression constructs have been submitted to GenBank: pcDNA3.1/Zeo(+)-K-AQP4b (EF437949), pcDNA3-K-AQP4d (EF437950), pcDNA3-wt-AQP4d (EF437951), pXoom-AQP4d (EF437952), pcDNA3-AQP4e (EF437953), pcDNA3.1/Zeo(+)-K-AQP4e (EF437954), pXoom-AQP4e (EF437955), pcDNA3-AQP4f (EF437956), and pcDNA3.1/Zeo(+)-K-AQP4f (EF437957).

Cell culture and plasmid transfections

HeLa cells and CRL2006 astrocyte cell lines were maintained in Dulbecco's modified Eagle's medium supplemented with 1% L-glutamine (Cambrex) and 10% fetal calf serum (Gibco BRL). The cells were passaged at subconfluency and transfected with plasmids using 200 μl OptiMem (Gibco), FuGENE6 (Roche) at a 3:1 ratio (μl:μg nucleic acid) and 0.17 μg DNA per cm^2. Medium was changed after 4 h. Cells were harvested 24 h after transfection.

SDS–PAGE/immunoblotting

Tissue lysates were obtained from male PVG rats. The tissue of interest was dissected and homogenized in lysis buffer (10 mM sodium phosphate, pH 7.4, 1% SDS, 150 mM NaCl, 5 mM EDTA, protease inhibitor cocktail (Roche)). Lysates were sonicated and centrifuged at 1000*g* for 10 min and pellets discarded. Total protein was determined using DC total protein kit (Bio-Rad) according to the manufacturer's instructions. LiDS-loading buffer (Invitrogen) supplemented with 100 mM ditheothreitol was used.

Cell lysates were made by washing cells twice with phosphate-buffered salins (PBS) and lysing in the well for 30 min at room temperature with agitation using 200 μl lysis buffer per 9.4 cm^2. Lysates were pipetted off, sonicated, and centrifuged at 23,000*g* for 5 min. To concentrate the lysate, the supernatant was concentrated to ~30 μl using 10-kDa cutoff spin columns (PALL) at 12,000*g*. This concentration step did not affect the isoform ratio (data not shown). The resulting retentate was subjected to total protein determination and loaded on gels in the same way as tissue lysates.

SDS–PAGE and immunoblotting has been optimized with respect to AQP4 resolution [23]. Briefly, gels were 12% Laemmli PAGE supplemented with 3 M urea in the stacking gel and the resolving gel. MagicMarkXP (Invitrogen LC5602) was employed as molecular weight marker. Proteins were blotted onto polyvinylidene

fluoride membranes that were blocked with 5% milk, incubated with primary antibody diluted in the blocking solution overnight at 4 °C, and detected with the ECF kit according to the manufacturer's instructions (Amersham).

Antibodies

Rabbit-anti-AQP4 (Chemicon AB3068), goat-anti-AQP4 (Santa Cruz SC9888), and rabbit-anti-AQP1 (Alpha Diagnostics AQP11-A) were used for immunoblots at a final concentration of 1 μg/ml. AB3068 diluted 1:500 was used for cell immunofluorescence. Mouse-anti-GM130 (BD Biosciences 610822) was used at a final dilution 1:1000. Mouse-anti-α-tubulin (Molecular probes A11126) was used at 1:1000 dilution. Cell nuclei were stained with TOTO-3 (Invitrogen T3604) diluted 1:500. Secondary antibodies for cell staining included Alexa488 conjugates (Molecular probes) and Cy3 conjugates (Jackson ImmunoResearch) used according to the manufacturers' instructions.

Cell immunofluorescence

Cells were fixed for 30 min at room temperature with 4% formaldehyde in 0.1 M sodium phosphate, pH 7.4, and then blocked for 15 min with PBS supplemented with 0.5% v/v Triton X-100 and 3% bovine serum albumin (PBST–BSA). Primary antibodies were diluted in PBST–BSA and cells were incubated 3 h or overnight at 4 °C. Following three 5-min washes with PBS, secondary antibodies were diluted in PBST–BSA and incubated 1 h at room temperature. Finally, cells were washed three times for 5 min or more in PBS and mounted using ProLong AntifadeGold with DAPI (Invitrogen). Pictures were captured on a Leica LSM5 confocal microscope.

Expression in Xenopus laevis oocytes

The rat AQP4 isoforms AQP4d, AQP4e, and AQP4f were subcloned into the pXOOM vector [13]. The rAQP4.M23 isoform (AQP4c), subcloned into pBlueScript SK(-), was a gift from Dr. Søren Nielsen. The rAQP4 cDNA constructs were linearized with *XhoI*, and cRNA was synthesized by the mMessage mMachine *in vitro* transcription kit (Ambion, Austin, TX) with T7 or T3 RNA polymerase. Stage V–VI *Xenopus* laevis oocytes were isolated and defolliculated with collagenase (Boehringer Mannheim) as described [14]. All oocyte collection procedures conformed to Danish Animal Ethics Regulations. cRNA (50 ng) was injected into defolliculated oocytes, and these were incubated for 3–6 days at 19 °C in Kulori buffer (90 mM NaCl, 1 mM KCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, 5 mM Hepes, pH 7.4) before water permeability measurements.

The water permeabilities of the various AQP4 isoforms were investigated as described [14,15]. Briefly, a single oocyte was placed in a continuously perfused 30-μl chamber and impaled with two microelectrodes that served to determine the membrane potential and stabilize the oocyte. Oocyte volume was monitored on-line from below via an inverted microscope with a time resolution of 1 s. The oocytes were illuminated from above via a Plexiglas rod. Measurements were performed at 20–23 °C. The control bathing solution contained (mM): 90 NaCl, 20 mannitol, 2 KCl, 1 $CaCl_2$, 1 $MgCl_2$, 10 Hepes, 214 mosm/L, pH 7.4.

Osmotic water permeability, Lp [μm/s], was calculated from the initial oocyte volume (1.35 μl), the initial slope of the relative volume increase observed upon removal of 20 mOsm of mannitol from the bathing solution (d(V/V0)/dt, usually measured within 10 s), the molecular water volume (Vw = 18 cm^3/mol), the osmotic gradient (osmin - osmout), and the true oocyte surface area (S = 0.53 cm^2) by the equation $Lp = V0 \times d(V/V0)/dt\ [S \times Vw \times (osmin - osmout)]$ [16].

Acknowledgments

Jan Gunnar Sorbo and Svein Erik Moe are supported by grants from Forskerlinjen, the Medical Student Research Programme at the University of Oslo. Torgeir Holen has a grant from The Norwegian Cancer Society. We thank Søren Nielsen for the M1 and M23 plasmids. This work was supported by the Nordic Centre of Excellence Program in Molecular Medicine, the Research Council of Norway (Storforsk Program), and FUGE (Norwegian functional genomics program).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ygeno.2007.12.003.

References

[1] H. Hasegawa, T. Ma, W. Skach, M.A. Matthay, A.S. Verkman, Molecular cloning of a mercurial-insensitive water channel expressed in selected water-transporting tissues, J. Biol. Chem. 269 (1994) 5497–5500.
[2] J.S. Jung, et al., Molecular characterization of an aquaporin cDNA from brain: candidate osmoreceptor and regulator of water balance, Proc. Natl. Acad. Sci. USA 91 (1994) 13052–13056.
[3] S. Nielsen, et al., Specialized membrane domains for water transport in glial cells: high-resolution immunogold cytochemistry of aquaporin-4 in rat brain, J. Neurosci 17 (1997) 171–180.
[4] L.S. King, D. Kozono, P. Agre, From structure to disease: the evolving tale of aquaporin biology, Nat. Rev. Mol. Cell Biol 5 (2004) 687–698.
[5] A.S. Verkman, D.K. Binder, O. Bloch, K. Auguste, M.C. Papadopoulos, Three distinct roles of aquaporin-4 in brain function revealed by knockout mice, Biochim. Biophys. Acta 1758 (2006) 1085–1093.
[6] M. Amiry-Moghaddam, O.P. Ottersen, The molecular basis of water transport in the brain, Nat. Rev. Neurosci 4 (2003) 991–1001.
[7] G.T. Manley, et al., Aquaporin-4 deletion in mice reduces brain edema after acute water intoxication and ischemic stroke, Nat. Med 6 (2000) 159–163.
[8] C.S. Furman, et al., Aquaporin-4 square array assembly: opposing actions of M1 and M23 isoforms, Proc. Natl. Acad. Sci. USA 100 (2003) 13609–13614.
[9] Y. Hiroaki, et al., Implications of the aquaporin-4 structure on array formation and cell adhesion, J. Mol. Biol 355 (2006) 628–639.
[10] J.D. Neely, B.M. Christensen, S. Nielsen, P. Agre, Heterotetrameric composition of aquaporin-4 water channels, Biochemistry 38 (1999) 11156–11163.
[11] C. Silberstein, et al., Membrane organization and function of M1 and M23 isoforms of aquaporin-4 in epithelial cells, Am. J. Physiol. Renal Physiol. 287 (2004) F501–F511.
[12] T. Holen, et al., Tolerated wobble mutations in siRNAs decrease specificity, but can enhance activity in vivo, Nucleic Acids Res 33 (2005) 4704–4710.
[13] T. Jespersen, M. Grunnet, K. Angelo, D.A. Klaerke, S.P. Olesen, Dual-function vector for protein expression in both mammalian cells and *Xenopus laevis* oocytes, Biotechniques 32 (2002) 536–538 540.
[14] T. Zeuthen, et al., Water transport by the Na+/glucose cotransporter under isotonic conditions, Biol. Cell 89 (1997) 307–312.
[15] A.K. Meinild, D.A. Klaerke, T. Zeuthen, Bidirectional water fluxes and specificity for small hydrophilic molecules in aquaporins 0-5, J. Biol. Chem 273 (1998) 32446–32451.
[16] L.M. Holm, et al., NH3 and NH4+ permeability in aquaporin-expressing *Xenopus* oocytes, Pflugers Arch 450 (2005) 415–428.
[17] M. Lu, et al., The human AQP4 gene: definition of the locus encoding two water channel polypeptides in brain, Proc. Natl. Acad. Sci. USA 93 (1996) 10908–10912.

[18] A.N. Mhatre, R.E. Stern, J. Li, A.K. Lalwani, Aquaporin 4 expression in the mammalian inner ear and its role in hearing, Biochem. Biophys. Res. Commun 297 (2002) 987–996.
[19] S. Zelenin, E. Gunnarson, T. Alikina, A. Bondar, A. Aperia, Identification of a new form of AQP4 mRNA that is developmentally expressed in mouse brain, Pediatr. Res 48 (2000) 335–339.
[20] K. Murata, et al., Structural determinants of water permeation through aquaporin-1, Nature 407 (2000) 599–605.
[21] H. Wen, et al., Ontogeny of water transport in rat brain: postnatal expression of the aquaporin-4 water channel, Eur. J. Neurosci 11 (1999) 935–945.
[22] M. Kozak, Pushing the limits of the scanning mechanism for initiation of translation, Gene 299 (2002) 1–34.
[23] J.G. Sorbo, S.E. Moe, T. Holen, Early upregulation in nasal epithelium and strong expression in olfactory bulb glomeruli suggest a role for Aquaporin-4 in olfaction, FEBS Lett 581 (2007) 4884–4890.
[24] R. Madrid, et al., Polarized trafficking and surface expression of the AQP4 water channel are coordinated by serial and regulated interactions with different clathrin-adaptor complexes, EMBO J. 20 (2001) 7008–7021.
[25] L.S. King, S. Nielsen, P. Agre, Aquaporins in complex tissues. I. Developmental patterns in respiratory and glandular tissues of rat, Am. J. Physiol. 273 (1997) C1541–C1548.
[26] B. Yang, D. Brown, A.S. Verkman, The mercurial insensitive water channel (AQP-4) forms orthogonal arrays in stably transfected Chinese hamster ovary cells, J. Biol. Chem 271 (1996) 4577–4580.
[27] J.E. Rash, T. Yasumura, C.S. Hudson, P. Agre, S. Nielsen, Direct immunogold labeling of aquaporin-4 in square arrays of astrocyte and ependymocyte plasma membranes in rat brain and spinal cord, Proc. Natl. Acad. Sci. USA 95 (1998) 11981–11986.
[28] T.L. Butler, et al., Cardiac aquaporin expression in humans, rats, and mice, Am. J. Physiol. Heart Circ. Physiol. 291 (2006) H705–H713.
[29] A. Warth, et al., Upregulation of the Water Channel Aquaporin-4 as a Potential Cause of Postischemic Cell Swelling in a Murine Model of Myocardial Infarction, Cardiology 107 (2007) 402–410.
[30] D.A. Gorelick, J. Praetorius, T. Tsunenari, S. Nielsen, P. Agre, Aquaporin-11: a channel protein lacking apparent transport function expressed in brain, BMC. Biochemistry 7 (2006) 14.
[31] M. Kozak, An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs, Nucleic Acids Res 15 (1987) 8125–8148.