The Major Intrinsic Protein Family of Arabidopsis Has 23 Members That Form Three Distinct Groups with Functional Aquaporins in Each Group

Alfons Weig, Corine Deswarte, and Maarten J. Chrispeels*
Department of Biology, 0116, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093–0116

Aquaporins, proteins that enhance the permeability of biological membranes to water, are widely distributed in living organisms. They are 26- to 29-kD proteins that belong to the major intrinsic protein (MIP) family of channels. By searching the Arabidopsis thaliana expressed sequence tag database and by using the polymerase chain reaction with oligonucleotides to conserved plant aquaporin domains, we identified 23 expressed Arabidopsis MIP genes. Eight of these had been previously identified as active aquaporins, and two additional ones are now reported to have water-transport activity in Xenopus laevis oocytes. One of these is highly expressed in suspension-cultured cells. On a dendrogram these 23 MIP sequences cluster into three groups: the first group has 11 members and contains the plasma membrane aquaporins, the second group also has 11 members and contains the tonoplast aquaporins, and the third group has only a single member. This MIP protein, provisionally called At-NLM1, is most closely related to the Gm-NOD26 protein that is found in the bacteroid membranes of soybean (Glycine max L.) nodules; At-NLM1 is an active aquaporin when expressed in oocytes. With a semiquantitative slot-blot analysis technique, we determined the expression levels of 22 MIP genes in the various organs. The quantitative polymerase chain reaction was used to determine the effects of various stress treatments on the expression of NLM1.

In growing plants nonlimiting water movement through cellular membranes is required by a large number of cellular and physiological processes such as cell enlargement, stomatal movement, phloem loading, and transpiration, to name but a few. Yet, until recently, the molecular basis of these processes remained obscure. The identification of integral membrane proteins that act as water-selective channels or aquaporins, whereas others, such as MIP itself (Zampighi et al., 1985) and the soybean nodule peribacteroid membrane protein Gm-NOD26 (Weaver et al., 1994), may be ion channels. However, whether a protein of the MIP family is classified as an ion channel or a water channel may depend on the method used to establish its function. When bovine lens MIP is expressed in Xenopus laevis oocytes, it is clearly an aquaporin (Mulders et al., 1995), but when the protein is incorporated into a lipid bilayer, it behaves as an ion channel (Zampighi et al., 1985). Bacteria have MIPs that are glycerol channels (Maurel et al., 1994), and some mammalian MIPs are bifunctional, transporting water and ions (Yool et al., 1996) or glycerol, urea, and water (Ishibashi et al., 1994).

In humans five different aquaporin cDNAs have been identified. Aquaporins are very abundant in the kidney, where they are responsible for the concentration of urine (Chrispeels and Agre, 1994; van Os et al., 1994; Brown et al., 1995). They are also present in other organs, where their functions remain to be elucidated. In plants MIP homologs were identified in at least 15 different species by a recent search of the EMBL/GenBank/DBJ databases.

In plants aquaporins are likely to function in a variety of processes that require regulated water flow (for review, see Maurel, 1997). We can distinguish transport processes requiring transcellular water flow from those that accompany intracellular osmotic adjustments. For example, the transpiration stream and phloem loading are processes that must be accompanied by transcellular water flow through living tissues to sustain them. Cell enlargement, opening of stomates, and movement of pulvini are examples of processes that require short-range transcellular water movement, as well as osmotic adjustments between the vacuole and the cytosol or the apoplast. Osmotic adjustments undoubtedly also occur in cells that hydrolize their vacuolar reserves (cotyledons and fruit mesocarp) or take up salts (root hairs) or metabolites (transfer cells and aquaporins). Plant Physiol. (1997) 114: 1347–1357

Abbreviations: EST, expressed sequence tag; MIP, major intrinsic protein; MS, Murashige-Skoog basal medium with Gamborg’s vitamins; PIP, plasma membrane intrinsic protein; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcription-coupled PCR; TIP, tonoplast intrinsic protein.

1 This work was supported by grants to M.J.C. from the National Science Foundation (Cell Biology Program) and from the University of California Biotechnology Program. C.D. was supported by a Lavoisier fellowship from the government of France.
2 Present address: Plant Physiology, University of Bayreuth, 95440 Bayreuth, Germany.
* Corresponding author; e-mail mchrispeels@ucsd.edu; fax 1–619–534–4052.
scutellar epithelium). To understand the multiple roles that aquaporins play in plants, we will have to catalog them and study their expression patterns.

In plant cells aquaporins are found in the plasma membrane and in the vacuolar membrane (tonoplast) (Höfte et al., 1992; Kammerloher et al., 1994; Daniels et al., 1996; Robinson et al., 1996) and we refer to them as PIPs and TIPs, respectively (see Chrispeels and Maurel, 1994). A few plasma membrane and tonoplas membrane aquaporins have been described and there is some evidence that individual genes are usually not expressed throughout the entire plant and that expression is sometimes enhanced or triggered by environmental stimuli. For example, promoter-GUS fusions showed that At-γTIP is expressed primarily in roots but also in leaves, bolts, and flowers (pedicel) (Ludevid et al., 1992), whereas the expression of At-8TIP, which is especially high in the vascular system of the shoot, is barely detectable in the root (Daniels et al., 1996). The expression of At-γTIP appears to be associated with cell elongation (Ludevid et al., 1992; Phillips and Huttly, 1994).

Genes encoding PIPs are also differentially expressed. Promoter-GUS fusions with At-PIP1b (AthH2) showed expression in root tips, cotyledons, and newly formed organs, with high expression in the vascular system. This aquaporin is expressed at much higher levels in the guard cells compared with the other epidermal cells (Kaldenhoff et al., 1995). Expression of At-RD28, a plasma membrane aquaporin, is much higher in shoots than in roots (Daniels et al., 1994).

Some environmental factors enhance the expression of aquaporins. For example, At-RD28 was first identified in a cDNA library of water-stressed Arabidopsis plants (Yamaguchi-Shinozaki et al., 1992), and a related MIP, not yet shown to be a water-channel protein, is induced in water-stressed pea (Pisum sativum) plants (Guerrero et al., 1990). Yamada et al. (1995) identified a number of salt-repressed MIPs in Mesembryanthemum crystallinum. All of these results suggest that plants modulate the abundance of aquaporin mRNAs in their tissues. Such modulation of aquaporin abundance and regulation of aquaporin activity by phosphorylation (Maurel et al., 1995) may be the major ways that plant cells alter their membrane permeability to water and transcellular water flow.

In this article we describe the MIP gene family of Arabidopsis as consisting of at least 23 members, comprising three major subfamilies according to their amino acid sequence. The mRNA levels of the individual genes were investigated in different organs of Arabidopsis. We also report the water-channel activity of At-PIP3, a new aquaporin that is highly expressed in Arabidopsis suspension-cultured cells, and of At-NLM1, a new aquaporin in a new subgroup of plant MIPs.

MATERIALS AND METHODS

Database Searches

Arabidopsis ESTs in the EMBL database were searched with the TFASTA program (Pearson et al., 1988) using the amino acid sequences of At-γTIP or At-RD28. From both searches, 174 hits with significant similarity to the query sequences were further analyzed. In a first step, the EST DNA sequences were compared with published full-length DNA sequences of Arabidopsis MIP genes; transcripts of the same gene could be identified using the program GELOVERLAP (Genetics Computer Group, Madison, WI) with the following parameters: identity = 0.98, overlap size = 40. In a second step, these clusters and the remaining EST sequences were individually tested using the programs FASTA and GAP (Genetics Computer Group). Finally, only those cDNA sequences were taken as transcripts of individual genes that overlapped with all of the other genes or partial cDNAs but were obviously not identical to one another.

At-NLM1, which was not identified using the computer analysis described above, was identified from a list of nonredundant ESTs by Cooke et al. (1996), and a full-length clone was isolated using the RACE technique (see below).

Cell-Suspension Culture

An Arabidopsis thaliana (var Columbia) cell-suspension culture (Axelos et al., 1992) was maintained at 25°C with a 16-h photoperiod and under constant agitation (120 rpm). Subculturing was performed every 7 d by adding 10 mL of the suspension to a 250-mL flask containing 60 mL of Gamborg's B5 basal medium with minimal organics (Sigma), 20 g/L of Suc, and 2.5 μM 2,4-D.

Plant Growth and RNA Extraction

Roots, rosette leaves, bolts, flowers, and siliques were harvested from A. thaliana grown for 5 to 6 weeks on soil, as described by Daniels et al. (1994). The effect of ABA, salt, or osmotic stress and etiolation or drought were measured in 2-week-old Arabidopsis seedlings grown on MS agar (0.8%, Sigma) plates; seeds were sterilized subsequently in 50% ethanol for 1 min and in 50% bleach and 0.1% Tween 20 for 10 min. The seeds were washed five times by low-speed centrifugation with sterile water and finally suspended in sterile water. The sterile seeds were pipetted onto a sterile nylon membrane that was laid on the MS agar. The seeds were vernalized in the cold room for 2 d and then transferred to room temperature with long daylight (16-h light/8-h dark). The seedlings were grown vertically for 12 to 14 d. For hormone or stress treatment, the nylon membranes carrying the seedlings were transferred onto a new MS agar plate containing 0.1 mM ABA, 0.3 M NaCl, or 0.3 M mannitol. The seedlings were grown for another 24 h, harvested, and stored in liquid nitrogen. Drought stress was applied by leaving the nylon membrane with the plants on the bench for 1 to 2.5 h.

Total RNA was extracted using Tip-20 columns (Qiagen, Chatsworth, CA), according to the protocol of the manufacturer. Poly(A⁺) RNA was isolated from total RNA using the poly(A⁺)Tract kit (Promega) following the instructions of the manufacturer.
Identification of MIP cDNAs by RT-PCR

cDNA was synthesized from 1 to 5 μg of total RNA using oligo(dT)12-18 as a primer and Moloney murine leukemia virus reverse transcriptase (GIBCO-BRL). MIP-specific cDNAs were amplified in a 50-μL reaction mixture containing 2 μL of cDNA, 1× PCR buffer (Perkin-Elmer), 0.2 mm deoxynucleotidetriphosphate (each), 50 μM TIP2 primer (5'-GG[AT]CC[AT][AG]CCCA[AG]TA[AG][A][CT][CCA -3'], 50 μM TIP4 primer (5'-CA[CT][GT][GT][A][A][CT][CC][AT][GT][GT][AT][AC-3'], and 1 unit of AmpliTaq DNA polymerase (Perkin-Elmer). Amplification was done for 30 cycles at 95°C for 60s, 55°C for 60s, and 72°C for 90s. The reaction products were separated on a 2% agarose gel, and visible bands of the expected size (about 422 bp for TIPs and 452 bp for PIPs) were eluted and subcloned into a pCRII vector using the TA-cloning kit (Invitrogen, San Diego, CA).

cDNA Cloning of NLM1

A NLM1 full-length cDNA was obtained using the 5'/3' RACE kit (Boehringer Mannheim). About 100 ng of polyadenylated RNA was reverse-transcribed using 1 μg of NLM1-specific antisense primer ATNODR3 (5'-CAAAGGTTGTTGAAAGTTG). A 900-bp PCR product was cloned into pCRII (TA cloning kit, Invitrogen) and sequenced. The DNA sequence on one side was identical to the 3' end of clone 170P5T7 and showed on the other end two putative start codons on the same frame, which might represent the start of the coding region. To minimize PCR amplification errors, a new full-length NLM1 cDNA was amplified, using Pfu polymerase (Stratagene) with proofreading activity and 10 μM specific primers to the 5' and 3' noncoding regions (ATNODF3B: 5'-CCCCGATCGTTCAATCCTTCTTGCATTCC; ATNODR4B: 5'-CCGGTACCGAGGTGGTCG-TGAAAAGTGTG, incorporating BamHI sites on both ends, and subcloned into the BamHI site of plBluescript II SK+(Stratagene). The cDNA source for this final PCR was prepared from two independent RNA preparations from the roots of 2-week-old light- or dark-grown Arabidopsis seedlings, respectively. The cDNA sequence of the two independently generated cDNA clones was fully identical.

Osmotic Water Permeability Assay

Fully grown oocytes (stages V and VI) were isolated from Xenopus laevis and incubated in Barth's solution (88 mm NaCl, 1 mm KCl, 2.4 mm NaHCO3, 10 mm Hepes-NaOH, 0.33 mm Ca(NO3)2, 0.41 mm CaCl2, and 0.82 mm MgSO4, pH 7.4) as described previously (Cao et al., 1992).

In vitro transcripts (1 mg/mL) or nuclease-free water in 50 nL were injected and the oocytes were kept at 18°C.

Three days after injection the oocytes were transferred from Barth's solution (200 mosmol) to the same solution diluted to 40 mosmol with distilled water. Changes in cell volume were followed with a microscope by taking photographs at 5-s intervals. The images were captured (NIH Image software, National Institutes of Health, Bethesda, MD), the oocyte surface areas were calculated using a program designed in the laboratory of J.E. Hall (Department of Biophysics and Physiology, University of California, Irvine), and they were used for the calculation of oocyte relative volume. The osmotic permeability coefficient (Pf) was calculated using the equation:

\[ Pf = V_o\left(\frac{d(V/V_o)dt}{S}\right)\left[\frac{V_m}{Osm_{in}} - \frac{Osm_{out}}{V_m}\right], \]

where the initial volume \(V_o\) = 9 × 10^{-4} cm³, the initial oocyte surface area \(S\) = 0.045 cm², and the molar volume of water \(V_m\) = 18 cm³/mol (Zhang and Verkman, 1991).

Slot-Blot Analysis

Plasmids for each type of MIP gene were either from our laboratory or obtained from the Arabidopsis Biological Resource Center (Ohio State University, Columbus) or from Institut National de la Recherche Agronomique (Versailles, France). For unpublished sequences, the GenBank accession number is given throughout the text. Slot blots (Schleicher & Schuell) on nylon membranes were prepared with 100 ng of plasmid in 100 μL of Tris-EDTA buffer. The membrane with the plasmid DNA was denatured by transferring for 3 min onto filter paper soaked in 0.5 M NaOH with 1.5 mM NaCl. The membrane was neutralized for 5 min on filter paper, soaked in 1 M Tris-HCl (pH 7.5) with 1.5 M NaCl, and washed for 30 s in 2× SSC (1× SSC = 0.15 M NaCl and 0.015 M sodium citrate). The membrane was baked for 30 to 60 min at 80°C in a vacuum oven. Prehybridization and hybridization were carried out at 42°C in 50% formamide, 2% blocking reagent (Boehringer Mannheim), 5× SSC, 0.1% SDS, and 0.02% sarcosyl. Digoxigenin-labeled cDNA was synthesized using 10 μg of total RNA in the reaction mixture described for RT-PCR above, in which the unlabeled nucleotides were replaced with 2.5 μL of the digoxigenin DNA-labeling mixture (Boehringer Mannheim), and denatured by boiling prior to the addition of the hybridization solution. After an overnight hybridization the membranes were washed once in 1× SSC with 0.1% SDS for 15 min at room temperature and twice in 0.1× SSC with 0.1% SDS at 68°C for 15 min. Hybridized cDNA was detected using the chemiluminescence of CSPD (Boehringer Mannheim) according the manufacturer's instructions. X-ray films were exposed from 10 min to several hours. The signals were scanned and quantified using the NIH image 1.58 software.

Quantitative PCR

Quantitative PCR was performed as described by Weig and Komor (1996). One microgram of total RNA together
with 10^6 copies of AW109 RNA (Perkin-Elmer) was reverse-transcribed into cDNA in a 50-µL reaction (see “Identification of MIP cDNAs by RT-PCR” above). From this cDNA pool, specific fragments of At-RD28, At-γTIP, At-ΔTIP, and β-tubulin-4 were amplified with gene-specific primers over 25 cycles (At-RD28, At-γTIP, and At-ΔTIP) or 30 cycles (At-ΔTIP and β-tubulin-4). The six-step PCR profile was 95°C (ramp, fast; hold, 30 s), 50°C (ramp, 120 s; hold, 30 s), and 72°C (ramp, 30 s; hold, 60 s). A short stretch of the synthetic control RNA AW109 was simultaneously amplified with the primers DM151 and DM152 (Perkin-Elmer). Gene-specific primers were At-γTIP (forward: 5'-GCATGCCGATCAGAAACATC; reverse: 5'-TCAGTAGTCTGTGGTTGG), At-STIP (forward: 5'-GCTTACCTCGCTGAGTTC; reverse: 5'-GACAGTGATGACCAAAGCGAAGG), At-RD28 (forward: 5'-CACAAACACTTACCG; reverse: 5'-GAAGTAGTCTGTGGTTGG).

RESULTS

Identification of At-PIP3 as a Functional Aquaporin

All aquaporins that have been characterized have a predicted topology, with six membrane-spanning domains connected by five loops and short N-terminal and C-terminal domains in the cytoplasm. A comparison of their amino acid sequences showed the presence of two conserved sequences: HVNPAVT in the cytoplasmic loop between transmembrane domains 2 and 3 and WVYWVGP in the sixth transmembrane domain (for the most recent review of the MIP family, see Park and Saier, 1996). These conserved amino acids were used to design degenerate oligonucleotide primers for RT-PCR. Using these oligonucleotides with mRNA from suspension-cultured Arabidopsis cells, we obtained a clone with a 0.42-kb insert. Nucleotide sequencing showed that it encoded a portion of an MIP protein and that the same cDNA was present in the A. thaliana EST database (accession no. T21533). The complete nucleotide sequence and the derived amino acid (Fig. 1A) indicate that this is a typical member of the MIP family. The actual amino acid sequences corresponding to the conserved region that were used to construct the degenerate oligonucleotides are underlined in Figure 1A. In the first underlined sequence six of the seven amino acids are identical, and in the second one, five of the seven amino acids are identical to the sequence defined by the nucleotide primers. Because of its close relationship to the PIP1 and PIP2 families (68.0 and 77.5% identity at the amino acid level to PIP1a and PIP2a, respectively), we have named it PIP3, although we do not yet have evidence that it is a PIP.

To find out whether PIP3 is an aquaporin, the cDNA was cloned in a vector for in vitro synthesis of cRNA, and the cRNA was injected into X. laevis oocytes. Three days later the oocytes were used for swelling assays by exposing them to a hypertonic solution (Fig. 2). The presence of PIP3 in the oocytes increased the osmotic permeability coefficient (Pf) of the oocyte plasma membrane from 2.0 \times 10^{-3} to 11.1 \times 10^{-3} cm s^{-1}.

The Arabidopsis MIP Gene Family

The search for Arabidopsis entries in the EST database with the At-γTIP and At-RD28 amino acid sequences led to the identification of more than 170 entries with similarity to plant TIPs and PIPs. Multiple sequence analysis and inspection of overlapping cDNA regions resulted in the identification of cDNA clones of individual genes representing

Figure 1. Nucleotide and deduced amino acid sequence of PIP3 cDNA (A) and NLM1 cDNA (B). The full-length clone of PIP3 was obtained from the Arabidopsis Biological Resource Center (clone 93P21 T7). The numbering refers to the nucleotide sequence. The cDNA nucleotide sequences will appear in the GenBank, EMBL, and DDBJ databases under the accession numbers U78297 (PIP3) and Y07625 (NLM1).
22 different gene transcripts (including PIP3). By comparing nonredundant sequence tags with databases, Cooke et al. (1996) identified a partial cDNA clone, VBVGF07, from Arabidopsis as a close homolog to a MIP from rice (Oryza sativa L.). By repeating the database search, we found that the proteins encoded by this Arabidopsis sequence and the rice cDNA sequence are also very similar to Gm-NOD26 but are only distantly related to other Arabidopsis PIPs and TIPs. Since no full-length clone identical to VBVGF07 could be identified in the EST databases, we obtained a full-length cDNA clone from Arabidopsis using the RACE method and provisionally named it At-NLM1 (Fig. 1B). The deduced amino acid sequence shows 64.2% identity to NOD26 from soybean (Fortin et al., 1987), but only 31.1 and 29.9% identity to PIP3 (this study) and γTIP (Hofte et al., 1992), respectively.

An amino acid sequence alignment of NLM1 to known plant MIPs (Fig. 3) shows a hydrophilic N-terminal tail with approximately the same number of negative and positive charges as found in PIPs. The C-terminal tail, however, has two negative charges and is more similar to plant TIPs. Unique to the NOD26/NLM1 subgroup is a highly conserved Cys residue in or close to the end of the first putative transmembrane domain. It is interesting that the NPA motifs, characteristic for the MIP family, are not fully conserved in NLM1. The second NPA motif reads NPG in NLM1 (identical in both independently isolated NLM1 clones) and is not found in any of the other published MIPs. We injected At-NLM1 cRNA into X. laevis oocytes to test whether the protein has water-channel activity. The oocyte swelling assays clearly showed that At-NLM1 encodes an aquaporin that is as active as other aquaporins, although we cannot rule out other transport activities such as ion transport. The presence of At-NLM1 in X. laevis oocytes increased the value of the osmotic permeability coefficient of the plasma membrane from $3.0 \times 10^{-3}$ to $16.4 \times 10^{-3}$ cm s$^{-1}$. To find additional MIPs, we carried out RT-PCR with oligonucleotides designed for PIPs and TIPs (see “Materials and Methods”) with mRNA from A. thaliana plants. This search yielded only one new MIP, clone pCR55.

On a dendrogram, all of these MIP cDNAs cluster into three groups: the tonoplast aquaporins (e.g. At-αTIP, At-γTIP, and At-δTIP) form one large group and the plasma membrane aquaporins (e.g. At-PIP1a, At-PIP1b, At-PIP1c, and others) form a second large group (Fig. 4). The third group contains one clone from the rice (At-γTIP 44) and the At-NLM1 clone obtained from Arabidopsis, which is distantly related to the other MIPs. The At-NLM1 clone could not be identified on the basis of similarity to any published sequence. However, it is clearly not a MIP and may represent a novel type of membrane protein.
Figure 4. Amino acid sequence relationship of the 23 members of the *A. thaliana* MIP gene family. Continuous lines connect those MIPs for which the complete cDNA sequence has been determined. Dashed lines connect only partially sequenced MIP cDNAs. Boxes indicate the MIPs that were identified as aquaporins. The bars on the right stand for the (putative) subcellular localization. The subcellular localization of At-NLM1 is not known. PIPla, PIP2a, and PIP2b (Kammerloher et al., 1994); PIP1b is identical to TMP-A (Shagan and Bar-Zvi, 1993) and Athf2 (Kaldenhoff et al., 1993); PIP1c (Kammerloher et al., 1994) is identical to TMP-B (Shagan et al., 1993); TMP-C (Kinoshita et al., 1994); pCR55, NLM1, and PIP3 (this study); RD28 (Yamaguchi-Shinozaki et al., 1992); α-TIP and γ-TIP (Höfte et al., 1992); β-TIP (H. Höfte and M.J. Chrispeels, unpublished data); δ-TIP (Daniels et al., 1996); T04164, T22419, T76151, Z30833, TT21060, T20432, Z18142, T21742, and T22237 (GenBank accession nos.). The experiment was performed three times and the results varied less than 10%.

At-PIP2a, At-PIP2b, and At-RD28) form the other large group, whereas At-NLM1 is on a branch by itself (Fig. 4). An amino acid sequence alignment of full-length TIPs and PIPs shows that TIPs exhibit three gaps compared with PIPs in hydrophilic loops between transmembrane domains 3/4, 4/5, and 5/6, that TIPs are more heterogeneous than PIPs, and that TIPs have a shorter N-terminal cytoplasmic tail.

Relative Abundance of MIP mRNA in Different Organs

Aquaporins enhance the hydraulic conductivity of membranes and may be involved in long-distance and short-distance transcellular water flows, as well as in the osmotic adjustments between the vacuole and the cytosol or the cytosol and the apoplast. We do not yet know whether all of the 23 MIP sequences identified to date are aquaporins, but the 10 Arabidopsis proteins tested so far have proven water-channel activity. It is our assumption that most, if not all, of the others will also turn out to be water channels.

To identify the most abundantly expressed sequences, we measured the amount of individual mRNAs of 22 different genes by slot-blot analysis (Fig. 5). With this method, the entire mRNA population of one tissue is reverse-transcribed and labeled as cDNA and used as the probe under high stringency. Since the specific target DNA is present in molar excess as slots on a membrane, we can assume that binding of a specific probe is directly correlated with the concentration of the original mRNA. Therefore, abundant mRNAs result in intense signals after slot-blot hybridization.

To compare signals from the slot blots of different organs, all signals were normalized against the β-tubulin-4 mRNA (Marks et al., 1987). This β-tubulin mRNA is expressed in all organs, although not at exactly the same level. Therefore, the relative amount of β-tubulin-4 mRNA was determined by quantitative PCR using an internal standard and was set at 1 for leaves, 0.539 for roots, 0.771 for bolt tissue, and 0.387 for reproductive organs (flowers and siliques). The most abundant mRNAs (about 10- to 20-fold of β-tubulin-4 mRNA) are those for the PIP1 and PIP2 plasma membrane aquaporins. This result is in agreement with the work of Kammerloher et al. (1994), who isolated PIP1 and PIP2 cDNAs by using antibodies against abundant plasma membrane proteins of roots. At-PIP3, which also encodes an active aquaporin, is very abundant in flowers and siliques of Arabidopsis. Abundant TIP isoforms are At-γTIP and At-δTIP, already identified aquaporins (Maurel et al., 1993; Daniels et al., 1996), but the mRNA of four other MIPs (Z30833, T21060, Z18142, and
T22237) are equally abundant. Especially in leaves, T22237 seems to be the most abundant TIP. Most of the other MIP transcripts were found in at least two, but often in all, organs included in this survey.

The level of mRNA for At-PI13 in tissue culture cells was compared with that of all known and fully sequenced aquaporins using the same slot-blot method. It was found to be more highly expressed in cultured cells than in any of the eight fully sequenced and previously known Arabidopsis aquaporins. Transcripts for some MIPs, however, could not be detected using slot-blot analysis. This is probably due to their low transcription rate, an assumption that is supported by the observation that the respective cDNA sequences were found only a few times in the EST database (e.g., only one entry for At-<i>α</i>TIP or At-<i>β</i>TIP), whereas abundant MIPs were found up to 50 times. However, each of the rare MIP cDNAs was cloned independently several times, indicating that they are not sequencing artifacts but low-abundance transcripts of MIP genes.

**Quantitative PCR of At-NL1M1 Expression**

At-NL1M1 is a newly identified aquaporin that is expressed at relatively low rates in Arabidopsis. Since At-NL1M1 is the sole member of a third subgroup of Arabidopsis MIPs, we were interested in the quantitative data on the expression of this particular gene. Because of the relatively low expression rate of At-NL1M1, we performed quantitative PCR analysis for At-NL1M1 and compared it with At-<i>γ</i>TIP, At-<i>δ</i>TIP, and At-RD28 (Fig. 6), which have been investigated in detail in our laboratory (Ludevid et al., 1992; Daniels et al., 1994, 1996). ABA, NaCl, etiolation, and desiccation did not have much influence on the levels of mRNA for At-<i>γ</i>TIP, At-<i>δ</i>TIP, and At-RD28. The application of osmotic stress by adding 0.3 M mannitol to the growth medium dramatically reduced the mRNA transcript of these three genes. In contrast, transcription of At-NL1M1 is strongly affected by ABA, NaCl, and etiolation but not very much by growth on mannitol (Fig. 6). The amount of an amplified control RNA (AW109) used as an internal standard was nearly identical in all of the experiments, indicating uniform PCR efficiency in all samples.

**DISCUSSION**

**Why Are There So Many Aquaporins?**

The recent discovery that the membranes of plant cells contain aquaporins (water-channel proteins) is shedding new light on the processes of transmembrane and transcellular water movement. Indeed, this discovery has changed how we envisage the control of water transport in plants. Aquaporins belong to the family of MIP proteins, and we report here that Arabidopsis contains no less than 23 expressed MIP genes, eight of which had previously been identified as active aquaporins of the plasma membrane and the tonoplast. The presence of so many actively transcribed MIP genes, most of which could not be detected by aquaporins, compels us to ask the question: why is the Arabidopsis

Why does a plant have so many aquaporins? (b) Are there functional differences between the proteins themselves or in the manner of their regulation? Aquaporins are known to occur in the tonoplast and the plasma membrane, and the derived amino acid sequences of new MIPs allowed us to postulate their subcellular location based on comparison with the amino acids in which the location has already been determined. Indeed, most of the Arabidopsis
sequences fall into two large groups: TIPs and PIPs. At least one reason why there are so many MIPs is that half are probably in the plasma membrane and the other half are in the tonoplast.

**NLM1 Is the First Member of the Third Subgroup of Aquaporins in Arabidopsis**

One advantage of EST random sequencing is that more distantly related cDNAs of a gene family can be identified by computational methods. This approach circumvents the problems encountered with in vitro experiments, such as library screening with labeled probes or RT-PCR with degenerated primers; these approaches detect closely related sequences more easily than distant ones. Computational sequence comparison of EST clones performed by Cooke et al. (1996) identified one Arabidopsis gene to be closely related to an rMIP1 from rice (Liu et al., 1994). The Arabidopsis homolog NLM1 clusters together with rMIP1, NOD26 (Fortin et al., 1987), and a third homolog from Persian tobacco (*Nicotiana alata*; GenBank accession no. U20490) into a distinct subgroup within the MIP dendrogram (Fig. 7). The *Escherichia coli* AQPZ (Calamita et al., 1995) lies on the same branch of this tree. Most of the sequence characteristics of the third subgroup (see above) are also found in Ec-AQPZ, including the highly conserved Cys residue at the end of the first hydrophobic domain. The identification of a third subgroup of MIP genes in four plant species (one monocotyledon and three dicotyledons) allows us to ask new questions. From the growing number of reports of TIPs and PIPs in plants we can assume that each plant species has these proteins, but is this also true for the NLM1 subgroup? What is the transport specificity of these proteins? NOD26 has been reported to be an ion channel when reconstituted in liposomes, and more recent reports have shown its aquaporin activity (Chandy et al., 1996; Rivers et al., 1996). NLM1 is also a water-channel protein, but we do not know whether this is its sole transport activity. The rat AQP protein transports not only water but also urea and glycerol, and protein sequence comparisons of this human isolog show that it lies outside of the aquaporin group and is more closely related to Ec-GLPF, the glycerol facilitator from *E. coli* (Fig. 7). We do not know in which subcellular compartment NLM1 is located and we cannot rule out the possibility that it may be located in the ER, the chloroplast envelope, or some other intracellular membrane. Its location may give us additional clues regarding the physiological role of the third subgroup of plant aquaporins.

Transmembrane water flow is likely to be important in many physiological processes, including the transpiration stream, phloem loading, stomatal opening, pulvinar movement, and osmotic adjustments. It would be to the plant’s advantage to be able to regulate such water flow by modulating the abundance or the activity of water channels. This may be the second reason why so many different aquaporins are needed. Evidence is accumulating that individual aquaporins are differentially expressed in certain cell types, tissues, or organs (see the introduction). A similar situation is encountered with plasma membrane H⁺-ATPases: Arabidopsis has at least 10 expressed AHA genes and some of those are uniquely expressed in certain cell types. For example, AHA10 is expressed primarily in developing seeds (Harper et al., 1994), AHA3 is expressed in the phloem of the root and the shoot (DeWitt et al., 1991), and AHA2 is expressed in the root epidermis (J. Harper, personal communication).

**All MIPs Are Not Equally Expressed**

To determine the expression levels of all MIPs simultaneously, we used an RNA-blot technique in which the plasmid DNA is attached to the membrane and the RNA is labeled. The slot-blot technique, as used here, is a useful method to measure relative amounts of many mRNAs simultaneously, but it has certain limitations. First, different mRNAs have different nucleotide compositions and therefore the amount of digoxigenin incorporated during the labeling reaction varies. Second, the plasmids fixed on the nylon membrane were mostly full-length clones but varied from about 850 to 1100 bp in length; the signals obtained with clone pCR55 were multiplied by 3 because this cDNA fragment covers only one-third of the putative full-length cDNA. When making a comparison between different experiments, these limitations do not need to be
taken into account because they apply to all experiments in the same manner. It is more important to normalize the intensities of each slot so that different blots can be compared. For this, we determined the relative intensity of the β-tubulin-4 mRNA in all of the organs and then adjusted the overall measurement of each slot according to these values. The measurement of gene expression on all genes simultaneously makes it possible to distinguish, in a first approach, abundant from nonabundant MIP mRNAs. Under the stringent hybridization conditions used here (hybridization in 50% formamide; washing at 68°C in 0.1X SSC), there is unlikely to be cross-hybridization between the different genes. The highly divergent signal intensities confirm this assumption. These signal intensities are also related to the number of times a sequence appears in the EST database: very weakly expressed genes are represented by very few independent accessions.

Sixteen of the 23 identified MIP transcripts represent more than 98% of aquaporin-related messages in Arabidopsis. The remaining 2% are encoded by 7 other MIP genes that are present at barely detectable levels in plants grown under laboratory conditions. Some MIPs may have important roles in specific nonabundant cell types or may be up-regulated under certain environmental conditions. Mesophyll cells are the most abundant in leaves, and this may be where the highly expressed PIP1 subgroup is mainly found. Immunocytochemical observations show that PIP1 is present in the multivesicular bodies and plasma membranes of mesophyll cells (Robinson et al., 1996). In roots the PIP1 and PIP2 subgroups are both abundant, but this is not the case in shoots, where PIP2 sequences are poorly represented. Are these differences in expression related to different cellular controls of water flow? Leaves contain large intercellular spaces filled with water-saturated air. Such a compartment is absent in roots. As a result of such morphological differences, we can imagine that different aquaporins, the activity of which can be regulated in different ways, could be required for water flow through the root cortex into the vascular system and out of the leaves.

Aquaporin Activity May Be Differentially Regulated

Although the hydraulic conductivity of tissues could be regulated by changing the level of specific aquaporins, alternatively, regulation could be exerted posttranscriptionally by changing the activity of aquaporins. The seed-specific aquaporin α-TIP has been shown to be regulated by phosphorylation (Maurel et al., 1995). Regulation by phosphorylation has also been proposed for GmNOD26 (Lee et al., 1995) and for a spinach (Spinacia oleracea L.) plasma membrane MIP (Johansson et al., 1996), recently shown to be an aquaporin (I. Johansson and M.J. Chrispeels, unpublished data). Protein phosphorylation is already known to be an important regulator of drought stress, and drought stress signaling probably involves a MAP-kinase cascade (Urao et al., 1994; Jonak et al., 1996; Mizoguchi et al., 1996; Shinozaki and Yamaguchi-Shinozaki, 1996). Thus, it will be most interesting to find out whether this cascade has aquaporins as a downstream target.

Whether the different aquaporins and other MIPs differ in their transport properties is not known. Individual aquaporins may differ in their water-transport efficiency, but this property is difficult to measure in oocytes, because it requires us to know that the same number of active molecules are present at the plasma membrane of two oocytes injected with mRNAs for two different aquaporins. Alternatively, some aquaporins may transport solute molecules in addition to water. This appears to be the case for AQPD, a mammalian aquaporin that also transports glycerol.

ACKNOWLEDGMENTS

We thank François Chaumont for his help with some of the figures, Lyn Alkan for the production of the typescript, and Christopher Maurel for his careful reading of the manuscript and his many valuable suggestions.

Received December 23, 1996; accepted May 14, 1997.

LITERATURE CITED


Urao T, Katagiri T, Mizoguchi T, Yamaguchi-Shinozaki K, Hayashida N, Shinozaki K (1994) Two genes that encode Ca^{2+}-dependent protein kinases are induced by drought and high-salt stresses in Arabidopsis thaliana. Mol Gen Genet 244: 331–340


