

Sazzad Karim · Dan Lundh · Kjell-Ove Holmström
Abul Mandal · Minna Pirhonen

Structural and functional characterization of AtPTR3, a stress-induced peptide transporter of *Arabidopsis*

Received: 2 August 2004 / Accepted: 27 January 2005 / Published online: 12 May 2005
© Springer-Verlag 2005

Abstract A T-DNA tagged mutant line of *Arabidopsis thaliana*, produced with a promoter trap vector carrying a promoterless *gus* (*uidA*) as a reporter gene, showed GUS induction in response to mechanical wounding. Cloning of the chromosomal DNA flanking the T-DNA revealed that the insert had caused a knockout mutation in a PTR-type peptide transporter gene named *At5g46050* in GenBank, here renamed *AtPTR3*. The gene and the deduced protein were characterized by molecular modelling and bioinformatics. Molecular modelling of the protein with fold recognition identified 12 transmembrane spanning regions and a large loop between the sixth and seventh helices. The structure of AtPTR3 resembled the other PTR-type transporters of plants and transporters in the major facilitator superfamily. Computer analysis of the AtPTR3 promoter suggested its expression in roots, leaves and seeds, complex hormonal regulation and induction by abiotic and biotic stresses. The computer-based hypotheses were tested experimentally by exposing the mutant plants to amino acids and several stress treatments. The *AtPTR3* gene was induced by the amino acids histidine, leucine and phenylalanine in cotyledons and lower leaves,

whereas a strong induction was obtained in the whole plant upon exposure to salt. Furthermore, the germination frequency of the mutant line was reduced on salt-containing media, suggesting that the AtPTR3 protein is involved in stress tolerance in seeds during germination.

Keywords Promoter trapping · Fold recognition · Peptide transporter · Wounding · Salt stress · *Arabidopsis*

Introduction

An increasing amount of sequence information is becoming available from various organisms. When a genomic sequence has been completed, the next phase is to assign functions to the predicted proteins. Presently, one approach for predicting function of a protein is based on analysis of protein structure by molecular modelling. If no homology exists to experimentally determined three-dimensional structures, fold recognition offers an alternative way to get a testable hypothesis of the function. Fold recognition is used to predict the fold by matching the sequence to an already known fold [1–3]. The approach usually involves undertaking several steps; fold assignment, template selection, alignment, model generation and evaluation/refinement [4]. The limitation of the method is that it can only recognize experimentally determined folds. However, the estimated probability of any new protein belonging to a previously characterized fold may reach more than 95% [5]. Fold recognition and comparative modelling have been applied successfully for functional prediction of individual proteins [6–8] and used at genome scale for many organisms [9].

The first fully sequenced plant was the common wall cress, *Arabidopsis thaliana*, a small weed in the mustard family [10]. It is used as one of the model systems for plant molecular biology due to its small genome (7×10^7 bp) and short generation time (5–8 weeks).

S. Karim · K.-O. Holmström · A. Mandal
School of Life Sciences,
University of Skövde,
541 28 Skövde, Sweden

S. Karim
Department of Plant Biology and Forest Genetics,
Swedish University of Agricultural Sciences,
750 07 Uppsala, Sweden

D. Lundh
School of Humanities and Informatics,
University of Skövde,
541 28 Skövde, Sweden

M. Pirhonen (✉)
Department of Applied Biology,
University of Helsinki, Box 27,
FIN, 00014 Helsinki, Finland
E-mail: minna.pirhonen@helsinki.fi

Although its sequence was published several years ago, a substantial part of the annotated proteins still lack functional prediction. In plants, one of the possible ways to get functional prediction of an uncharacterized gene is offered by *Agrobacterium*-mediated insertion mutagenesis. This method can be used to generate both T-DNA tagged insertion mutants with recognizable phenotypes and reporter gene fusions to plant promoters [11, 12]. When a promoter-trapping T-DNA construct is inserted into the plant nuclear genome, it can produce a knock-out mutant, where the activity of the reporter gene can be used to study the regulation of the mutated gene. Several strategies have been developed for cloning of the chromosomal DNA flanking the insertion, leading to rapid identification of the tagged genes [8, 12–14]. For *A. thaliana*, T-DNA tagged mutant lines are available for most of the open reading frames in the genome [15].

The major facilitator superfamily (MFS) is one of the biggest groups of transporters and membrane proteins and is ubiquitously present in all three groups of living organisms [16]. These transporters use proton-motive force to drive the transport of a wide range of molecules, such as ions, sugars, sugar-phosphates, nucleosides, amino acids, peptides, neurotransmitters and drugs in a uniport, symport or antiport mode. Despite the vast range of substrates and different transport mechanisms, the MFS transporters are 400–600 amino acids long and likely to be composed of 12 transmembrane helices with both the N- and C-termini located on the cytoplasmic side [17]. Three-dimensional crystal structures of two MFS superfamily proteins of *Escherichia coli*, lactose permease and glycerol-3-phosphate transporter, were recently determined [18, 19], and a third bacterial protein, oxalate transporter, was characterized by cryo-electron microscopy [20]. The reported structures of these proteins show that they are composed of pseudo-symmetric N- and C-terminal halves, each composed of six transmembrane helices, and a substrate binding site, which is located in a pore between the two halves. The two parts of the protein are connected by a long cytoplasmic loop, which allows the halves to move relative to each other in a “rocker-switch”-type of movement. Binding of the substrate changes the conformation of the protein and causes the substrate-binding cavity to open to the other side of the membrane, where the substrate is released [18–20]. Regardless of the low sequence similarity between the characterized members in MFS superfamily, it seems that this fold and transport mechanism may be shared by all MFS proteins [21].

The acquisition and allocation of nitrogenous compounds, including amino acids and peptides, is an essential process for most living organisms. Twenty characterized transporter families can transport amino acids, whereas peptides can be transported by members in four families [22]. Peptide transporters in di-/tripeptide transporter family (PTR, also called the POT family for proton-dependent oligopeptide transporter) are about 450–600 amino acids long and usually exhibit 12 putative α -helical transmembrane regions and cause

symport of one or more H^+ simultaneously with the transported molecule [22–25]. It has been suggested that the PTR family is a distant element of the MFS, in spite of the fact that sequence similarities are insufficient to establish homology [16]. Instead, this conclusion is based on the similarity between shorter segments as well as similarities in topology and transport mechanism [16, 26].

Phylogenetic analyses of transporters in PTR family indicate that they can be classified into four groups according to the presence and location of a hydrophilic loop in the proteins [27]. Groups I–III contain peptide transporters of bacteria, animals and fungi, respectively, whereas group IV contains a mixture of peptide and nitrate transporters of plants as well as histidine/peptide transporters from mammals. Though homologous, the nitrate transporters and peptide transporters of plants seem to form functionally distinct groups in this family [27–29]. However, some PTR transporters of plants appear to be able to transport both types of compounds as well as other molecules, such as amino acids or chlorate [30–32]. Furthermore, a nodule-specific PTR-type dicarboxylate transporter of alder (*Alnus glutinosa*) was recently characterized, expanding the list of possible molecules transported by PTR-type transporters [33].

Analysis of the *Arabidopsis* genomic sequence has revealed that this plant has ten times more peptide transporters than any other sequenced organism [34]. Approximately 50 putative PTR-type transporters have been identified in the *Arabidopsis* genome with sequence analysis (<http://www.biology.ucsd.edu/~ipaulsen/transport/>). The first characterized was a di- and tripeptide transporter AtPTR2 (previously called as AtPTR2-B), which was identified by functional complementation of a yeast transport mutant *PTR2* with *Arabidopsis* clone [35]. *Arabidopsis* mutants expressing the *AtPTR2* cDNA in antisense orientation exhibited delay in flowering and an arrest in seed development [36]. The same transporter was identified independently as histidine transporter and designated NTR1 [37]. Recently, a new PTR-type di- and tripeptide transporter, named AtPTR1, was identified from *Arabidopsis* by yeast complementation [38]. The corresponding gene was shown to be expressed during seed germination and in vascular tissues. Besides PTR-type peptide transporters, nitrate transporters of PTR-type have been identified in *Arabidopsis*. The first of them was a nitrate and chlorate transporter CHL1, recently renamed as AtNRT1.1 [39, 40]. Later another homologue in the same gene family was identified and named AtNRT1.2 [28]. In total, four PTR-type nitrate transporters in AtNRT1 family have now been identified in *Arabidopsis* [41].

In this paper we describe cloning and in silico analysis of a wound-induced *Arabidopsis* transporter AtPTR3 belonging to the PTR family of peptide transporters. Molecular modelling showed that the AtPTR3 protein resembled MFS proteins and computer-based promoter analyses indicated a stress-induced expression pattern of the corresponding gene. The computer-based predictions

could be verified by experimental data showing that, in addition to the wound-induction, the transcription of the gene could also be activated when plants were exposed to amino acids or salt. Furthermore, the germination of the *Atptr3* knockout-mutant was shown to be sensitive to salt stress.

Materials and methods

Plant material and growth conditions

Mutant line 9 of *A. thaliana*, of ecotype C24, originates from screening of T-DNA tagged mutant lines for induced expression of the reporter gene *uidA* present in the vector pMHA2 [42]. Mutant lines carrying a T-DNA inserted in *At5g46050* in ecotype Columbia-0 was ordered from Salk Insertion Sequence Database (<http://signal.salk.edu>). All plant lines were grown in jars on basal MS medium (Sigma-Aldrich) [43] supplemented with 1% sucrose. Plants were grown in Petri dishes or in jars in growth chambers at 22°C and 70% relative humidity with 12 h day length at 100 µE light intensity and used when 3–4 weeks old. For induction of line 9 by amino acids, glutamic acid, histidine, leucine, methionine, phenylalanine or tryptophan were pipetted to final concentration of 10 mM on the surface of the MS medium. For salt experiments, plants were cultivated on Petri dishes on MS medium containing 20–200 mM NaCl.

Inducing treatments

For the analysis of wound-induced expression of the *AtPTR3* gene, plants were wounded in leaves by forceps and samples collected 2, 4, 8, 12 and 24 h later for GUS staining or RNA isolation. To test the induction of GUS activity by drought, line 9 plants were kept in a laminar flow hood at 40% relative humidity and 22°C with open lid and samples were collected at 2, 4, 8, 12 and 24 h. In order to examine if the GUS activity can be induced by cold, unwounded plants were kept at +4°C 2–4 h before using them in GUS analysis. Induction of GUS activity by amino acids was studied by pipetting the amino acids from stock solutions on the media and collecting the plants for GUS analysis 2, 4, 8, 12 and 24 h later. For the salt experiments plants were grown 4 weeks on MS medium supplemented with 20–200 mM NaCl (at 20-mM intervals). In each experiment, control plants were either kept in original growth conditions or exposed to sterile water instead of the inducing agent. All experiments were repeated at least four times.

Histochemical GUS staining

The histochemical GUS staining was done by immersing detached leaves or whole plants in 0.05 M PO₄ buffer,

pH 7.0, containing 0.5 mg/ml X-GlcA (Duchefa). The tube was wrapped with aluminium foil and incubated at 37°C for 4 h, followed by incubation in chlorophyll washing solution (95% ethanol : acetic acid; 3 : 1) for 1 h at 37°C, after which the samples were preserved in 95% ethanol. Induced GUS expression was identified based on appearance of blue colour in the plant tissue.

Southern blotting and cloning

About 3 µg of plant genomic DNA was digested with the restriction enzymes *Bam*HI, *Eco*RI, *Hind*III and their combinations and run in a 1% agarose gel, followed by blotting on membrane. Hybridization was performed as described by Southern [44] and Sambrook et al. [45] using a 1.9 kb PCR fragment containing the *gus* coding region as a probe. Labelling was done by Rediprime II DNA labelling kit using redivue dCTP α -³²P label (Amersham Biosciences). After hybridization, *Eco*RI–*Hind*III-digested genomic DNA was run similarly in 1% gel and the DNA corresponding to the size of the hybridizing fragment (based on the previous Southern) was cut out, purified (QIAquick Gel Extraction Kit, Qiagen) and cloned into *Eco*RI–*Hind*III-cut cloning vector pBlueScriptSK⁺ and transformed into TOP10 competent *E. coli* cells (Invitrogen). Colonies were blotted on membrane and hybridized with the same *gus* probe. After isolating the plasmid DNA from positive colonies universal primers of the vector and *gus* reverse primer 5'-CACGGGTTGGGGTTTCT-3' were used to sequence the ends of the clone and the DNA upstream of the right border of the T-DNA.

Similarity search and homology determination

The predicted peptide transporter *At5g46050* sequence was obtained from GenBank [46]. The Non-Redundant DataBase (NRDB), and the Protein Data Bank (PDB) [47] were searched for functionally related sequences using BLAST [48] and PSI-BLAST [49]. The secondary database tool Interpro [50] was used for identifying known motifs. The identified homologous sequences were aligned using the default parameters in the ClustalW tool [51].

Reverse transcriptase-polymerase chain reaction

RNeasy plant mini kit (Qiagen) was used to purify total RNA from line 9 and C24-plants (five leaves per plant) 2, 4, 8, 12 and 24 h after induction by wounding. RT-PCR analysis was performed with 1 µg of total RNA with RobusT II RT-PCR kit (M-MuLV RT RNase H⁻/DyNAzyme EXT) from Finnzymes. The reverse transcription was performed according to the manufacturer's instructions (48°C for 30 min followed by 94°C for 2 min). For the subsequent PCR, primers 5'-AA-

CAGTTGATCTTCAAGGCAATCCT-3' and 5'-GGTGGCTTAATTCGCGGTATAGTTA-3' were designed from two different exons of the At5g46050 gene, allowing distinction of amplified products arising from mRNA and genomic DNA that contains an intron. The PCR was performed for 30 cycles under the following conditions: 94°C for 30 s, 60–65°C for 30 s, and 72°C for 1 min, followed by a final extension at 72°C for 10 min. Samples were visualized on 2% agarose gels. To show that the GUS transgene was present in line 9 and absent in wild-type C24, RT-PCR was run with GUS forward primer 5'-CCGGGTGAAGGTTATCTCTATGAAC-3' and GUS reverse primer 5'-GAGCGTCGCAGAACATTACATTGAC-3' by using the same program as above. All primers used in this work were designed and analysed using the software Oligo v.6 (Molecular Biology Insight Inc.).

Transmembrane topology prediction

The topology of the PTR was predicted by using TMHMM [52], Phobius [53], Memsat [54, 55], Top-Pred [56] and PSORT [57]. The consensus of these tools was used as a baseline for the suggested topology.

Fold recognition

To identify possible templates for the predicted sequence of the AtPTR3 protein, we used 3D-PSSM [58] as a working tool. Based on the results from previous steps, the template prediction was made by threading the sequence using the 3D-PSSM tool [58]. The reason for choosing the 3D-PSSM tool was based on the experience that THREADER [2, 59] does not perform well on integral membrane proteins. The aim of threading the sequence was to obtain a template for structure prediction. The resulted threading, i.e., alignment and suggested fold, were compared to the predicted topology and threading score for selection of suitable template. The structure prediction was then made by MODELLER [60]. To obtain an accurate model, we also searched for long non-threaded regions in PDB (here, long is defined as exceeding eight amino acids).

Structure prediction

The alignment of structurally related regions of the template from 3D-PSSM and AtPTR3 was checked manually for possible improvements in comparison to the suggested topology. To obtain a structure model similar to the native fold, the homologous tool MODELLER 7 v.7 was used [60]. For the estimation of the prediction quality the software PROCHECK [61] was employed together with manual inspection of energy and solvent accessibility. The selected model, out of several, was identified by having the lowest energy (total as well

as stereochemical clashes), and containing no suspicious structures (knots). This model was investigated further by energy minimization by molecular dynamics before a final evaluation.

Results

Identification and cloning of a wound-induced PTR-type transporter gene of *A. thaliana*

A wound-induced *A. thaliana* line was identified when T-DNA tagged mutant lines, transformed with a promoter-probing vector containing a promoterless *gus* (*uidA*) reporter gene, were screened for cold-induced expression of GUS marker in detached leaves [42]. Further characterization of this line 9 suggested that the GUS activity was induced by wounding the plant when a leaf was removed for the assay. GUS activity could be detected even without cold treatment 4 h after wounding, with increasing activity up to 24 h after treatment (Fig. 1). In the mutant line, a weak constitutive GUS expression could be seen in the hydathodes. No blue colour could be seen in the wild-type plant, suggesting that the blue colour seen in the mutant plant was due to the *gus* transgene (Fig. 1). These results suggest that in mutant line 9 the promoterless *gus* reporter gene had been inserted downstream of a wound-induced *Arabidopsis* promoter enabling the transcription of the *gus* reporter gene in the mutant line.

To characterize the insert in line 9, the genomic DNA was isolated from plants of T₃ generation and used in a Southern analysis (Fig. 2). When the genomic DNA was cleaved with *EcoRI* and *BamHI*, which do not cut the promoter-trapping vector pMHA2, only one large band with size larger than 14 kb was observed. With *HindIII*, which cleaves between the *gus* coding sequence and the rest of the T-DNA in the binary vector, one clear band with size 4.3 kb was observed. A *HindIII*–*EcoRI* double digestion reduced the size of the fragment with 500 bp (Fig. 2), indicating the presence of an *EcoRI* site originating from plant DNA present in this fragment. The fragment was cloned and sequenced, revealing *Arabidopsis* DNA upstream of the start codon of the promoterless *gus* gene in the T-DNA. Homology search (BLASTN) with the cloned *Arabidopsis* fragment revealed 100% identity with *Arabidopsis* genomic DNA in chromosome 5 and suggested that the T-DNA had caused a knockout-mutation in an *Arabidopsis* PTR-type transporter gene *At5g46050*. This gene is renamed here as *AtPTR3* because it is the third characterized gene coding for PTR-type peptide transporters in *Arabidopsis*.

To verify that the correct chromosomal DNA had been cloned, expression of the *AtPTR3* was studied by RT-PCR in the mutant line 9 and in the corresponding wild-type accession C24. Furthermore, the existence of a *gus* reporter gene in the mutant line was verified. RT-PCR with *gus*-specific and *AtPTR3* gene-specific primers

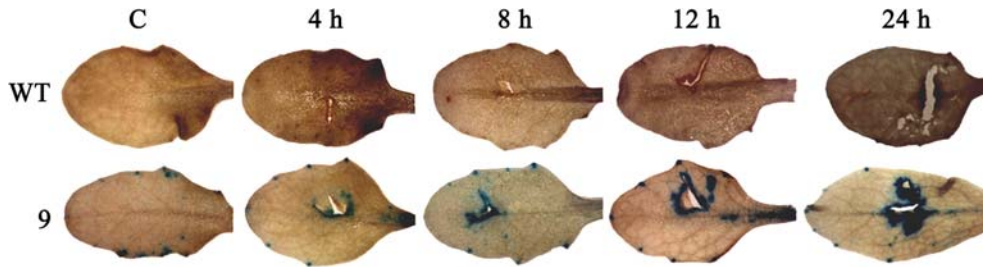


Fig. 1 Wound-inducible GUS expression in the mutant line 9 carrying a promoter probe *uidA* insert. *WT*, leaves of wounded wild-type *A. thaliana* plants; 9, leaves of axenically grown T-DNA tagged plants of line 9 wounded by forceps. Treated leaves were exposed to GUS assay 4, 8, 12 and 24 h after treatment. Induced GUS expression was identified based on appearance of blue colour in treated leaves. *C* stands for untreated control leaves

confirmed the existence of a *gus* transgene in the mutant line 9 and *AtPTR3* transcript in wild-type C24, respectively (Fig. 3a). In the wild-type plant, a gradually increasing expression of the *AtPTR3* gene was observed up to 24 h after wound induction, whereas in mutant line 9 no signal was detected even 24 h after wounding (Fig. 3b). These results suggest that the wound-induced gene *AtPTR3* had been mutated in line 9, in which the induction of the gene could be followed easily by GUS staining.

Similarity search and homology determination of the *AtPTR3*

In public databases the *Arabidopsis* protein At5g46050 (*AtPTR3*) is labelled as a peptide transporter protein in the PTR family. Homology search with BLASTP revealed high similarity between *AtPTR3* and other transporter proteins in the PTR transporter family. The E-values ranged from e^{-136} to $2e^{-73}$ and both peptide and nitrate transporters were among the closest homologues (data not shown).

The search for structural homologues with blast in the PDB database by sequence similarity did not provide any significantly related structure. The best match found in PDB had E-value of 4 ($1E12$; Halorhodopsin, a light-driven chloride pump). The high E-value disabled the structure for being a template. The search for

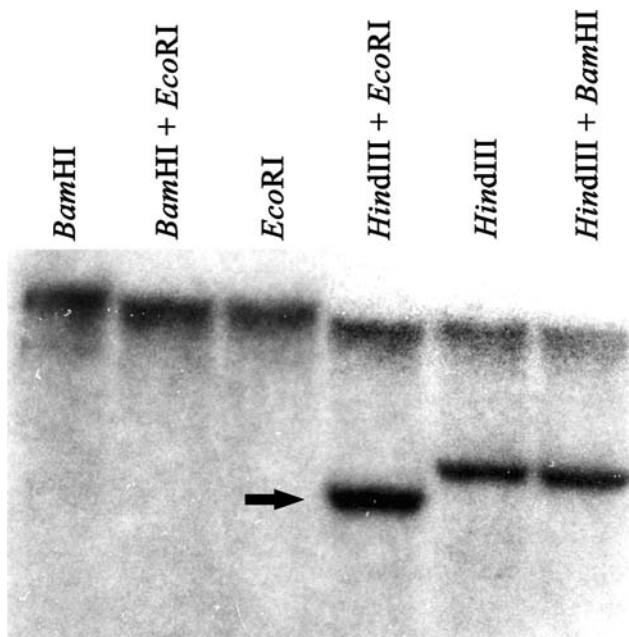


Fig. 2 Southern analysis of the genomic DNA of line 9. About 3 μ g of plant genomic DNA was loaded in each lane of 1% agarose gel after cleaving with restriction enzymes, *Bam*HI, *Eco*RI, *Hind*III and their combinations. A 1.9 kb PCR fragment containing the *gus* coding sequence was used as a probe. Arrow shows the 3.8 kb fragment which was cloned to identify the *Arabidopsis* genomic DNA flanking the T-DNA

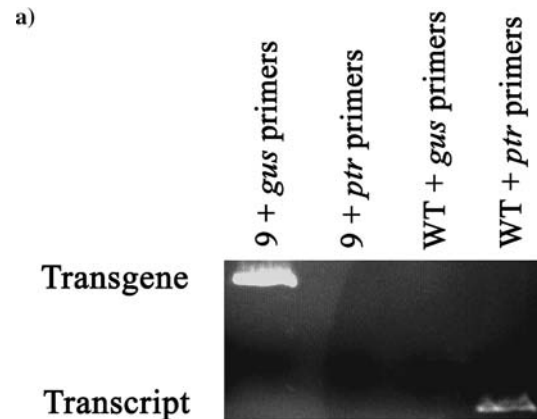


Fig. 3 a RT-PCR with *gus*-specific (*gus* primers) and *AtPTR3*-specific primers (*ptr* primers) verified the *gus* transgene in the mutant line 9 and *AtPTR3* transcript in wild-type C24 (*WT*), respectively. **b** The expression of the *AtPTR3* transcript in wounded wild-type C24 (*WT*) plants was studied by RT-PCR. In wild-type plants, a gradually increasing expression of *AtPTR3* was observed up to 24 h after wounding. In mutant line 9, no signal was detected after wound induction. *SM* indicates size marker

common motifs with Interpro resulted in a hit in region 99–501 that matches the POT (proton-dependent oligopeptide transport; PTR2) family (PFAM:PF00854), and a hit in region 40–55 that matches a domain of beta and gamma crystallins (PROSITE: PS00225, *CRYSTALLIN_BETAGAMMA*). However, the domain hit of beta and gamma crystallins in region 40–55 was in conflict with a transmembrane region suggested for amino acids 46–63 and was therefore omitted. The PTR2 family signatures in Pfam consist of two consensus sequences, i.e., PTR2 family proton/oligopeptide symporters signature 1: [GA]-[GAS]-[LIVMFYWA]-[LIVM]-[GAS]-D-x-[LIVMFYWT]-[LIVMFYW]-G-x(3)-[TAV]-[IV]-x(3)-[GSTAV]-x-[LIVMF]-x(3)-[GA] and PTR2 family proton/oligopeptide symporters signature 2: [FYT]-x(2)-[LMFY]-[FYV]-[LIVMFYWA]-x-[IVG]-N-[LIVMAG]-G-[GSA]-[LIMF] [62]. The AtPTR3 sequence matches the PTR2 Pfam motifs with an E-value of 8e-68.

Cellular localization and transmembrane topology prediction of AtPTR3 protein

Cellular localization of the AtPTR3 protein was studied using the PSORT and SignalP (<http://www.cbs.dtu.dk/services/SignalP/>) programs. PSORT suggested localization of AtPTR3 in the endoplasmic reticulum with a certainty of 66.7%. This suggestion was in line with the prediction from SignalP which indicated a cellular location other than chloroplast, mitochondrion and secretory pathway.

The prediction of a transmembrane topology resulted in a sequential order of membrane-spanning helical

elements. This topology was useful for identification of a suitable template. We used a number of tools based on different techniques to generate a set of topologies. The topology that was predicted, by taking the consensus of these tools, suggested 12 transmembrane elements. The general agreement in placement of the segments resulted in the following topology (Fig. 4a). All predictions resulted in a large loop of 90–100 amino acids, most likely with cytoplasmic localization, between helix VI and helix VII. The predicted topology matched with the topology of other PTR-type transporters of plants [27].

Fold recognition and structure prediction of AtPTR3 protein

Since the initial results from a BLAST search indicated that AtPTR3 protein is an integral membrane protein, we threaded the sequence with 3D-PSSM (Table 1). According to our experience THREADER does not perform well on transmembrane proteins. The results showed that the best and third-ranked folds, glycerol-3-phosphate transporter (1PW4) and lactose permease (1PV7), were transporters in MFS superfamily. The other top-ranked proteins were either different cytochrome-*c*-oxidases (1EHK, 2OCC, 1AR1 and 1BCC) or a chloride channel (1KPK). Functional keywords identified by PSSM indicate: peptide transport (weight 37), symport (29), transport (26), transmembrane (22), glycoprotein (17), herbicide resistance (12) and nitrate assimilation (11).

The PSSM-E value for the top-ranked fold indicated 95% certainty. To derive a structure of the AtPTR3 protein, the software MODELLER 7 v.7 was used on the template structure 1PW4. The transmembrane regions suggested from the topology prediction were checked manually against the alignment of 1PW4 with AtPTR3 so that corresponding elements were suggested. When using MODELLER, the suggested transmembrane regions were restrained according to the suggested topology. MODELLER generated a number of models from which the best was selected, based on energy (compared to template), visual inspection and PROCHECK. The resulting predicted structure is shown in Fig. 4b.

Fig. 4 a A schematic diagram of the predicted topology of AtPTR3. The transmembrane helices correspond to the following amino acids: *I*, 46–63; *II*, 74–93; *III*, 100–120; *IV*, 143–163; *V*, 190–211; *VI*, 220–238; *VII*, 332–352; *VIII*, 370–390; *IX*, 410–430; *X*, 452–473; *XI*, 493–515; *XII*, 538–558. For clarity, the long loop between helices *VI* and *VII* has been shortened and is indicated by *L*. **b** The predicted structure of AtPTR3 using 1PW4 as template. A long cytoplasmic loop connecting helices *VI* and *VII* has been omitted between amino acid 247–327. The cuts are indicated by *L*. The letters *N* and *C* denote the positions of the N- and C-terminal, respectively

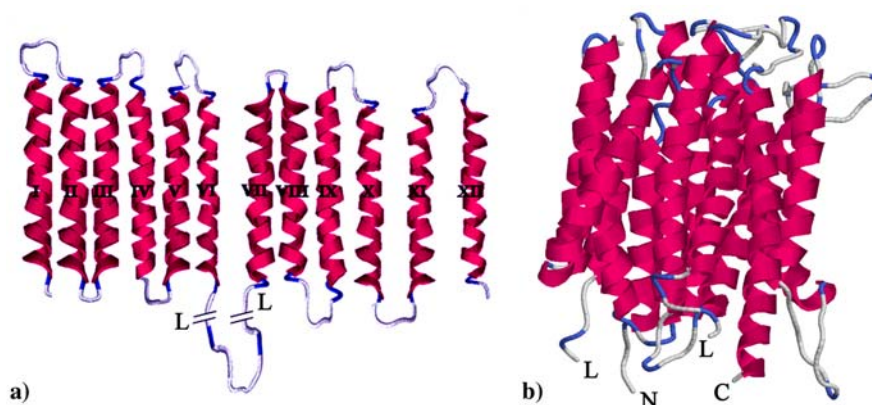


Table 1 The results from 3D-PSSM with significant hits

E-value (certainty) ^a	Fold ^b	Function ^c	Sequence identity ^d
0.0169 (95%)	1PW4 (chain A) membrane protein	Glycerol-3-phosphate transporter	15%
0.0859 (90%)	1EHK (chain A) membrane protein	Aberrant ba3-cytochrome- <i>c</i> oxidase2	14%
0.103 (90%)	1PV7 (chain A) transport protein	Lactose permease	14%
0.269 (70%)	1KPK (chain A) membrane protein	Chloride channel	16%
0.337 (70%)	1FFT (chain A) membrane all-alpha	Homologue to 2occ	13%
0.395 (50%)	2OCC (chain A) membrane all-alpha	Cytochrome- <i>c</i> -oxidase	12%
0.4 (50%)	1AR1 (chain A) membrane all-alpha	Cytochrome- <i>c</i> -oxidase	13%
0.647 (50%)	1BCC (chain C) membrane all-alpha	Cytochrome bc1 transmembrane subunits	17%

^aProbability for random match and the stated certainty for the given fold to be correct by 3D-PSSM.

^bIdentified fold in PDB.

^cFunction of the fold.

^dSequence identity between fold sequence and *AtPTR3*.

The energy of the predicted structure resembles the energy of 1PW4, i.e., no large peaks are identified. The software PROCHECK reports that 97.7% of the distribution of the phi and psi angles are in allowed areas (2.1% in generously allowed areas). The Morris et al. [63] classification for the prediction is 1–2–1 (phi–psi distribution, chi 1 standard deviation, and H-bond energy standard deviation; 1 is one standard deviation below average, 2 is average, 3 is one standard deviation above average). The template used 1PW4 (chain A) has 98.9% of the phi and psi angles in allowed areas, and 1.1% in generously allowed areas, and a Morris et al. [63] classification of 1–1–2. Based on these data, the structure of *AtPTR3* was predicted (Fig. 4b).

The long cytoplasmic loop spanning helical segment VI–VII, indicated in Fig. 4a, was further investigated by BLAST and 3D-PSSM. Homology search with the loop sequence indicated similarity with PTR-type proteins transporting peptides or nitrate/chlorate (data not shown). The 3D-PSSM search resulted in non-significant scores because of the length of the region.

Analysis of regulatory elements on *AtPTR3* promoter

The chromosomal sequence corresponding to the *AtPTR3* gene and the upstream promoter sequence was retrieved from GenBank and the upstream regulatory region of the gene was scanned by PLACE (<http://www.dna.affrc.go.jp/PLACE/>) to find known regulatory elements in plant promoters. The results indicated several putative promoter regulatory elements, 569 sites distributed on 79 different regulatory elements. Frequent keywords among these elements indicated tissue specific expression (seed, root, shoot and leaf), regulation by light and several hormones, involvement of several abiotic stress responses and defence mechanisms against pathogens (Table 2).

Induction of *AtPTR3* transcription with amino acids and abiotic stresses

Because in the databases the *AtPTR3* protein was indicated to be a peptide transporter, it was tested whether

amino acids can induce the expression of the *AtPTR3* gene. It would be expected that the transported molecules or its constituents should be able to induce the production of the components in the uptake system. GUS analysis of the mutant line 9 offers an easy way to study gene expression and regulation of the *AtPTR3* gene. Plants of the mutant line were treated with six amino acids: glutamic acid, histidine, leucine, methionine, phenylalanine and tryptophan. Plants treated with phenylalanine, histidine and leucine showed GUS activity in the cotyledons and lower leaves, indicating that transcription of *AtPTR3* gene can be activated by these amino acids (Fig. 5a). The three other amino acids, glutamic acid, methionine and tryptophan, did not show

Table 2 Frequencies of the putative regulatory elements in the *AtPTR3* promoter

Regulation specific to	Keywords	Number of binding sites	Number of elements
Green parts of the plant	Leaf	105	18
	Shoot	86	16
Seed	Seed	70	15
	Storage	19	3
Root	Root	18	14
Light	Light	42	9
	Circadian	5	2
	Chloroplast	4	1
	Etiolation	4	2
Hormones	ABA	27	7
	GA	7	4
	Ethylene	4	3
	SA	3	1
	Auxin	2	2
Abiotic stresses	Cold	18	5
	Drought	12	6
	Salt	1	1
Plant–pathogen interaction	SAR	16	3
	HR	14	1
	Resistance	3	1
	Disease	3	1
	Flavonoids	7	1
Secondary metabolism	PAL	1	1

ABA, abscisic acid; GA, gibberellic acid; SA, salicylic acid; SAR, systemic acquired resistance; HR, hypersensitive response; PAL, phenylalanine ammonia lyase. Elements and binding sites are not unique occurrences, i.e., an element may be involved in regulation by more than one feature, e.g., roots and hormones.

any GUS induction in the same experiment (data not shown).

Bioinformatics analysis of the regulatory elements upstream of the *AtPTR3* coding sequence suggested that the gene is involved in abiotic stresses. To test this hypothesis, induction of the GUS activity in mutant line 9 plants was studied after exposing unwounded plants to cold, drought and salt stress. When axenically grown plants were placed in a laminar flow hood with open lid, the mutant plants showed weak induction of GUS expression especially in the petioles, whereas cold treatment at +4°C did not induce any expression (data not shown). Growth on NaCl-containing medium, 20–200 mM, caused strong induction of GUS activity in line 9 plants. The intensity of GUS activity increased with rising salt concentration and on the higher salt concentrations the whole plants, including cotyledons and roots, turned intensive blue (Fig. 5b). Growth of both the line 9 knockout mutants and the C24 control plants was reduced on higher salt concentration (Fig. 5c).

Germination of *Atptr3* mutant seeds on salt-containing medium

Seeds of mutant line 9 and control plants demonstrated no difference in germination when sown on sterile media with salt concentrations up to 100 mM, but on higher salt concentrations the germination frequency of the mutant was lower than that of the wild-type seeds (Fig. 5c). At 140 and 160 mM NaCl concentrations 86.5 and 68.6% of the wild-type seeds but only 40.1 and 15.7% of the mutant seeds germinated, respectively. On the highest NaCl concentration, 200 mM NaCl, 13.9% of the wild-type seeds and only 1.7% of the mutant line seeds germinated (Fig. 6). To verify that the results were due to insert in the *AtPTR3* gene and not by other unrelated factors, the corresponding Salk (SIGnAL) mutant lines were ordered from line collection and used in similar experiments. These mutants also showed reduced germination on high salt when compared to the corresponding wild-type plants, *A. thaliana* accession Columbia-0. On the highest NaCl concentrations, 180 and 200 mM, 95 and 80% of the Columbia-0 plants germinated, while the germination of the Salk line was reduced to 65 and 40%, respectively. These results suggest that the *AtPTR3* protein, besides being induced by salt, is also involved in germination under salt stress.

Discussion

A mutant *Arabidopsis* line harbouring a promoter-trapping T-DNA with a promoterless gus gene, exhibited GUS activity when induced by wounding. Cloning, sequencing and bioinformatics analyses of the chromosomal DNA upstream of the insert revealed a knockout mutation in the gene *At5g46050*, annotated as a PTR-type peptide transporter, renamed *AtPTR3* in this paper.

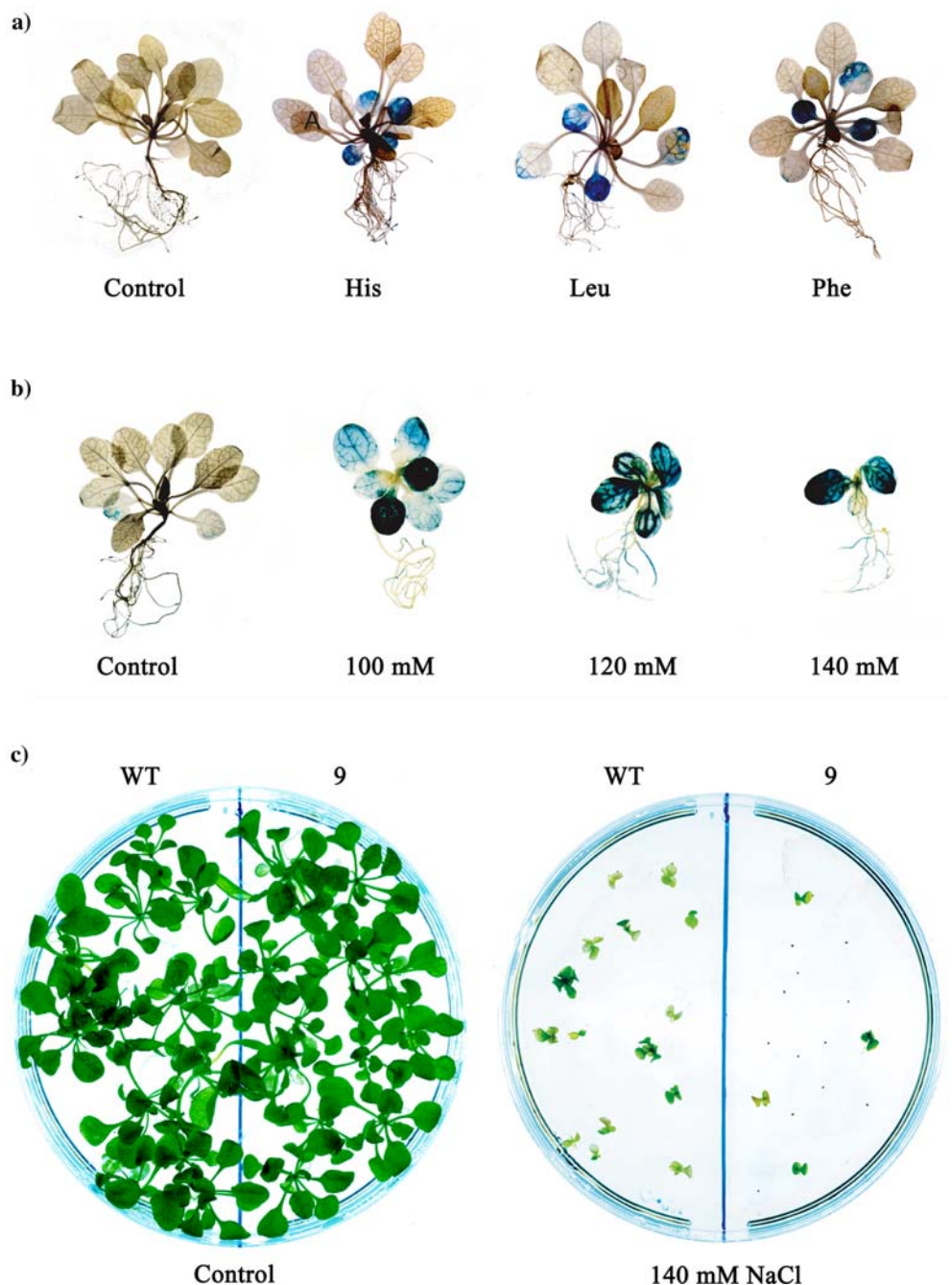
The annotation of the protein function as peptide transporter is apparently based on its homology to the peptide transporters in the PTR family. However, in BLAST analyses, several nitrate transporters of PTR-type could be found among the homologues, suggesting that *AtPTR3* might transport other molecules as well. Activation of the *AtPTR3* gene after exposure of the plants to amino acids suggests that it is likely that the protein is involved in transporting amino acids or peptides rather than nitrate. For conclusive identification of the activity and the transported molecule the *AtPTR3* protein should be expressed in a heterologous system.

To gain further functional clues, we performed a protein-structure prediction using the MODELLER software with a template obtained from fold recognition by the software 3D-PSSM. In our experience, 3D-PSSM is the best tool for threading transmembrane proteins. An alternative tool is the THREADER software. However, it does not work well for this class of proteins according to the THREADER manual. Transmembrane protein structures are difficult to determine experimentally, and the number of experimentally determined structures is consequently few. By using fold recognition, we obtained a template that we merged with a transmembrane topology prediction. From this model we could derive structural properties that would be difficult to obtain otherwise, for example, the large cytoplasmic loop in the middle of the protein.

The suggested fold from 3D-PSSM, 1PW4, is probably the best template for the *AtPTR3* available at the moment. The fold recognition by 3D-PSSM indicates a 95% certainty. However, higher certainty might follow when more experimental structures of membrane proteins are available. The transmembrane topology of *AtPTR3* shows 12 helical regions, which is in agreement with the structures of the PTR/POT and MFS family proteins, both stated to have 12 putative transmembrane regions. The topology and helical packing suggest that the predicted fold is a member of the PTR family and also most likely a member of the MFS superfamily, as suggested recently [26]. However, the secondary database search by Interpro does not report the sequence of *AtPTR3* to match the MFS family. When the topology of *AtPTR3* was derived through consensus of the topology predictions, ten helices were predicted by all the tools, whereas Memsat and Phobius suggested two additional helices in the N-terminal. The topology and structure prediction enabled the identification of the large cytoplasmic loop in *AtPTR3*. A similar loop in the same position in MFS transporters has been suggested to have a structural function in keeping the two parts of the protein together, while still allowing them to move relative to each other in different conformations of the proteins [18–20]. In MFS protein lactose permease, a similar loop has been suggested to be involved in membrane insertion of the protein [64] as well as binding of regulatory protein [65].

To get more information about the function of *AtPTR3*, the upstream regulatory regions of the gene were

Fig. 5 **a** Induction of *AtPTR3* gene by amino acids. GUS staining of line 9 plants 8 h after induction with amino acids. *Control* indicates plant treated with water. *His*, *Leu* and *Phe* indicate plants treated with 10 mM amino acids histidine, leucine or phenylalanine, respectively. **b** Induction of *AtPTR3* gene by salt. GUS staining of line 9 plants grown on MS medium on different salt concentrations. *Control* indicates plant grown on MS medium and 100, 120 and 140 mM indicate plants grown on MS medium supplemented with the indicated NaCl concentrations. Size of the plants grown on salt medium has been magnified. **c** Germination frequency of *Atptr3* knockout mutant line is reduced on salt medium. *Atptr3* knockout mutant (9) and wild-type C24 (*WT*) sown on MS medium (*Control*) and MS medium supplemented with salt (140 mM NaCl)



analysed with bioinformatics. Analysis of the promoter regulatory elements with PLACE suggested myriads of possible promoter elements. A high proportion of the possible promoter elements indicated seed-, root-, leaf- and shoot-specific expression and regulation by light. Furthermore, abiotic stresses, especially cold and dehydration (drought), defence-related gene expression and regulation by several hormones involved in stress responses were also indicated. Only some of the elements are reported to be present in the promoter of *AtPTR3* by AGRIS (The Ohio State University *Arabidopsis* Gene Regulatory Information Server, <http://arabidopsis.med.ohio-state.edu>), possibly reflecting a more

stringent use of experimentally proven regulatory elements present in the promoters of *Arabidopsis*.

The result of the computer analysis of the promoter suggests several stress-related elements probably regulating the *AtPTR3* gene. When expression of the *AtPTR3* gene was studied after cold, drought and salt treatment, salt stress was found to cause substantial induction of the *AtPTR3* gene, thereby supporting the results of computer analysis. Furthermore, increased seed sensitivity to salt in the growth media was observed in the knockout mutant. Even the modelling indicated similarity to salt-transporting proteins (halorhodopsin and chloride channel), which was ignored due to insig-

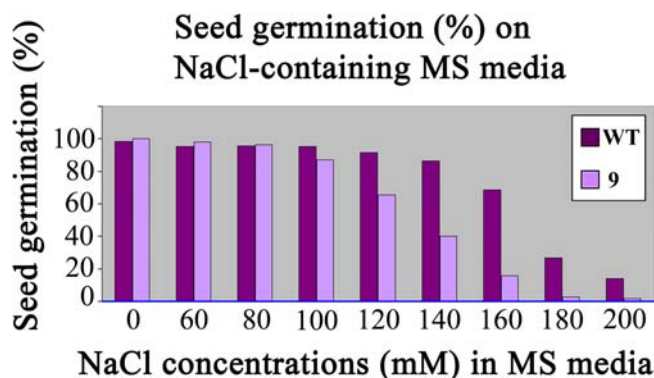


Fig. 6 Frequencies of wild-type C24 (*WT*) and *Atptr3* knockout mutant line (9) seed germination on MS media supplemented with different concentrations of NaCl (60, 80, 100, 120, 140, 160, 180 and 200 mM). The plants were grown as in Fig. 5c and the results presented as cumulative mean of germination percentage on four plates. The control treatment without added NaCl is indicated as 0

nificant E-values. The precise interaction, if any, between salt and the transporter activity of *AtPTR3* is unknown at the moment. *AtPTR3* is the first characterized peptide transporter that is regulated by stresses, such as wounding and most likely osmotic stress. Previously, the *Arabidopsis* PTR-type nitrate transporter *AtNRT1.1* has been shown to be associated with stress, contributing to drought susceptibility due to its involvement in stomatal opening in the presence of nitrate [66]. Many stress responses caused by drought, freezing and osmotic stress have partially overlapping signalling pathways with wounding, which causes crosstalk between the different stress responses and interactions [67]. Tissue damage usually induces local osmotic stress responses often found to be a key component in the response to mechanical wounding [68, 69]. About 8% of the 8,200 genes studied in *Arabidopsis* were altered by wounding at steady-state mRNA levels. Among these are many osmotic stress- and heat shock-regulated genes [67].

Subcellular localization of PTR-type transporters is not well studied. *AtPTR1* of *Arabidopsis* was shown to be localized to plasma membrane in tobacco protoplasts [38], and in barley peptide transporter *HvPTR1* present in germinating seed was localized to plasma membranes of scutellum cells [70]. To characterize the subcellular localization of the *AtPTR3* protein, PSORT and SignalP were used. This analysis suggested that the *AtPTR3* protein could reside in the endoplasmic reticulum or in the membrane of organelles other than chloroplast, mitochondrion or secretory pathway. The results of computer analysis of *AtPTR3* are in agreement with the information in the Transporter Classification System (<http://tcdb.ucsd.edu>), where some of the transporters are mentioned to reside in organellar membranes such as lysosomes.

The induction of the *AtPTR3* gene by wounding and salt stress could suggest that upon attack and osmotic stress peptide transport is needed for relocation or uptake of nutrients for stress response and defence. Sugar

transporter genes *STP4* and *AtSU* are similarly induced by wounding, suggesting that also the utilization of other nutrients than amino acids is regulated by wounding [71, 72]. The reduced germination of *Atptr3* mutant on salt-containing medium and the activity of the gene in cotyledons suggest that this gene is active during germination, possibly in utilization of peptides originating from digested seed storage proteins. Both the characterized PTR-type peptide transporters of *Arabidopsis*, *AtPTR1* and *AtPTR2*, and the homologous transporters in faba bean and in barley have been suggested to be expressed either in developing or germinating seeds [36, 38, 70, 73].

With bioinformatics and modelling new knowledge can be developed fast. Unfortunately, computer-based analyses do not always produce entirely correct data. However, computer-based methods can produce novel and unexpected hypotheses that can be tested by molecular biological experiments. Furthermore, exploration and validation of data created with bioinformatics is necessary for further refinement of bioinformatic tools.

Acknowledgements The research grant from Nilsson-Ehle Foundation is gratefully acknowledged.

References

- Bowie JU, Luethy R, Eisenberg D (1991) *Science* 253:164–170
- Jones DT (1999) *J Mol Biol* 287:797–815
- Jones DT (2000) *Curr Opin Struct Biol* 10:371–379
- Sánchez R, Pieper U, Melo F, Eswar N, Martí-Renom MA, Madhusudhan MS, Mirkovic N, Šali A (2000) *Nat Struct Biol* 7:986–990
- Brenner SE, Chothia C, Hubbard TJP (1997) *Curr Opin Struct Biol* 7:369–376
- Šali A, Matsumoto R, McNeil HP, Karplus M, Stevens RL (1993) *J Biol Chem* 268:9023–9034
- Fetrow JS, Godzik A, Skolnick J (1998) *J Mol Biol* 282:703–711
- Svensson M, Lundh D, Ejdebäck M, Mandal A (2004) *J Mol Model* 10:130–138
- Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, Šali A (2000) *Annu Rev Biophys Biomol Struct* 29:291–325
- Arabidopsis* Genome Initiative (2000) *Nature* 408:796–815
- Springer PS (2000) *Plant Cell* 12:1007–1020
- Valentine L (2003) *Plant Physiol* 133:948–955
- Lindsey K, Wei W, Clarke MC, McArdle HF, Rooke LM, Topping JF (1993) *Transgenic Res* 2:33–47
- Mathur J, Szabados L, Schaefer S, Grunenberg B, Lossow A, Jonas-Straube E, Schell J, Koncz C, Koncz-Kalman Z (1998) *Plant J* 13:707–716
- Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, Gadriab C, Heller C, Jeske A, Koesema E, Meyers CC, Parker H, Prednis L, Ansari Y, Choy N, Deen H, Geralt M, Hazari N, Hom E, Karnes M, Mulholland C, Ndubaku R, Schmidt I, Guzman P, Aguilar-Henonin L, Schmid M, Weigel D, Carter DE, Marchand T, Risseuw E, Brogden D, Zeko A, Crosby WL, Berry CC, Ecker JR (2003) *Science* 301:653–657
- Pao SS, Paulsen IT, Saier MH Jr (1998) *Microbiol Mol Biol Rev* 62:1–34
- Abramson J, Iwata S, Kaback HR (2004) *Mol Membr Biol* 21:227–236

18. Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S (2003) *Science* 301:610–615
19. Huang Y, Lemieux MJ, Song J, Auer M, Wang D-N (2003) *Science* 301:616–620
20. Hirai T, Heymann JAW, Maloney PC, Subramaniam S (2003) *J Bacteriol* 185:1712–1718
21. Vardy E, Arkin IT, Gottschalk KE, Kaback HR, Schuldiner S (2004) *Protein Sci* 13:1832–1840
22. Saier MH Jr (2000) *Microbiol Mol Biol Rev* 64:354–411
23. Saier MH Jr (2000) *Microbiology* 146:1775–1795
24. Paulsen IT, Skurray RA (1994) *Trends Biochem Sci* 19:404–405
25. Steiner HY, Naider F, Becker JM (1995) *Mol Microbiol* 16:825–834
26. Chang AB, Lin R, Studley WK, Tran CV, Saier MH Jr (2004) *Mol Membr Biol* 21:171–181
27. Chiang C-S, Stacey G, Tsay Y-F (2004) *J Biol Chem* 279:30150–30157
28. Huang N-C, Liu K-H, Lo H-J, Tsay Y-F (1999) *Plant Cell* 11:1381–1392
29. Lin C-M, Koh S, Stacey G, Yu S-M, Lin T-Y, Tsay Y-F (2000) *Plant Physiol* 122:379–388
30. Galván A, Fernández E (2001) *Cell Mol Life Sci* 58:225–233
31. Frommer WB, Hummel S, Rentsch D (1994) *FEBS Lett* 347:185–189
32. Zhou J-J, Theodoulou FL, Muldin I, Ingemarsson B, Miller AJ (1998) *J Biol Chem* 273:12017–12023
33. Jeong J, Suh SJ, Guan C, Tsay Y-F, Moran N, Oh CJ, An CS, Demchenko KN, Pawlowski K, Lee Y (2004) *Plant Physiol* 134:969–978
34. Stacey G, Koh S, Granger C, Becker JM (2002) *Trends Plant Sci* 7:257–263
35. Song W, Steiner H-Y, Zhang L, Naider F, Stacey G, Becker JM (1996) *Plant Physiol* 110:171–178
36. Song W, Koh S, Czako M, Marton L, Drenkard E, Becker JM, Stacey G (1997) *Plant Physiol* 114:927–935
37. Rentsch D, Laloi M, Rouhara I, Schmelzer E, Delrot S, Frommer WB (1995) *FEBS Lett* 370:264–268
38. Dietrich D, Hammes U, Thor K, Suter-Grotemeyer M, Fluckiger R, Slusarenko AJ, Ward JM, Rentsch D (2004) *Plant J* 40:488–499
39. Tsay YF, Schroeder JI, Feldmann KA, Crawford NM (1993) *Cell* 72:705–713
40. Guo FQ, Wang R, Chen M, Crawford NM (2001) *Plant Cell* 13:1761–1777
41. Okamoto M, Vidmar JJ, Glass ADM (2003) *Plant Cell Physiol* 44:304–317
42. Mandal A, Sandgren M, Holmström K-O, Gallois P, Palva ET (1995) *Plant Mol Biol Rep* 13:243–254
43. Murashige T, Skoog F (1962) *Phys Plantarum* 15:473–497
44. Southern E (1975) *J Mol Biol* 98:503–517
45. Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY
46. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Rapp BA, Wheeler DL (2000) *Nucleic Acids Res* 28:15–18
47. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) *Nucleic Acids Res* 28:235–242
48. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) *J Mol Biol* 215:403–410
49. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) *Nucleic Acids Res* 25:3389–3402
50. Apweiler R, Attwood TK, Bairoch A, Bateman A, Birney E, Biswas M, Bucher P, Cerutti L, Corpet F, Croning MDR, Durbin R, Falquet L, Fleischmann W, Gouzy J, Hermjakob H, Hulo N, Jonassen I, Kahn D, Kanapin A, Karavidopoulou Y, Lopez R, Marx B, Mulder NJ, Oinn TM, Pagni M, Servant F, Sigrist CJA, Zdobnov EM (2001) *Nucleic Acids Res* 29:37–40
51. Thompson JD, Higgins DG, Gibson TJ (1994) *Nucleic Acids Res* 22:4673–4680
52. Sonnhammer ELL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. In: Glasgow J, Littlejohn T, Major F, Lathrop R, Sankoff D, Sensen C (eds) *Proc ISMB*. AAAI Press, Menlo Park, CA, pp175–182
53. Käll L, Krogh A, Sonnhammer ELL (2004) *J Mol Biol* 338:1027–1036
54. Jones DT, Taylor WR, Thornton JM (1994) *Biochemistry* 33:3038–3049
55. Jones DT (1998) *FEBS Lett* 423:281–285
56. von Heijne G (1992) *J Mol Biol* 225:487–494
57. Nakai K, Horton P (1999) *Trends Biochem Sci* 24:34–35
58. Kelley LA, MacCallum RM, Sternberg MJE (2000) *J Mol Biol* 299:499–520
59. Jones DT (1997) *Curr Opin Struct Biol* 7:377–387
60. Sánchez R, Šali A (1997) *Curr Opin Struct Biol* 7:206–214
61. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *J Appl Cryst* 26:283–291
62. Meredith D, Boyd CAR (2000) *Cell Mol Life Sci* 57:754–778
63. Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) *Proteins* 12:345–364
64. Weinglass AB, Kaback HR (2000) *Proc Natl Acad Sci USA* 97:8938–8943
65. Seok Y-J, Sun J, Kaback HR, Peterkofsky A (1997) *Proc Natl Acad Sci USA* 94:13515–13519
66. Guo F-Q, Young J, Crawford NM (2003) *Plant Cell* 15:107–117
67. Cheong YH, Chang H-S, Gupta R, Wang X, Zhu T, Luan S (2002) *Plant Physiol* 129:661–677
68. Reymond P, Weber H, Damond M, Farmer EE (2000) *Plant Cell* 12:707–719
69. Denekamp M, Smeekens SC (2003) *Plant Physiol* 132:1415–1423
70. Waterworth WM, West CE, Bray CM (2000) *J Exp Bot* 51:1201–1209
71. Truernit E, Schmid J, Epple P, Illig J, Sauer N (1996) *Plant Cell* 8:2169–2182
72. Meyer S, Lauterbach C, Niedermeier M, Barth I, Sjolund RD, Sauer N (2004) *Plant Physiol* 134:684–693
73. Miranda M, Borisjuk L, Tewes A, Dietrich D, Rentsch D, Weber H, Wobus U (2003) *Plant Physiol* 132:1950–1960