The Plant Cell, Vol. 2, 525–532, June 1990 © 1990 American Society of Plant Physiologists

An Intrinsic Tonoplast Protein of Protein Storage Vacuoles in Seeds Is Structurally Related to a Bacterial Solute Transporter (GlpF)

Kenneth D. Johnson,[1] Herman Höfte, and Maarten J. Chrispeels[2]

Department of Biology and Center for Molecular Genetics, University of California, San Diego, La Jolla, California 92093-0116

The tonoplast mediates the transport of various ions and metabolites between the vacuole and cytosol by mechanisms that remain to be elucidated at the molecular level. The primary structure of only one tonoplast protein, the H^+-ATPase, has been reported to date. Here we report the primary structure of tonoplast intrinsic protein (TIP), a 27-kilodalton intrinsic membrane protein that occurs widely in the tonoplasts of the protein storage vacuoles (protein bodies) of seeds [Johnson, K.D., et al. (1989). Plant Physiol. 91, 1006–1013]. Hydropathy plots and secondary structure analysis of the polypeptide predict six membrane-spanning domains connected by short loops and hydrophilic, cytoplasmically oriented N- and C-terminal regions. TIP displays significant homology with several other membrane proteins from diverse sources: major intrinsic polypeptide from bovine lens fiber plasma membrane; NOD 26, a peribacteroid membrane protein in the nitrogen-fixing root nodules of soybean; and interestingly, GlpF, the glycerol facilitator transport protein in the cytoplasmic membrane of *Escherichia coli*. Based on the homology between TIP and GlpF and the knowledge that the protein storage vacuolar membrane and the peribacteroid membrane are active in solute transport, we propose that TIP transports small metabolites between the storage vacuoles and cytoplasm of seed storage tissues.

INTRODUCTION

The most conspicuous of plant organelles, the vacuole, is a multifunctional, polymorphic compartment. During differentiation of the parenchyma cell, many small vacuoles fuse into a single, large, central vacuole that serves as a reservoir for cytoplasmic homeostasis and contributes to osmotic maintenance of cell turgor. With its complement of acid hydrolases and its acidic pH, the vacuole also functions as the lysosomal compartment of the cell. The sequestration of ions and metabolites in vacuoles has an important dynamic dimension, as illustrated by dramatic seasonal changes in vacuolar protein contents of vegetative storage tissues, diurnal fluctuations of organic acids in mesophyll vacuoles of Crassulacean acid metabolism plants, and the rapid changes occurring in guard cells during stomatal movement. This dynamic aspect of vacuolar compartmentation implies that the tonoplast, the vacuole-delimiting membrane, is actively engaged in the transport of ions and metabolites. Indeed, isolated vacuoles have been shown to possess transport systems for a variety of substances (reviewed in Boller and Wiemken, 1986). With the exception of the H^+-ATPase (Nelson and Taiz, 1989), however, we have very few of the molecular details of the proteins that carry out the transport functions of the tonoplast (Barbier-Brygoo et al., 1986).

In the storage tissues of developing seeds, the vacuolar system becomes specialized for the storage of proteins and other metabolites. Soon after the central vacuole begins to accumulate reserve protein, it gradually breaks up into thousands of smaller, densely filled bodies called protein storage vacuoles (PSV) or protein bodies. These PSVs store not only storage proteins and plant defense proteins (e.g., lectins and enzyme inhibitors), but also minerals (in the form of phytin) and metabolites such as sucrose and other oligosaccharides (Muller and Jacks, 1983; Plant and Moore, 1983).

As part of our goal toward understanding the molecular basis of vacuolar compartmentation and storage in seeds, we have become interested in the molecular mechanisms of metabolite transport across the tonoplast of the PSV. It seems probable that vacuolar import of metabolites and ions occurs during PSV ontogeny, and events following seed germination must trigger a massive export of metabolites, especially amino acids, sugars, and the products of phytin digestion from the storage vacuoles. With this in mind, we (Johnson et al., 1989) recently purified and characterized the most abundant intrinsic membrane protein of the PSV in bean cotyledons (Pusztai et al., 1979; Maeder and Chrispeels, 1984). With a specific antiserum,

[1] On leave from the Department of Biology and Molecular Biology Institute, San Diego State University, San Diego, CA 92182-0057.
[2] To whom correspondence should be addressed.

we demonstrated that this protein (previously called TP25, but now designated TIP) is seed specific and highly conserved among seed plants—TIP analogs were found in the seeds of all monocots and dicots tested. In this paper we report the sequence of a full-length TIP cDNA as well as the partial sequence of a closely related cDNA, both obtained from the common bean *Phaseolus vulgaris*. Analysis of the deduced amino acid sequence reveals that TIP is a hydrophobic polypeptide with six putative membrane-spanning domains. TIP bears a striking resemblance to several other intrinsic membrane proteins from diverse sources, including the major intrinsic protein (MIP) from bovine lens fiber plasma membrane, a plant-encoded peribacteroid membrane protein (NOD 26) in soybean root nodules, and the glycerol facilitator (GlpF) located in the cytoplasmic membrane of *Escherichia coli*. These homologies together with TIP's specific pattern of expression in seeds lead us to suggest that TIP is a channel-type transport protein in the tonoplast of protein storage vacuoles.

RESULTS

Nucleotide and Deduced Amino Acid Sequence of cDNAs Encoding TIP

Previous work from our laboratory has shown that TIP in the common bean (*P. vulgaris* cv Greensleeves) is expressed during the mid-maturation stage of seed development (Johnson et al., 1989). Thus, we prepared a cDNA expression library in λgt11 using poly(A)$^+$ RNA from mid-mature bean cotyledons and screened the library with antiserum against TIP. Twelve antibody-positive clones were isolated but none had full-length inserts. The longest of these, denoted λTP1, was sequenced. Analysis of the sequence data revealed a truncated 307-bp open reading frame and a 257-bp 3′-noncoding region followed by the polyA tail. The partial deduced amino acid sequence (120 residues) showed agreement with the amino acid composition of TIP (Pusztai et al., 1979), and a hydropathy plot (Kyte and Doolittle, 1982) revealed two membrane-spanning, hydrophobic domains expected for this intrinsic membrane protein (data not shown).

By using bean seed poly(A)$^+$ RNA and oligonucleotides complementary to sequences in the middle of the open reading frame of λTP1, we employed the RACE (rapid amplification of cDNA ends) procedure of Frohman et al. (1988) to synthesize and amplify cDNA toward the 5′ end. The RACE cDNA products were cloned into pUC 19, and two recombinants with inserts that hybridized differentially to λTP1 were selected and sequenced. The sequencing data revealed two different but related cDNA molecules (designated rTPα and rTPβ). The nucleotide and deduced amino acid sequences of the α and β (partial) TIP cDNAs are shown in Figure 1. The full-length α cDNA sequence is the combination of αTP1 (3′ end) and rTPα (5′ end) sequences, which overlap by 125 bp.

The αTIP cDNA consists of 1208 bp upstream of the polyA tail, which includes a 183-bp 5′ leader sequence, followed by 768 bp of open reading frame encoding 256 amino acids, and, finally, a 257-bp 3′-noncoding region. The calculated molecular mass of the derived polypeptide is 27,092 D, which compares favorably to the estimated M_r of 25 kD for TIP on SDS-PAGE (Johnson et al., 1989). The polyA addition site lies 23 bp downstream from the AATAAA polyadenylation signal. The designated translation start site is the first ATG encountered downstream from the 5′ end. To confirm that the αTIP cDNA encodes the vacuolar membrane protein described earlier (Johnson et al., 1989), we purified the protein, subjected it to trypsin digestion followed by peptide separation on HPLC, and determined the amino acid sequence of one of the tryptic

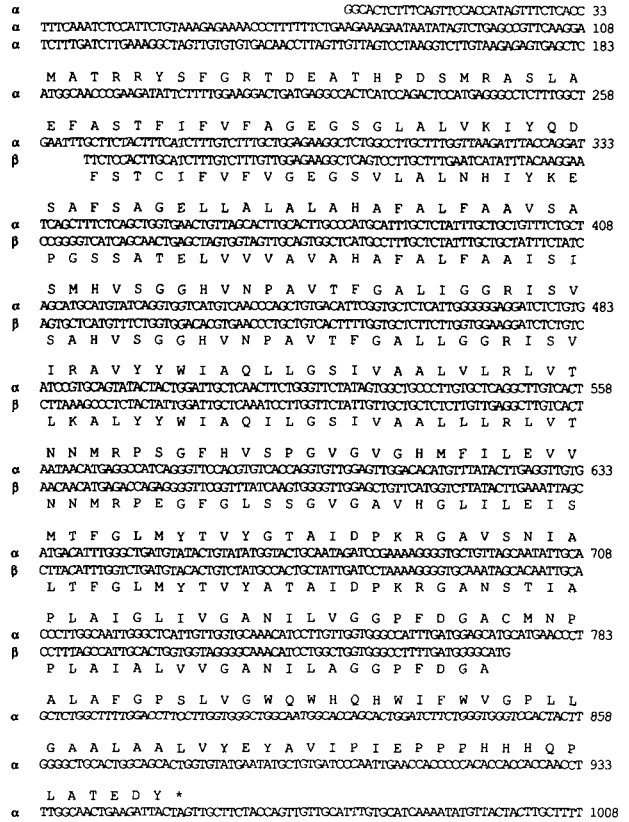

Figure 1. The cDNA and Deduced Amino Acid Sequences of αTIP (Complete) and βTIP (Partial).

The numbering refers to the nucleotide sequence of αTIP. The putative polyadenylation signal is underlined. The polyA tract is not shown.

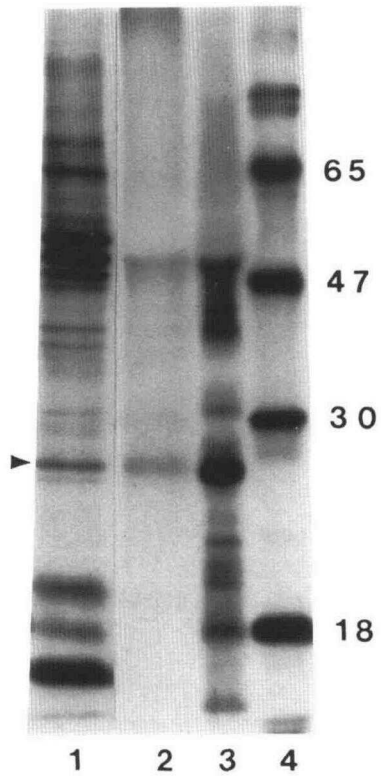

Figure 2. SDS-PAGE Fluorogram of TIP from Isolated Microsomes and in Vitro Translation Products.

^{35}S-methionine-labeled polypeptides from bean cotyledon microsomes before (lane 1) or after (lane 2) immunoprecipitation with TIP antiserum. Lane 3, In vitro translation products after immunoprecipitation; lane 4, molecular weight standards. Arrowhead indicates position of TIP.

fragments. The sequence obtained—TDEATHPDSMR—was also found near the N terminus (residues 11 to 21; see Figure 1) of the deduced amino acid sequence. Finally, the derived sequence corresponded closely with amino acid composition data on the purified protein (Pusztai et al., 1979).

The rTPβ clone yielded a truncated cDNA sequence for βTIP. It was colinear with the αTIP sequence and displayed 75% sequence identity at the amino acid level. Most of the differences represent conservative amino acid replacements, but the divergence suggests that the α and β cDNAs represent different genes. In this regard, DNA gel blot analysis showed that several of the λgt11 cDNAs hybridized weakly to labeled λTP1 (K.D. Johnson and M.J. Chrispeels, unpublished data). The nucleotide sequences of these cDNAs all showed homology to λTP1 in their truncated open reading frames, but the 3′-noncoding regions were entirely different. This lends support to the idea that at least two related TIP genes are expressed in developing bean seeds. As the primer-extended β cDNA sequence did not overlap with the cDNA sequences found in this second class of λgt11 clones, we do not know whether they represent the same TIP (β) gene.

Structural Characteristics

Earlier work from our laboratory showed that TIP is synthesized on the rough endoplasmic reticulum (Maeder and Chrispeels, 1984). A comparison of the mobilities on SDS-PAGE gels of TIP immunoselected from in vivo labeled and in vitro translated polypeptides indicated that they have similar M_r values as shown in Figure 2. Because TIP is not a glycoprotein (Pusztai et al., 1979; Maeder and Chrispeels, 1984), the similar M_r values indicate the absence of a cleaved signal peptide. This conclusion is also supported by the fact that the deduced sequence of the N-terminal 21 amino acids does not contain a hydrophobic domain typical of cleaved endoplasmic reticulum signal peptides; rather, it is quite hydrophilic. Furthermore, most cleavable signal sequences range between 13 and 30 residues in length (Verner and Schatz, 1988); because mature, vacuolar TIP contains the tryptic peptide corresponding to residues 11 to 21 in the derived amino acid sequence, it is unlikely that any cleavage event at the N terminus takes place.

The most striking structural characteristic of the TIP polypeptide can be seen in the hydropathy plot in Figure 3. There are six hydrophobic domains of sufficient length and hydrophobicity to span a lipid bilayer. Analyses of the amino acid sequence with secondary structure algorithms (Chou and Fasman, 1974; Garnier et al., 1978) predicted that all the putative transmembrane domains are predom-

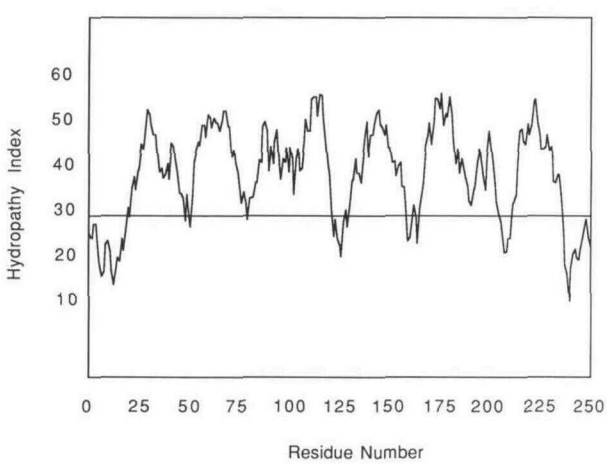

Figure 3. Hydropathy Plot of αTIP.

The computer-generated plot was carried out using a moving window of 9 residues (Kyte and Doolittle, 1982).

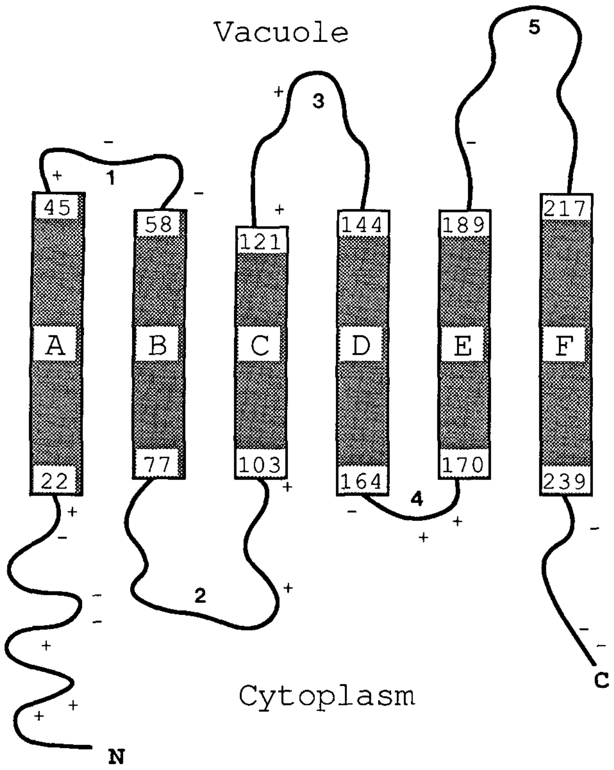

Figure 4. Topological Model of αTIP.

The bars (A to F) represent putative transmembrane domains and the numbers (1 to 5) indicate consecutive extramembrane loops. Charged amino acid positions are indicated for the extramembrane regions only. Domains A, C, and F each have single glutamic acid residues as the only charged residues in the membrane-spanning segments.

inated by either helix or sheet configurations. Based on these data and homologies with other known polytopic membrane proteins (see below), we believe that the topology of TIP in the vacuolar membrane can be represented by the model shown in Figure 4. Each transmembrane domain consists of 19 to 24 residues and together accounts for half of the total amino acids. The connecting loops between the membrane-spanning segments are relatively short, hydrophilic stretches; all have either charged amino acids, histidines, or helix-breakers at both ends. The secondary structure algorithms indicate a high probability of random coil and turns in these loops and in the N- and C-terminal domains. Most (8 of 11) of the positive-charged amino acids are on the cytoplasmic side of the membrane in accordance with the "positive-inside" rule that holds for most polytopic membrane proteins (von Heijne and Gavel, 1988). Although these theoretical considerations are consistent with the membrane orientation model shown, other interpretations are possible. Moreover, there is no direct evidence for this model at the present time.

Homologous Proteins

A search of the PIR protein data base and sequences reported in the literature yielded three proteins whose deduced amino acid sequences show considerable homology to TIP: the *E. coli* GlpF (Muramatsu and Mizuno, 1989); MIP in the junctional membranes of bovine lens fiber cells (Gorin et al., 1984); and nodulin-26 (NOD 26), a peribacteroid membrane protein in soybean nodules (Sandal and Marcker, 1988). The sequence homologies among GlpF, NOD 26, and MIP were reported recently by Baker and Saier (1990). Figure 5 shows the amino acid sequence alignment for TIP and GlpF, which indicates 32% sequence identity over a 213-residue span. The regions of greatest sequence homology appear in the transmembrane domains of TIP (underlined in Figure 5) and especially in two of TIP's putative extramembrane loops (loops 2 and 5 in Figure 4). Interestingly, MIP and NOD 26 also share a high level of sequence identity with TIP and GlpF in these two extramembrane regions. These highly homologous segments are shown for all four proteins in Figure 6, in which the identities relative to GlpF are emphasized. The potential significance of these highly conserved loop domains is heightened by the fact that they exhibit significant internal

```
TIP                                           M A T R R Y S F G R T D   12

glpF  M S Q T S T L K G Q C I A E F L G T G L L I F F G V G C V A A L   32
                              . : : :     :   . . . :   :   :       : :
TIP   E A T H P D S M R A S L A E F A S T F I F V F A G E G S G L A L   44

glpF    K V A G A S   F G Q W E I S V I W G L G V A M A I Y L T A G     61
        : .       :   :       : .   .       :     : . : .
TIP   V K I Y Q D S A F S A G E L   L A L A L A H A F A L F A A V S A   75

glpF        V S G A H L N P A V T I   A L W L F A C F D K R K V I P F   89
            : : :   : . : : : : : .   : : .               :     .   .
TIP   S M H V S G G H V N P A V T F G A L I G G R I S V I R   A V Y Y   106

glpF  I V S Q V A G A F C A A A L V Y G L Y Y N L F F D F E Q T H H I   121
      . .   : .   :   .     : : : :     :   . :
TIP   W I A Q L L G S I   V A A L V L R L V T N N                       127

glpF  V R G S V E S V D L A G T F S T Y P N P H I N F V Q A F A V E M   153
        :   :       :                       :   .   .       :   . : .
TIP   M R P S G F H V S                     P G V G V G H M F I L E V   149

glpF  V I T A I L M G L I L A L T V D G N G V P R G P L A P L L I G L   185
      : . :     : :     .         . :                 . : : :   : : :
TIP   V M T F G L M Y T V Y G T A I D P K R G A V S Y I A P L A I G L   181

glpF  L I A V I G A S M G P L T G F A M N P A R D F G P K V F A W L A   217
      . .               : : .   :     : : : :     : : :   . .   :
TIP   I V G A N I L V G G P F D G A C M N P A L A F G P S L V G W Q W   213

glpF  G W G N V A F T G G R D I P Y F L V P L F G P I V G A I V G A F   249
              .   .   :                   : : . :     .   : . :       .
TIP   H Q H W I F W V G ................. P L L G A A L A A L V Y E Y   236

glpF  A Y R K L I G R H L P C D I C V V E E K E T T T P S E Q K A S L * 281
      :       .                         : .
TIP   A V I P I E P P P H H H Q P L A T E D Y *                         256
```

Figure 5. Sequence Comparison of TIP and GlpF.

Double and single dots correspond to identities and conservative replacements, respectively; underlined amino acids indicate the six putative membrane-spanning domains of TIP, as shown in Figure 4.

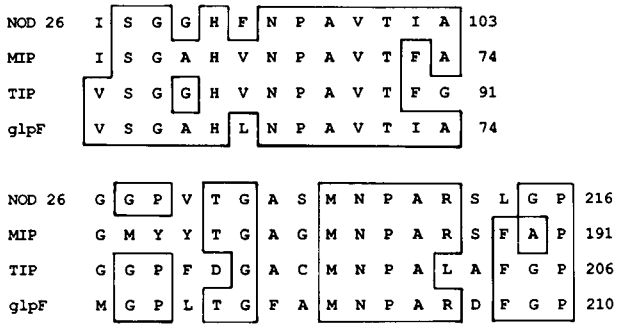

Figure 6. Sequence Comparisons of Two Highly Conserved Domains in NOD 26, MIP, TIP, and GlpF.

The boxed residues indicate identities with GlpF amino acids. In reference to TIP structure (see Figure 4), the upper panel domain derives from cytoplasmic loop 2; the lower panel domain is found in vacuolar loop 5.

similarity to one another. The two 16-residue spans beginning at positions 81 and 196 (see Figure 1) of TIP contain 7 identical residues and 5 conservative replacements.

The level of amino acid sequence identity between MIP and NOD 26 is 35% (Sandal and Marcker, 1988). The sequence alignments for TIP, MIP, and NOD 26 are illustrated in Figure 7 and indicate that TIP shares 35% to 36% sequence identity with both MIP and NOD 26.

DISCUSSION

The tonoplast plays an important role in the metabolic compartmentation of the plant cell. In seeds, metabolite fluxes across the tonoplast, the limiting membrane of the protein storage vacuole (protein body), occur both during seed development and during seedling growth. Our interest in understanding vacuolar compartmentation in seeds and the molecular mechanisms of metabolite flux across the tonoplast led us to characterize TIP, an abundant, seed-specific, and highly conserved intrinsic protein of the protein storage vacuole. The hydropathy plot of the deduced amino acid sequence indicates that TIP is a 27-kD protein with six membrane-spanning domains. Most remarkable is the sequence homology between TIP and three other intrinsic membrane proteins: GlpF, MIP, and NOD 26. Do these proteins have a common evolutionary origin and a common function?

Structure of TIP and Orientation in the Tonoplast

A model of the orientation of TIP in the tonoplast is shown in Figure 4, with cytosolic N and C termini. The model is based in part on TIP's sequence homology with MIP and the extensive topographical studies carried out on MIP (Paul and Goodenough, 1983; Gorin et al., 1984). Like MIP (Gorin et al., 1984), TIP apparently lacks a cleaved signal peptide. Insertion into the endoplasmic reticulum membrane is presumably specified by the first transmembrane domain. The transmembrane segments of multispanning proteins have been proposed to function alternatively as signals for membrane insertion (signal-anchor sequences) and stop-transfer (Blobel, 1980; Friedlander and Blobel, 1985), but only the most N-terminal signal-anchor sequence interacts with the signal-recognition particle (Rapoport, 1985; Wessels and Spiess, 1988). Thus, the orientation of the first transmembrane domain dictates the transmembrane topology of the entire multispanning polypeptide (i.e., a multispanning protein with a cytosol-facing N terminus and six transmembrane segments would have its C terminus oriented toward the cytosol as well, as depicted in Figure 4 for TIP). But what determines the orientation of the first signal-anchor sequence? A recent analysis of many eukaryotic multispanning proteins synthesized on the rough endoplasmic reticulum indicates that the charge difference between the domains flanking the first internal signal-anchor sequence determines the transmembrane topology of these proteins, with the more positive flank oriented toward the cytosol (Hartmann et al., 1989). When the length of each flanking sequence was set at 15 residues, greater than 90% of the analyzed proteins with reasonably well-established membrane topologies fit this rule. A significant exception was bovine lens fiber MIP, significant because TIP also does not fit the rule when the 15-residue flank criterion is applied. If, however, the entire N-terminal flanking region (21 amino acids) of the first transmembrane domain is considered,

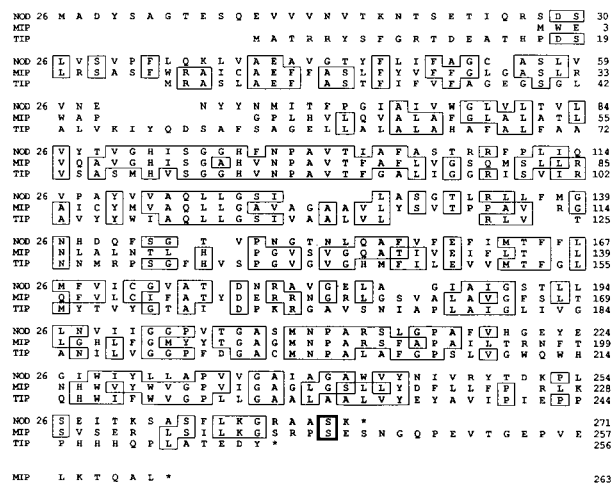

Figure 7. Sequence Comparison of NOD 26, MIP, and TIP.

The boxed residues indicate sequence identities. The alignment of NOD 26 to MIP is taken from Sandal and Marcker (1988).

the charge difference rule does predict a cytosolic orientation of the N terminus of TIP and, thus, a topology consistent with the model shown in Figure 4. Recent evidence shows that intervening domains not embedded in the membrane may also modulate the topology of a multispanning protein (Lipp et al., 1989). It is interesting in this respect that the domains with the greatest sequence identity between TIP, MIP, NOD 26, and GlpF are two external loops, one on each side of the membrane. Whether these or other loops contain topological information remains to be determined. Confirmation that the model presented in Figure 4 depicts the correct orientation of TIP must await topographical studies with TIP itself.

We have emphasized the structural similarities between TIP and three other proteins (MIP, NOD 26, and GlpF), whose sequences have been published. Preliminary reports in the literature indicate the existence of two other genes in this gene family. These are the so-called big brain protein in *Drosophila* (Rao et al., 1989) and the B7 cDNA from tobacco roots (Yamamoto et al., 1989). Thus, this group of homologous proteins has representatives in the bacteria, plants, insects, and vertebrates, suggesting that these proteins play some fundamental role in the cell. Our immunological studies (Johnson et al., 1989) point to a wide distribution of TIP among seed plants, indicating a protein of fundamental importance to the function of seeds.

TIP Is a Putative Metabolite Transporter in the Tonoplast

Of all these structurally related proteins, only the function of GlpF is known. This cytoplasmic membrane protein of *E. coli* is responsible for the channel-type transport of glycerol and other small molecules (Heller et al., 1980). On the basis of the observed sequence homologies, it is tempting to speculate that they all function as small metabolite channels. MIP is a junctional membrane protein that when added to preformed lipid films induces voltage-dependent channels (Zampighi et al., 1985). NOD 26 is a major component of the peribacteroid membrane, the tonoplast-like barrier that separates the bacteroids from the host cell's cytoplasm in the root nodules of soybeans. This membrane is the site of metabolite transport in both directions: transport of photosynthetic products from the cytoplasm to the bacteroids and transport of amino acids and/or other compounds containing reduced nitrogen from the bacteroids to the cytoplasm. On the basis of their sequence homologies with GlpF, NOD 26 and TIP may be involved in metabolite transport in their respective membranes. TIP's location in the tonoplast of seed reserve tissues makes it a candidate for metabolite transport into the vacuole during seed development and/or mobilization of metabolites (in part, the products of macromolecule digestion) during seedling growth.

Does TIP Function after Germination?

Seed development is characterized by the active accumulation of proteins and metabolites such as phytin in the protein storage vacuoles. These stored molecules are hydrolyzed early during seedling growth and the resulting small molecules are exported to the growing axis. In live storage tissues, this hydrolysis depends on the synthesis of hydrolytic enzymes, such as protease (Chrispeels et al., 1976) and phytase (Walker, 1974), and their transport to the vacuoles. Hydrolysis occurs within the vacuoles (Baumgartner et al., 1978; Nishimura and Beevers, 1979) and the hydrolysis products are presumed to be transported out of the vacuoles at that time. In seeds having storage tissues that are dead at seed maturity, such as the cereal endosperm, hydrolytic enzymes are synthesized and secreted by adjoining viable tissues (the aleurone or the scutellum). Hydrolysis and export of metabolites do not occur in such a controlled fashion because the cellular membranes have broken down. It may be significant in this regard that TIP is found in all seed tissues that are alive at seed maturity (axis, cotyledons, endosperm of dicots, aleurone cells), but not in tissues that lose viability during seed maturation (starchy endosperm of cereals) (Johnson et al., 1989). On the basis of that distribution, we suggest that TIP functions during germination and seedling growth rather than during seed development.

TIP Homologs in Other Tissues

Tobacco roots express a protein that is homologous to MIP and NOD 26 (Yamamoto et al., 1989) and, therefore, presumably to TIP as well. Using specific antiserum directed against TIP, we found TIP to be seed specific (Johnson et al., 1989). However, by screening an *Arabidopsis thaliana* genomic library with the bean TIP α cDNA, we have found three genes homologous to TIP, one of which is expressed in the vegetative organs of this plant (γTIP) (H. Höfte, K.D. Johnson, and M.J. Chrispeels, unpublished observation). This γTIP gene may be similar to the one described for tobacco roots and may represent a metabolite transporter in vegetative cells. We will make specific antibodies to γTIP and determine its subcellular location.

METHODS

Construction and Screening of the cDNA Library

Poly(A)$^+$ RNA was isolated from mid-maturation (20 days to 25 days after flowering) bean cotyledons as described (Prescott and Martin, 1987) and used to construct a cDNA library in λgt11 according to the manufacturer's instructions (Amersham). The

library was screened with a polyclonal antiserum raised against purified TIP (Johnson et al., 1989). Of 2×10^5 recombinants screened, 12 positive clones were isolated, six of which had inserts ranging in length from 400 bp to 580 bp. The longest of these (λTP1) was sequenced.

Extension and Amplification of the 5′ cDNA End

The RACE procedure of Frohman et al. (1988) was used with modification to synthesize and amplify TIP cDNA containing the 5′ end. First-strand cDNA was generated using bean cotyledon poly(A)$^+$ RNA, Moloney murine leukemia virus reverse transcriptase (Promega Biotec), and a 20-mer oligonucleotide primer complementary to an internal sequence in λTP1. This strand was polyadenylated using terminal nucleotidyl transferase (Bethesda Research Laboratories). Second-strand synthesis was carried out with Taq DNA polymerase (Bethesda Research Laboratories) and a primer containing an oligo(dT) stretch linked to a 19-base adapter with overlapping EcoRI, XhoI, and SalI sites. The double-stranded cDNA was then amplified over 35 polymerase chain reaction cycles using the 19-base adapter as the 5′ primer and, at the 3′ end, a 21-mer oligonucleotide complementary to a known sequence just downstream from an SphI site and upstream from the first-strand-synthesis primer site. The amplified cDNA appeared as a diffuse band (500 bp to 800 bp) on agarose gels and hybridized specifically to labeled λTP1. This cDNA was cut with EcoRI and SphI and cloned into pUC 19.

DNA Sequence Analysis

DNA inserts from antibody-selected λgt11 clones were inserted at the EcoRI site of pUC 19 according to standard procedures (Maniatis et al., 1982). These and the RACE clones were characterized by restriction enzyme analysis, and suitable length restriction fragments were subcloned into pUC 19 for sequencing. DNA sequence analysis was carried out by the dideoxy chain termination method on double-stranded DNA with Sequenase (United States Biochemical) according to the manufacturer's suggestions. In all cases both strands were sequenced.

In Vivo Labeling and Microsome Preparation

Mid-mature bean cotyledons were labeled for 4 hr with ^{35}S-methionine (10 μCi/cotyledon; Amersham) as described (Spencer et al., 1980). The cotyledons were rinsed and then homogenized in 0.1 M Tris-HCl (pH 7.8)/1 mM EDTA/12% (w/w) sucrose. The homogenate was centrifuged at 1000*g* for 10 min to remove cell walls and debris, and the supernatant was layered over a 16%/54% (w/w) discontinuous sucrose gradient and centrifuged for 2 hr at 100,000*g*. The microsomal membrane fraction at the 16%/54% sucrose interface was collected.

RNA Extraction, in Vitro Translation, Immunoprecipitation, and Electrophoresis

Total RNA from mid-mature bean cotyledons was extracted according to Prescott and Martin (1987). Poly(A)$^+$ RNA was purified using Hybond-mAP affinity paper (Amersham) according to the manufacturer's suggestions and translated in a rabbit reticulocyte lysate system (Jackson and Hunt, 1982) in the presence of ^{35}S-methionine.

TIP from in vivo labeled microsomes and in vitro translation products was immunoprecipitated with TIP antiserum after dissolving the membranes with detergent as described (Faye and Chrispeels, 1987). The total and immunoprecipitated products were analyzed by SDS-PAGE and fluorography (Bonner and Laskey, 1974).

Trypsin Digestion, Tryptic Peptide Separation, and Amino Acid Sequencing

TIP protein from the detergent-rich phase of Triton X-114-extracted protein body membranes from bean cotyledons (Johnson et al., 1989) was purified by SDS-PAGE, electroeluted into 1% SDS, and precipitated at −20°C in 90% acetone. After washing twice with 90% acetone and vacuum drying, the protein pellet (500 mg) was resuspended in 1 mL of 50 mM ammonium carbonate and heated at 100°C for 3 min. After cooling to room temperature, 50 mg of trypsin treated with tosylphenylalanyl chloromethyl ketone (Sigma) was added and the mixture was incubated for 2 hr at 37°C. Another 50 mg of trypsin was added followed by another 2 hr of incubation. The tryptic peptides were separated by reverse-phase HPLC and one of these peptides was sequenced by Edman degradation on an automated amino acid sequencer as described (Sturm et al., 1987).

ACKNOWLEDGMENTS

We thank Dr. Toni A. Voelker of Calgene, Inc. for his generous help with the RACE procedure, Dr. Brian W. Tague for assistance with the amino acid sequence comparisons, Dr. Milton H. Saier for pointing out to us that GlpF has sequence homology with MIP and NOD 26, and Roberta Bradford for the cover artwork. This work was supported by grants from the National Science Foundation (Cell Biology) and the United States Department of Agriculture (Competitive Research Grants Office). Herman Höfte is supported by a fellowship from the European Molecular Biology Organization.

Received March 13, 1990; revised April 3, 1990.

REFERENCES

Baker, M.E., and Saier, M.H., Jr. (1990). A common ancestor for bovine lens fiber major intrinsic protein, soybean nodulin-26 protein, and *E. coli* glycerol facilitator. Cell **60,** 185–186.

Barbier-Brygoo, H., Renaudin, J.-P., and Guern, J. (1986). The vacuolar membrane of plant cells: A newcomer in the field of biological membranes. Biochimie **160,** 417–425.

Baumgartner, B., Tokuyasu, K.T., and Chrispeels, M.J. (1978). Localization of vicilin peptidohydrolase in the cotyledons of mung bean seedlings by immuno-fluorescence microscopy. J. Cell Biol. **79,** 10–19.

Blobel, G. (1980). Intracellular protein topogenesis. Proc. Natl. Acad. Sci. USA **77,** 1496–1500.

Boller, T., and Wiemken, A. (1986). Dynamics of vacuolar compartmentation. Annu. Rev. Plant Physiol. **37,** 137–164.

Bonner, W.M., and Laskey, R.A. (1974). A film detection for tritium-labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46,** 83–88.

Chrispeels, M.J., Baumgartner, B., and Harris, N. (1976). Regulation of reserve protein metabolism in the cotyledons of mung bean seedlings. Proc. Natl. Acad. Sci. USA **73,** 3168–3172.

Chou, P.Y., and Fasman, G.D. (1974). Prediction of protein conformation. Biochemistry **13,** 222–245.

Faye, L., and Chrispeels, M.J. (1987). Transport and processing of the glycosylated precursor of concanavalin A in jack-bean. Planta **170,** 217–224.

Friedlander, M., and Blobel, G. (1985). Bovine opsin has more than one signal sequence. Nature **318,** 338–343.

Frohman, M.A., Dush, M.K., and Martin, G.R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. Proc. Natl. Acad. Sci. USA **85,** 8998–9002.

Garnier, J., Osguthorpe, D.J., and Robson, B. (1978). Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J. Mol. Biol. **120,** 97–120.

Gorin, M.B., Yancey, S.B., Cline, J., Revel, J.-P., and Horwitz, J. (1984). The major intrinsic protein (MIP) of the bovine lens fiber membrane: Characterization and structure based on cDNA cloning. Cell **39,** 49–59.

Hartmann, E., Rapoport, T.A., and Lodish, H.F. (1989). Predicting the orientation of eukaryotic membrane-spanning proteins. Proc. Natl. Acad. Sci. USA **86,** 5786–5790.

Heller, K.B., Lin, E.C.C., and Wilson, T.H. (1980). Substrate specificity and transport properties of the glycerol facilitator of *Escherichia coli*. J. Bacteriol. **144,** 274–278.

Jackson, R.J., and Hunt, T. (1982). Preparation and use of nuclease-treated rabbit reticulocyte lysate for the translation of eukaryotic messenger RNA. Methods Enzymol. **96,** 50–74.

Johnson, K.D., Herman, E.M., and Chrispeels, M.J. (1989). An abundant, highly conserved tonoplast protein in seeds. Plant Physiol. **91,** 1006–1013.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157,** 105–132.

Lipp, J., Flint, N., Haeuptle, M.-T., and Dobberstein, B. (1989). Structural requirements for membrane assembly of proteins spanning the membrane several times. J. Cell Biol. **109,** 2013–2022.

Maeder, M., and Chrispeels, M.J. (1984). Synthesis of an integral protein of the protein body membrane in *Phaseolus vulgaris* cotyledons. Planta **160,** 330–340.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Muller, L.L., and Jacks, T.J. (1983). Intracellular distribution of free sugars in quiescent cottonseed. Plant Physiol. **71,** 703–704.

Muramatsu, S., and Mizuno, T. (1989). Nucleotide sequence of the region encompassing the glpKF operon and its upstream region containing a bent DNA sequence of *Escherichia coli*. Nucl. Acids Res. **17,** 4378.

Nelson, N., and Taiz, L. (1989). The evolution of H^+-ATPases. Trends Biochem. Sci. **14,** 113–116.

Nishimura, M., and Beevers, H. (1979). Hydrolysis of protein in vacuoles isolated from higher plant tissue. Nature **277,** 412–413.

Paul, D.L., and Goodenough, D.A. (1983). In vitro synthesis and membrane insertion of bovine MP 26, an integral protein from lens fiber plasma membrane. J. Cell Biol. **96,** 633–638.

Plant, A.R., and Moore, K.G. (1983). The protein, lipid and carbohydrate composition of protein bodies from *Lupinus angustifolius* seeds. Phytochemistry **22,** 2359–2363.

Prescott, A., and Martin, C. (1987). A rapid method for the quantitative assessment of levels of specific mRNAs in plants. Plant Mol. Biol. Rep. **4,** 219–224.

Pusztai, A., Croy, R.R.D., Stewart, J.S., and Watt, W.B. (1979). Protein body membranes of *Phaseolus vulgaris* L. cotyledons: Isolation and preliminary characterization of constituent proteins. New Phytol. **83,** 371–378.

Rao, Y., Jan, L.Y., and Jan, Y.N. (1989). *Big brain*, a neurogenic gene in *Drosophila*, encodes a transmembrane protein. In Abstracts of Papers Presented at the Molecular Neurobiology of *Drosophila* Meeting, W. Pak and D. Ready, eds (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory), p. 15.

Rapoport, T.A. (1985). Extensions of the signal hypothesis—Sequential insertion versus amphipathic tunnel hypothesis. FEBS Lett. **187,** 1–10.

Sandal, N.N., and Marcker, K.A. (1988). Soybean nodulin 26 is homologous to the major intrinsic protein of the bovine lens fiber membrane. Nucl. Acids Res. **16,** 9347.

Spencer, D., Higgins, T.J.V., Button, S.C., and Davey, R.A. (1980). Pulse-labeling studies on protein synthesis in developing pea seeds and evidence of a precursor form of legumin small subunit. Plant Physiol. **66,** 510–515.

Sturm, A., Van Kuik, J.A., Vliegenthart, J.F.G., and Chrispeels, M.J. (1987). Structure, position, and biosynthesis of the high mannose and the complex oligosaccharide side chains of the bean storage protein phaseolin. J. Biol. Chem. **262,** 13392–13403.

Verner, K., and Schatz, G. (1988). Protein translocation across membranes. Science **241,** 1307–1313.

von Heijne, G., and Gavel, Y. (1988). Topogenic signals in integral membrane proteins. Eur. J. Biochem. **174,** 671–678.

Walker, K.A. (1974). Changes in phytic acid and phytase during early development of *Phaseolus vulgaris* L. Planta **116,** 91–98.

Wessels, H.P., and Spiess, M. (1988). Insertion of a multispanning membrane protein occurs sequentially and requires only one signal sequence. Cell **55,** 61–70.

Yamamoto, Y.T., Acedo, G.N., and Conkling, M.A. (1989). A tobacco root-specific gene: Characterization and regulation of its expression. J. Cell. Biochem. (Suppl.) **13D,** 313.

Zampighi, G.A., Hall, J.E., and Kreman, M. (1985). Purified lens junctional protein forms channels in planar lipid films. Proc. Natl. Acad. Sci. USA **82,** 8468–8472.