

Barley Nodulin 26-like intrinsic protein permeates water, metalloids, saccharides, and ion pairs due to structural plasticity and diversification

Received for publication, July 15, 2023, and in revised form, September 22, 2023. Published, Papers in Press, October 31, 2023,

<https://doi.org/10.1016/j.jbc.2023.105410>

Akshayaa Venkataraghavan[‡], Julian G. Schwerdt[†], Stephen D. Tyerman[†], and Maria Hrmova^{*†}

From the School of Agriculture, Food and Wine, and Waite Research Institute, Waite Research Precinct, University of Adelaide, Glen Osmond, South Australia, Australia

Reviewed by members of the JBC Editorial Board. Edited by Phillip A. Cole

Aquaporins can facilitate the passive movement of water, small polar molecules, and some ions. Here, we examined solute selectivity for the barley Nodulin 26-like Intrinsic Protein (HvNIP2;1) embedded in liposomes and examined through stopped-flow light scattering spectrophotometry and *Xenopus laevis* oocyte swelling assays. We found that HvNIP2;1 permeates water, boric and germanic acids, sucrose, and lactose but not D-glucose or D-fructose. Other saccharides, such as neutral (D-mannose, D-galactose, D-xylose, D-mannoheptaose) and charged (N-acetyl D-glucosamine, D-glucosamine, D-glucuronic acid) aldoses, disaccharides (cellobiose, gentiobiose, trehalose), trisaccharide raffinose, and urea, glycerol, and acyclic polyols, were permeated to a much lower extent. We observed apparent permeation of hydrated KCl and MgSO₄ ions, while CH₃COONa and NaNO₃ permeated at significantly lower rates. Our experiments with boric acid and sucrose revealed no apparent interaction between solutes when permeated together, and AgNO₃ or H[AuCl₄] blocked the permeation of all solutes. Docking of sucrose in HvNIP2;1 and spinach water-selective SoPIP2;1 aquaporins revealed the structural basis for sucrose permeation in HvNIP2;1 but not in SoPIP2;1, and defined key residues interacting with this permeant. In a biological context, sucrose transport could constitute a novel element of plant saccharide-transporting machinery. Phylogenomic analyses of 164 Viridiplantae and 2993 Archaeal, bacterial, fungal, and Metazoan aquaporins rationalized solute poly-selectivity in NIP3 sub-clade entries and suggested that they diversified from other sub-clades to acquire a unique specificity of saccharide transporters. Solute specificity definition in NIP aquaporins could inspire developing plants for food production.

Aquaporins (AQPs) are fundamental to water and other solute movements in nearly all living organisms. AQPs are membrane proteins classified in the family of major intrinsic proteins (MIPs) that transport water and neutral solutes (1, 2), although the repertoire of permeated solutes has recently expanded as some AQPs may also transport ions (3). The AQP

family consists of seven sub-families of plasma membrane intrinsic proteins (PIPs), tonoplast intrinsic proteins (TIPs), Nodulin 26-like Intrinsic Proteins (NIPs), and small and basic intrinsic proteins (SIPs). Other less common sub-families include Solanaceae X-intrinsic proteins (XIPs) (4) found in fungi, slime molds, and dicots, while GlpF-like intrinsic proteins (GIPs) (5) and hybrid (between PIPs and TIPs) intrinsic proteins (HIPs) are found in green algae, the moss *Physcomitrella* and lycopods (4). AQPs from each sub-family differ in physicochemical properties that underlie their complex roles in osmotic homeostasis, root water uptake from the soil, root, and leaf hydraulic conductance, lateral root emergence, motor cell movement, internode elongation, the diurnal regulation of leaf movements, petal development, biotic interactions, signaling and stomatal dynamics (3, 6, 7).

NIPs, named after soybean nodulin 26 (NOD26), an abundant protein in the peribacteroid membrane of nitrogen-fixing root nodules, are functional equivalents of aquaglyceroporins (8) and occur in barley (9), rice (10), wheat (11), and other plants. Besides water and boric acid (BA) (9, 10), NIPs transport other hydroxylated metalloids such as arsenious and germanic acid (9, 12–15), glycerol (12), urea (16), hydrogen peroxide (13), acyclic polyols, purines and pyrimidines (17), neutral lactic acid (18), selenious and antimonious acids (19), and aluminum malate (20). Some AQPs facilitate gas diffusions such as CO₂ (21), NH₃ (22–24), and O₂ (25). When NOD26 was incorporated in lipid bilayers, ionic currents were detected with slight anion over cation selectivity (26, 27). From an evolutionary point of view, plant NIPs (28) and Solanaceae XIPs (“metalloido-porins”) (29), permeating heavy metals and H₂O₂ but not water (30), were shown to evolve from cyanobacteria (31) some 1500 million years ago. Then a primordial AQP NIP-like (*aqpN*) gene was acquired by some plants via horizontal transfer shedding light on NIPs solute evolution and selectivity (32, 33).

The structural monomeric sub-unit of functional AQP folds into six tilted membrane-spanning α -helices and two short re-entrant α -helices, running in two repeats, with five inter-connecting loops forming a right-handed α -helical bundle. A permeating pore running through each monomer has a defined morphology and width in each AQP sub-family; for example, barley HvNIP2;1 has a wide pore along its entire

[‡] These authors contributed equally to this work.

^{*} For correspondence: Maria Hrmova, maria.hrmova@adelaide.edu.au.

Broad-selective barley HvNIP2;1 aquaporin permeates sucrose

length compared to PIPs and TIPs (2, 34, 35). In native environments, AQPs oligomerize into tetramers along a four-fold rotational axis and create a tightly fitted trapezoid that at structural and functional levels may form an additional central fifth pore. Each AQP monomer has its solute path and with a central pore, the tetramer potentially offers five permeating routes (3, 36). The broad solute selectivity of the monomer pore of NIPs is supported by Asn-Pro-Ala (NPA) motifs that are separated by 4 Å to 5 Å from the aromatic/R selectivity filter residues (2, 36).

To achieve the constant adjustment of water and solute permeability in fluctuating environments, cells developed multiple controls of AQP function through structural features. For example, extended N- and C-terminal regions in NIPs (9, 37) could impact transport, gating, and *in vivo* AQP expression levels. These regions and cytosolic loops (e.g., loop D) in SoPIP2;1 (PDB-Protein Data Bank 2B5F) (38, 39), AtPIP2;4 (PDB 6qim) (40), AtTIP2;1 (PDB 5i32) (41) and recently elucidated rice OsNIP2;1 (PDB 7cjs and 7nl4) (42, 43) have similar spatial dispositions suggesting that dynamics of PIPs, TIPs, and NIPs could be similar. In SoPIP2;1 loop D alongside His193 and dephosphorylated Ser residues trigger pore closure and other structural re-arrangements (38).

To understand the function of HvNIP2;1 in the context of metalloids toxicity, it is fundamental to understand that this trait is a major problem in cereal crops around the world (44). Although metalloids boron and silicon are essential micronutrients, other metalloids such as arsenic and germanium are toxic at certain concentrations, which together with high soil boron pose risks to populations (45). Metalloids exist in soils in the form of amphoteric oxides with atomic radii between 3.43 Å to 4.48 Å, and these differences could be exploited to engineer mutants with selective permeation properties (19).

In the present study, solute selectivity is explored in HvNIP2;1 solubilized from *Pichia pastoris* membranes through embedding in liposomes. The selectivity of HvNIP2;1 towards neutral solutes and ion pairs is examined, establishing that this AQP permeated water, BA, germanic acid, sucrose, and MgSO₄ and KCl ion pairs at relatively high rates, but also permeated at low rates acyclic polyols, urea and glycerol, other ion pairs, and some mono- and disaccharides of various stereochemistries. Molecular docking was performed on the monomeric HvNIP2;1 3D AlphaFold 3D model (accession AF-D8V828) to provide molecular-level descriptions of permeation, including identifying structural elements that underlie solute poly-specificity. These data are compared to those of water-selective SoPIP2;1 AQP in an open state conformation (38, 39). Structural data suggested that the wider pore of HvNIP2;1 lined with acidic and aromatic residues, and high pore flexibility, were the key features supporting the permeation of a wide range of solutes. These data were rationalized by phylogenomic analyses of 3157 Viridiplantae (green algae and land plants), Archaeal, bacterial, fungal, and Metazoan AQPs to show that the members of the NIP3 sub-clade evolved a unique solute specificity of saccharide transporters through neo-functionalization after the emergence of tracheophytes. The HvNIP2;1 substrate poly-specificity definition

underscores yet unidentified roles of this AQP in plant growth, metabolism, and development.

Results

Cloning, expression, solubilization, purification, and liposomal reconstitution of HvNIP2;1

The native sequence of HvNIP2;1 was cloned into the *HvNIP2;1*-Myc-6xHis-pPICZ-B DNA fusion (46–48) and expressed in *P. pastoris*. *Pichia* transformants were screened for high-level protein expression, where HvNIP2;1 appeared predominantly in alkaline soluble or insoluble alkaline/urea fractions (data not shown). HvNIP2;1 was solubilized in styrene-maleic anhydride co-polymer 3:1 (SMA) (49) and purified by Immobilized Metal Affinity Chromatography (IMAC) (46, 47). Near-homogenous HvNIP2;1 (Fig. 1A) was reconstituted in 1, 2-dimyristoyl-sn-glycero-3-phosphocholine (DMPC) liposomes that were floated on an iohexol (Accudenz) gradient (50, 51) to obtain homogenous populations of proteo-liposomes of around 100 µm in size (Fig. 1B) used to define permeation properties of HvNIP2;1.

The identity of purified HvNIP2;1 confirmed by immunoblot, SDS-PAGE, and IB analyses using anti-HvNIP2;1 antibody (Fig. 1) revealed that HvNIP2;1 folded predominantly into monomeric (apparent molecular mass of around 34 kDa) and dimeric (apparent molecular mass of around 70 kDa) forms, but tetrameric forms (apparent molecular mass of around 150 kDa) were also detected (Fig. 1A). The inspection of the tetrameric model of HvNIP2;1, built on the coordinates of OsNIP2;1 (43) (the structural aspects of the model discussed below), indicated that the four monomers displayed extended hydrophobic interactions, which included phenylalanine, leucine and valine residues that linked individual subunits. These types of interactions were noted with other aquaporins (38, 42, 43).

Electrophoretic profiles of purified HvNIP2;1 in the form of diffused bands suggested that HvNIP2;1 was N-glycosylated at putative S26 although O-glycosylation sites could also be occupied. No degradation of purified HvNIP2;1 was observed after 3 weeks at 4 °C indicating that it was amenable for permeation studies. To confirm that HvNIP2;1 was incorporated in DMPC liposomes and not simply aggregated during HvNIP2;1 reconstitution, liposomes with embedded HvNIP2;1 were subjected to floatation by iohexol density gradient ultracentrifugation. As indicated, fractions 2 to 4 containing proteo-liposomes with HvNIP2;1, detected with anti-His antibody, floated near the top of the gradient, and most of HvNIP2;1 incorporated in DMPC liposomes (Fig. 1B).

Permeation properties of HvNIP2;1

Measurements of water and solute permeability, based on osmotic gradients generated by co-incubating DMPC liposomes with embedded HvNIP2;1 or control liposomes lacking the protein, and osmolytes at high concentrations, were assessed by stopped-flow light scattering spectrophotometry (Figs. 2, S1 and S2 and Table S1). These co-incubations created an osmotic gradient that in the first phase led to increased

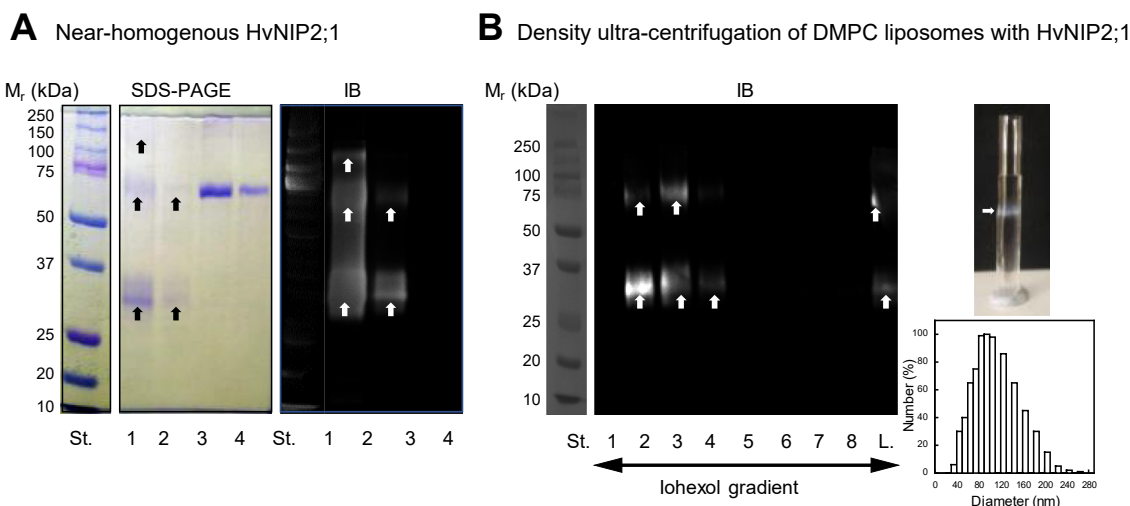


Figure 1. SDS-PAGE and immunoblot (IB) analyses of near-homogenous HvNIP2;1 purified by IMAC, and density gradient ultra-centrifugation of DMPC liposomes with reconstituted HvNIP2;1. A, monomeric (apparent molecular mass 34 kDa), dimeric (70 kDa), and tetrameric (150 kDa) forms of HvNIP2;1 (lanes one and 2) are indicated by *arrows* at approximately 0.5 and 1 μg protein loadings, respectively. Lanes three and four contain BSA (Fraction V) at one or 2 μg loadings, respectively. Immunoblot (IB) analysis proceeded with anti-His antibody, as described in [Experimental procedures](#). St. lanes indicate molecular masses of protein standards. B, density gradient ultra-centrifugation of DMPC liposomes with embedded HvNIP2;1 and proteo-liposomal sizing. Lanes 1 to 8 are fractions collected after ultra-centrifugation in the iohexol gradient, where L (lane 9) contains non-fractionated preparation. Lane with standard proteins (St.) indicates molecular masses of protein standards. HvNIP2;1 was detected by IB with an anti-His antibody. Right top panel shows the test tube of liposomes with reconstituted HvNIP2;1 (*arrow*) after the iohexol gradient floating, forming a white diffuse band. Bottom image displays the size distribution profile of proteo-liposomes analysed by the NICOMP 380 Particle Sizing System, as described in [Experimental procedures](#).

rates of water efflux causing the shrinkage of liposomes, while in the second phase, proteo-liposomes increased their volumes and swelled due to rate-limiting solute transport of water (52). Thus, as shown in [Figure 2](#), fast water flow (shrinkage of liposomes) is reflected by relative values at less than 0.1 s that approach 100% (real values are presented in [Figs. S1](#) and [S2](#)). Conversely, the steep downward slopes of light scattering traces after this time then indicate swelling slopes of HvNIP2;1 proteo-liposomes, where black and white curves indicate normalized and non-linearly fitted data using a one-phase decay model. This occurred with liposomes with reconstituted HvNIP2;1 that mediated solute permeation, but not for control liposomes or when liposomes with embedded HvNIP2;1 were incubated with 0.5 mM AgNO₃, a potent blocker of AQPs (53), for which a re-swelling phase was incomplete ([Fig. 1C](#)). HvNIP2;1 permeated BA and germanic acid at higher rates, while glycerol, D-mannitol and D-sorbitol, and urea were permeated at low rates ([Figs. 2, S1](#) and [S2](#) and [Table S1](#)). Surprisingly, HvNIP2;1 permeated at high rates the disaccharide sucrose, as previously observed (46), and some monosaccharides (D-xylopyranose, L-arabinofuranose), and hydrated ion pairs KCl and MgSO₄, while lactose, the CH₃COONa, and NaNO₃ ion pairs were transported at lower rates. Conversely, NaF, the neutral monosaccharides D-glucopyranose, D-fructopyranose, D-galactopyranose, D-mannopyranose, and D-mannopyranopheptaose, charged monosaccharides D-glucosamine, N-acetyl β-D-glucosamine and D-glucuronic acid, disaccharides trehalose, cellobiose and gentiobiose, and trisaccharide raffinose were permeated at even lower rates ([Fig. S2](#)). Permeation rate constants and derived permeability coefficients (P) ([Table S1](#)) for water, metalloids, saccharides, and other solutes, and ion pairs descended in these orders:

Water and metalloids: H₂O >> BA > germanic acid.
 Polyols and urea: urea > glycerol > D-mannitol > D-sorbitol.
 Saccharides: sucrose > L-arabinofuranose > lactose > D-glucose > D-fructose.
 Ion pairs: MgSO₄ > KCl > CH₃COONa > NaNO₃.
 All permeants: H₂O >> MgSO₄ ≥ BA > KCl ≥ sucrose > urea > glycerol > CH₃COONa > D-mannitol ≥ lactose > germanic acid > NaNO₃ > D-sorbitol > D-glucose > D-fructose.

Measurements of water and solute permeability for selected solutes by HvNIP2;1 were confirmed with *Xenopus laevis* oocytes that were injected with cRNA of HvNIP2;1 or water as controls ([Fig. 3](#)), where the swelling was measured in response to water gradient ([Fig. 3A](#)) confirming water permeation by HvNIP2;1, and isosmotic BA, sucrose, and D-glucose concentration gradients ([Fig. 3](#)). Transport of BA and sucrose was observed ([Fig. 3, B](#) and [C](#)), while D-glucose permeation did not occur ([Fig. 3D](#)), as also shown using the stopped-flow light scattering analyses ([Fig. S2](#)). No detectable interaction was seen between BA and sucrose, when permeated together ([Fig. 3C](#)). In oocytes, incubated with 0.5 mM H[AuCl₄] (53), which is the known histidine and thiol groups modifier, solute transport was inhibited.

Molecular model of HvNIP2;1 and the SoPIP2;1 crystal structure, and docking of sucrose

To investigate the structural features of HvNIP2;1 and work out the structural basis of sucrose permeation, we used the structural bioinformatics approach and aligned the sequences of HvNIP2;1 and SoPIP2;1 (PDB-Protein Data Bank 2B5F), and

Broad-selective barley HvNIP2;1 aquaporin permeates sucrose

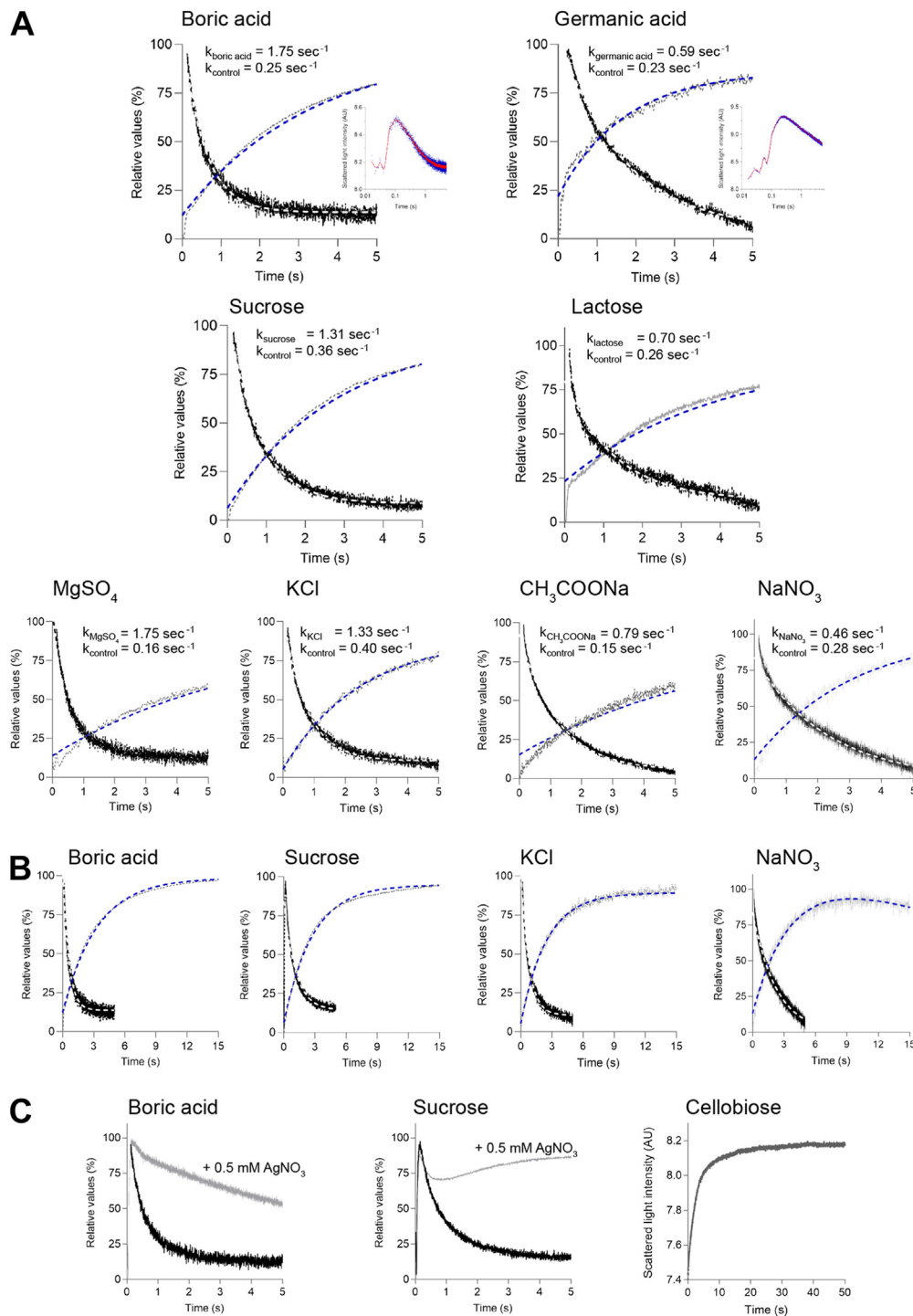


Figure 2. Transport of permeants by HvNIP2;1 embedded in liposomes. A–C, HvNIP2;1 solubilized by SMA and reconstituted in DMPC liposomes were exposed to gradients of solutes generating osmotic gradients with BA and germanic acid (*top panels*), disaccharides sucrose and lactose (*bottom-middle panels*), and MgSO₄, KCl, CH₃COONa, and NaNO₃ (*bottom panels*). The uptake of permeants was measured by stopped-flow spectrophotometry in the liposome buffer] at the concentration of 0.2 M (340 mOsmol/kg). Fast water flow (shrinkage of liposomes) is shown by relative values at less than 0.1 s that approach 100% (the real values indicated in Fig. S1 and S2). The steep downward slopes of light scattering traces after this time indicate swelling slopes of HvNIP2;1 proteo-liposomes, where *black curves* indicate normalised data (as relative values) and *white dashed curves* represent non-linearly fitted data using a one-phase decay model. Light scattering traces of control liposomes (lacking HvNIP2;1) are drawn in *grey* for the first five (panels A) and 15 s (panels B), and non-linear fits using a one-phase association model are shown in *blue dashed curves*. 0.5 mM AgNO₃ was used to inhibit transport in HvNIP2;1 (panels C; *grey curves*). In panel A, the calculated rate constants for each combination of permeants are indicated in s⁻¹ for liposomes with HvNIP2;1 and control liposomes lacking HvNIP2;1. Plots in panels B (control liposomes and proteo-liposomes for BA, sucrose, KCl, and NaNO₃) and panel C (right plot – data of control liposomes exposed to cellobiose) indicate that no solute leakage was observed; the same non-solute leakage profiles were observed for control liposomes. Data were plotted in relative values (%) in GraphPad Prism 9.

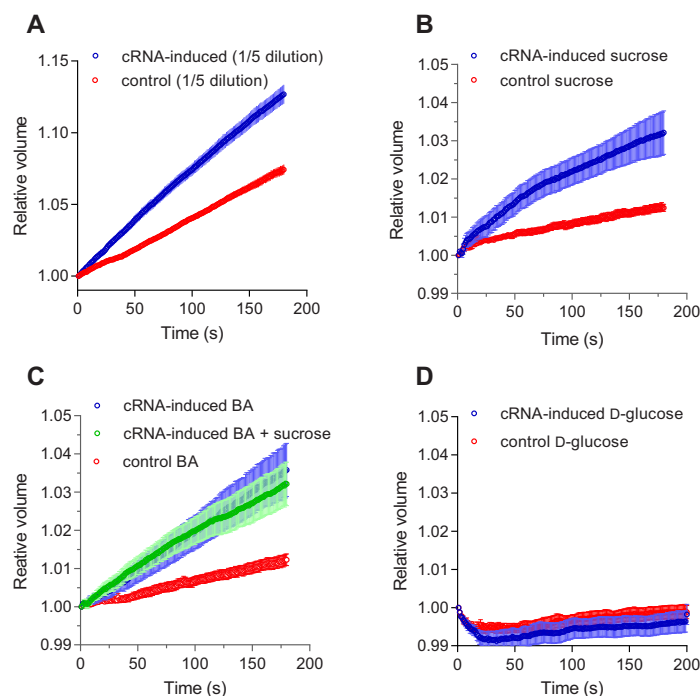


Figure 3. Uptake of permeants by *Xenopus laevis* oocytes transformed with cRNA of HvNIP2;1 and by water-injected (non-transformed) oocytes. *A*, an increase in relative volume (V/V_0) over time of transformed oocytes after replacing incubation solutions with 5-fold diluted ND96 ($n = 7$ oocytes). *B*, isosmotic solutions supplemented with 160 mM sucrose ($n = 7$ oocytes), *C*, 160 mM BA and 80 mM BA with 80 mM sucrose added together ($n = 7$ oocytes), and *D*, 160 mM D-glucose ($n = 8$ oocytes). Controls ($n = 7$ –eight oocytes) refer to water-injected (non-transformed) oocytes. Three batches of oocytes showed the same results. Error bars represent Standard Error of Measurements.

the plant structural homologs AtTIP2;1 (PDB 5i32), AtPIP2;4 (PDB 6qim), and OsNIP2;1 (PDB 7cjs and 7nl4). As shown in Figure 4A, the HvNIP2;1 model has nearly an identical secondary structure distribution to that of the crystal structures of Lsi1 (42) and OsNIP2;1 (43). Here, the distribution of secondary structures, such as α -helices and interconnecting loops coincided (Fig. 4A). We also utilised the 3D AlphaFold HvNIP2;1 model (AlphaFold Protein Structure Database accession D8V828; (54)), which features six tilted membrane-spanning α -helices, two short re-entrant α -helices in two repeats, and five interconnecting loops forming an α -helical bundle (Fig. 4B; HvNIP2;1 – cyan cartoon). Thus, the respective 0.67 Å and 0.45 Å root-mean-square deviation values between OsNIP2;1 and Lsi1, at the sequence similarity/identity of 74%/68% and 80%/71% (using Needle EMBOS algorithm; (55)) were unsurprising. In the HvNIP2;1 model, the selectivity filter residues G88, S207, G216, and R222 constituting the GSGR motif were positioned near the entrance to the pore (Fig. 4B), while the residues of the two conserved NPA1 and NPA2 motifs (Fig. 4B) were located deeper in the pore, in accordance with the positions of these structural elements in both crystal structures. A single solute-conducting pore in HvNIP2;1, investigated through Caver (56) in monomeric HvNIP2;1 was wide alongside its lengths (approximate diameter between 2 Å and 3 Å) without forming constricted regions (approximate diameter range between 2.2 Å and 2.6 Å), and contrary those observed in other AQPs near NPA motifs, such as *e.g.*, in the SoPIP2;1 crystal structure (38, 39) (Fig. 4B; SoPIP2;1 – yellow cartoon).

Individual monomers of HvNIP2;1 are expected to form a quaternary assembly, while the disposition of six α -helices followed a pseudo-two-fold axis that run along a membrane normal and was sub-divided an hour-glass fold with bipartite segments. A significant symmetrical distribution of repeating peptide motifs was observed in each bipartite HvNIP2;1 segment (Fig. 4C and Table S2) detected by the Multiple EM (Expectation Maximization) for Motif Elicitation (MEME) analysis (57). At least ten peptide motifs were identified in each segment based on *p*-values (Table S2).

To investigate the permeation of sucrose, observed in HvNIP2;1 through stopped-flow light scattering spectrophotometry and oocyte swelling, this disaccharide molecule was docked in HvNIP2;1 (Fig. 4B), by means of HDock, which is based on a hybrid algorithm of template-based modeling and *ab initio* free docking (58). The docking unambiguously showed that the sucrose can be accommodated in at least three favorable positions in the pore, making interactions with several residues. One such example indicated that sucrose interacted with carbonyl oxygens of M107 and H160 in the vicinity of the NPA1 motif, carbonyl oxygens of G216, and S217 near the NPA2 motif, and ND2 nitrogen of N219 participating in the NPA2 motif (Fig. 4B; left and right-bottom panels; HvNIP2;1 – cyan cartoon); these interactions were formed at separations between 2.2 Å and 3.2 Å.

To increase the accuracy of data interpretation with HvNIP2;1, the monomeric high-efficiency water-permeating SoPIP2;1 AQP in an open state conformation (38, 39) was used as an accessory structural target in the docking experiment

Broad-selective barley HvNIP2;1 aquaporin permeates sucrose

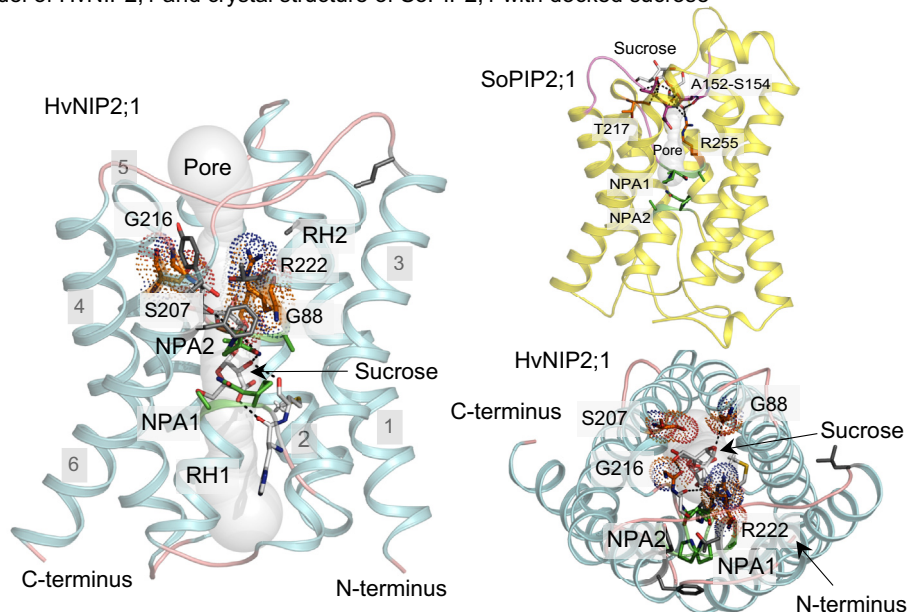
A Multiple sequence alignment of HvNIP2;1, SoPIP2;1 and related plant aquaporins

Conservation:	9	55	5	5	55	59555959555	5	5	6655	6	9	9
AtPIP2;1	1	MA-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	70
AtPIP2;4	1	MA-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	94
SoPIP2;1	1	MS-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	88
OsNIP2;1	1	MASNSR	NSRNSR	ANSNEI	HLDS	TVQNGT	MPTM	-----	-----	-----	-----	95
HvNIP2;1	1	MASNSR	NSRNSR	ATFSSEI	HDL	IGVQNS	TTFSM	-----	-----	-----	-----	95
Consensus aa:		M	-----	-----	-----	-----	-----	-----	-----	-----	-----	
Consensus ss:		hh	-----	-----	-----	-----	-----	-----	-----	-----	-----	hh

Conservation:	56	5	99959	99999	555	6	9	99	596665	5	5965	5	65	956	99	55555	5	5	6	65	69								
AtPIP2;1	71	VAIGANIS	GGHVN	PAVTF	GLAVG	QITVIT	VG	VFWIA	QLLG	STAA	CLFL	KVVT	GGLA	-----	-----	-----	-----	-----	-----	-----	-----	173							
AtPIP2;4	95	VYCTAGIS	GGHIN	PAVTF	GLF	LARKV	SLV	RTVLYI	VAQ	CLGAI	CGCG	VKAF	QSSYY	YTR	YGGG	ANEL	ADG	YNG	KTGL	GAEI	IGT	FVLY	VTFS	ATD	PKRN	ARDS	SHV	PVLA	204
SoPIP2;1	89	VYCTAGIS	GGHIN	PAVTF	GLF	LARKV	SLV	RTVLYI	VAQ	CLGAI	CGCG	VKAF	QSSYY	YTR	YGGG	ANEL	ADG	YNG	KTGL	GAEI	IGT	FVLY	VTFS	ATD	PKRN	ARDS	SHV	PVLA	198
OsNIP2;1	96	IYAVGHIS	GAHM	PAVTL	LAFAV	RHF	FWI	QVPP	VAAQ	FTGAI	CAS	FV	KA	VHP	VPD	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	195	
HvNIP2;1	96	IYAVGHIS	GAHM	PAVTL	LAFAI	RHF	FWI	QVPP	VAAQ	FTGAI	CAS	FV	KA	VHP	VPD	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	195	
Consensus aa:		I	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	
Consensus ss:		hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	

Conservation:	59559	9	55555	59555	99955	59555	6	5	9	55599	9	56	5	5	6	5																			
AtPIP2;1	174	PLAIGL	VGANI	LAAG	PFSG	GSNM	PARS	FGPA	VAAG	D	-----	-----	-----	-----	-----	-----	250																		
AtPIP2;4	205	PLPIG	FAVM	HLAT	IPIT	TGTG	IN	PARS	FGAA	VI	YNS	NEK	AWD	QW	IF	VWG	PMIG	AAAA	FYHQ	FIL	R	AAAI	K	ALG	S	F	G	S	F	R	S	F	A	-----	291
SoPIP2;1	199	PLPIG	FAVM	HLAT	IPIT	TGTG	IN	PARS	FGAA	VI	YNS	NEK	AWD	QW	IF	VWG	PMIG	AAAA	FYHQ	FIL	R	AAAI	K	ALG	S	F	G	S	F	R	S	F	A	-----	281
OsNIP2;1	196	GLAVG	SAVC	ITS	IFAGA	ISGG	SMP	ARTL	GP	ALAS	NK	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	298		
HvNIP2;1	196	GLAVG	SAVC	ITS	IFAGA	ISGG	SMP	ARTL	GP	ALAS	NK	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	295		
Consensus aa:		L	s	L	G	S	S	A	L	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h	h			
Consensus ss:		hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh	hh			

B 3D model of HvNIP2;1 and crystal structure of SoPIP2;1 with docked sucrose



C Distribution of peptide motifs in bipartite segments of 3D

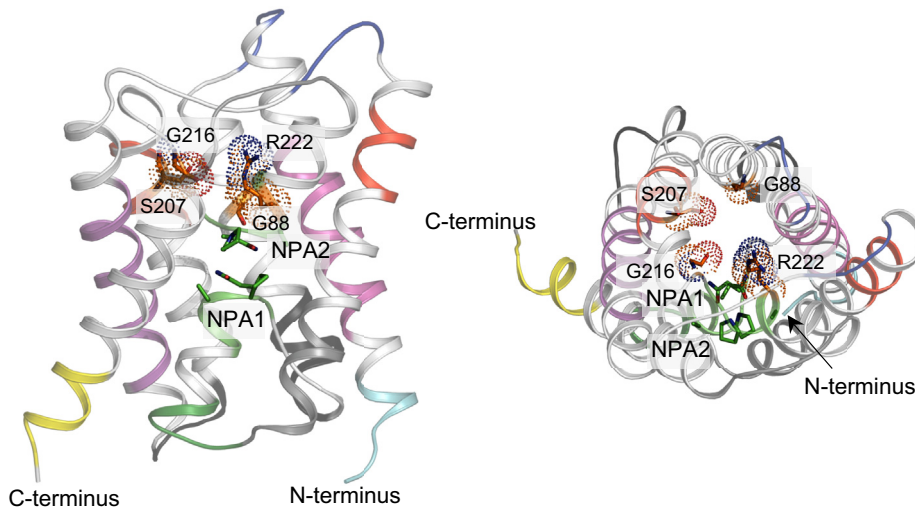


Figure 4. Structural model of HvNIP2;1 and the crystal structure of SoPIP2;1. A, Multiple sequence alignment of HvNIP2;1 and SoPIP2;1 (PDB-Protein Data Bank 2B5F), and plant structural homologues AtPIP2;1 (PDB 5132), AtPIP2;4 (PDB 6qim), and OsNIP2;1 (PDB 7cjs and 7n14). Sequences were aligned using ProMals3D (91), where the alignment indicates the level of conservation of residues (absolutely conserved residues on the scale 5–9 in brown and bold). Respective indices ‘s’, ‘p’, ‘l’ and ‘h’ indicate small (A, G, C, S, V, N, D, T, P), polar (D, E, H, K, N, Q, R, S, T), aliphatic (I, V, L) and hydrophobic (W, F, Y, M, L, I,

(Fig. 4B; right-top panel; SoPIP2;1 – yellow cartoon). With SoPIP2;1, the sucrose disaccharide could not localise inside the pore (approximate diameter range between 1.0 Å and 2.1 Å) and it was always contained within the surface or at the entry or exit to the pore of the SoPIP2;1 in all examined 100 poses. Here, the prominent GGGANSVALGNYK loop (residues 149–161), was localized at the entry of SoPIP2;1, which obstructed the pore and prevented the passage of the sucrose molecule through the pore. Underneath this loop, the second TGTGINPAR loop (residues 217–225) was identified, consisting of nine residues, which also carried the NPA motif and highly conserved R225. Residues from both loops, more specifically A152, N153 and S154 (first loop) and T217 and R225 (second loop) made close contacts with sucrose at separations between 2.2 Å and 3.5 Å, immobilised it and prevented its passage through the SoPIP2;1 pore (Fig. 4B; right top panel). Additionally, it would also be impossible for sucrose to bypass the aromatic/R constriction region point of around 2.1 Å (38), which would be too narrow to accommodate sucrose passage.

Phylogenomic analyses of MIP proteins

To define the phylogenetic relationships, explicitly the variations in selectivity filter residues of MIPs—which define solute selectivity and thus the function of plant aquaporins—we used Randomized Axelerated Maximum Likelihood (RAxML) phylogenetic analysis as a tool. This widely used approach embodies a fast-maximum likelihood tree search algorithm that reconstructs phylogenetic trees with exceptional likelihood scores. Here, we analyzed 164 Viridiplantae MIP sequences from *Chlamydomonas reinhardtii*, *Volvox carterii*, *Physcomitrella patens*, *Amborella trichopoda*, *Spirodela polyrhiza*, *Arabidopsis thaliana*, and *Hordeum vulgare* (Fig. 5). These analyses indicated that these 164 Viridiplantae (Fig. 5) and additional 2993 archaean, bacterial, fungal, and metazoan entries of the MIP family (Pfam database PF00230) (Fig. S3 and Dataset S1) clustered in the four major TIP, NIP and SIP clades, where the PIP and TIP clades covered the majority of MIP entries.

Barley HvNIP2;1 (Fig. 5; marked in bold) with the GSGR selectivity filter residue signature (that is included in square brackets in Fig. 5) was resolved to form a monophyletic group with *A. thaliana* and *A. trichopoda*. We also indicated that HvNIP2;1 split from *P. patens* that carried the FAAR signature, and that the three NIP1, NIP2, and NIP3 representative clades emerged prior to the evolution of tracheophytes (ferns and seed plants), where the NIP2 and NIP3 clades contained basal *Physcomitrella* sequences, and where a NIP1 representative

was perhaps lost. The functional evolution of the NIP clades was evident, and this was indicated by the fact that the selectivity filter in *Physcomitrella* (FAAR) diversified into the GSGR signature, specifically in the NIP3 sub-clade, following the tracheophyte/*Physcomitrella* split. This finding contrasted with the PIP entries, where the FHTR signature was highly conserved in all examined PIP entries (Fig. 5). From the structural point of view, and in accordance with this phylogenomics analysis, the GSGR selectivity filter residue pattern (three smaller side chains with conserved R) in NIP3 entries supports the permeation of larger solutes such as hydroxylated metalloids and certain saccharides (Figs. 2, 3, S1 and S2 and Table S1). Conversely, in PIPs, TIPs, and SIPs, these signatures contained bulky residues. These included aromatic phenylalanine, histidine, tryptophan, tyrosine and other bulky side-chains in the first positions, more specifically: F for the PIPs; His, N, Q, M or I for the TIPs; W, F, Y, V or T for the SIPs, and also W for the NIP1, and V, A or T for the NIP2 clade entries, compared to the NIP3 subclade, where the glycine residue was found (Fig. 5). These selectivity filter residue characteristics agreed with the published data (2, 33, 59–62), where HvNIP2;1 consistently clustered in the NIP3 sub-clade.

Discussion

In the present work, we observed that HvNIP2;1 when reconstituted in liposomes, permeated large cyclic saccharide molecules such as sucrose, *L*-arabinofuranose, and lactose but not *D*-glucose or *D*-fructose (Figs. 2, S1 and S2 and Table S1). This is a novel observation not reported previously for any AQP, including other structurally similar NIPs or NOD26. AQP9 from humans and a rat were found to transport large molecules including polyols (glycerol, *D*-mannitol, *D*-sorbitol), purines (adenine), pyrimidines (uracil, 5-fluorouracil), and thiourea, but not cyclic saccharides (63, 64). To define the saccharide specificity permeation by HvNIP2;1, an extensive panel of substrates was examined. Here, neutral (*D*-xylopyranose, *D*-glucopyranose, *D*-fructopyranose, *D*-galactopyranose, *D*-mannopyranose and *D*-mannopyranopheptaose) and charged (*D*-glucosamine, *N*-acetyl β -*D*-glucosamine, and *D*-glucuronic acid) monosaccharides, disaccharides (trehalose, cellobiose, gentiobiose), and the trisaccharide raffinose were permeated at lower rates than sucrose, *L*-arabinofuranose and lactose (Figs. 2, S1 and S2 and Table S1).

The next group of permeants transported by HvNIP2;1 included KCl and MgSO₄ ion pairs, while others such as CH₃COONa and NaNO₃ were permeated at lower rates

V, A, C, T, H) residues. Consensus amino acid residues (aa) and secondary structure elements (ss) are shown in three diversified aquaporin groups (*magenta*). Some of the key structural elements and residues for HvNIP2;1, such as NPA motifs, conserved R222 and Froger's P2 (T223), indicated in panel B and Table S2, are marked by *black bars*. B, Cartoon representation of monomeric HvNIP2;1 (*cyan cartoon*) in two orthogonal orientations (rotated by approximately 90°) features a predicted pore (cpk spheres; pore radii equal to sphere diameters) that contains the docked sucrose (cpk sticks). Separations (*dashed lines*) of sucrose with surrounding residues are between 2.4 Å and 3.2 Å. Selectivity filter residues G88, S207, G216, and R222 (*orange sticks with dots*) are near the pore entrance, while the NPA1 and NPA2 motifs (*green sticks*) are located in the pore. Residues in the Froger's P1–P5 positions (L148, T223, A227, Y239, and F240) are shown in *grey sticks*. Numbering of α -helices 1 to 6 and re-entrant α -helices (RH1 and RH2) are indicated. The crystal structure of SoPIP2;1 (*yellow cartoon*) with the predicted pore, NPA motifs and the docked sucrose molecule is shown with interacting residues (A152, N153, S153, T217, R255) at the distances between 2.2 Å and 3.5 Å (*dashed lines*). C, the bipartite symmetry distribution of peptide motifs in two orthogonal orientations (rotated by approximately 90°). Motif pairs and their positions in secondary structural elements are coloured identically in each half of the structure. Selectivity filter residues G88, S207, G216, and R222 (*orange sticks with dots*), residues in NPA1 and NPA2 motifs (*green sticks*) and in the Froger's P1–P5 positions are indicated. Six N- and 20 C-terminal residues are omitted for clarity in panels A and B.

permeability was completely suppressed. Notably, BA and sucrose did not compete during transport when applied together suggesting that they could be co-permeated (Fig. 3).

Sucrose permeation by the 3D AlphaFold model of *HvNIP2;1* was confirmed by docking, where *HvNIP2;1* contributed with M107, H160, G216, S217, and N219 as pivotal residues (Fig. 4), while no permeation was observed in *SoPIP2;1* AQP, also investigated through docking (Fig. 4B; right top panel). To explain the inability of *HvNIP2;1* to permeate smaller monomeric D-glucose or D-fructose molecules compared to the disaccharide sucrose (a larger molecule) that we observed, we suggest that the voluminous *HvNIP2;1* pore is decorated with hydrophobic and hydrophilic residues. We assume that in the *HvNIP2;1* pore, hydrophilic glucose could get trapped in pockets, while sucrose, due to its larger size avoids this trapping and slides through the pore. Conversely, a narrower and less capacious *SoPIP2;1* would collectively offer an unfavorable milieu for saccharide molecules to pass through.

In planta, *HvNIP2;1* AQP is expressed in roots and localized to epidermal and cortical cells of seminal roots and hypodermal cells in lateral roots (9, 10). The biological significance of our key findings of sucrose permeation means that *HvNIP2;1* could carry these permeants along with water and metalloids and that this could have profound importance in plant physiology (35, 76). Through this permeation, *HvNIP2;1* located on root cell membranes could recover sucrose directed from apoplastic (extracellular) to cellular environments. Consequently, this permeation route could constitute the novel element of a plant's cellular saccharide-transporting machinery.

To investigate the evolutionary origin of NIP AQPs, the phylogenetic reconstruction of 164 Viridiplantae MIPs was conducted to reveal that PIP, TIP, and SIP clades diversified before the chlorophyte (green algae) and embryophyte (land plants) split. This finding proposes that a single duplication event resulted in SIPs, or that multiple duplication events occurred before the chlorophyte and embryophyte split. These analyses used a relatively complex substitution model avoiding insufficient data for maximum likelihood but still kept accuracy during reconstruction, although there was little support for establishing deep relationships in MIPs as these sequences are relatively short. Hence, we restricted our maximum likelihood analyses to Viridiplantae only (Fig. 5), and separately to Viridiplantae, archaean, bacterial, fungal, and metazoan entries (Fig. S3 and Dataset S1). Notably, we observed increased gene duplication events in angiosperms (flowering plants) in all analyzed clades, our phylogenetic reconstruction showed that there were ancestral gene losses in chlorophytes (green algae), and where deep relationships between the NIP, PIP, and TIP entries were obscured by poor node support (Fig. 5 and Fig. S3).

Notably, and relevantly to the experimental permeation data obtained in this work, the NIP3 sub-clade segregated clearly and carried the GSSR selectivity filter signature. This NIP3 clade was lost or reduced in eudicots, although the deep phylogenetic relationships of the AQP clades were less

supported (Fig. S3). Nevertheless, the NIP3 clade segregation suggested that the pore residues were prone to residue alterations (variations) to modulate solute permeation selectivity in a NIP3 sub-clade, which would allow the permeation of hydroxylated metalloids and larger molecules such as sucrose (Fig. 2 and Fig. S1). Meanwhile, the PIP entries positioned on the short molecular branches, showed strong selectivity filter residue conservation in Viridiplantae, while the examined NIP, TIP, and SIP entries underwent more substitutions over time. In agreement with the published data (33, 60, 61), our analysis confirmed that the three NIP1-NIP3 sub-clades diversified early during embryophyte evolution, before the split from the tracheophytes. Further, the NIP3 sub-clade separated from those of NIP1 and NIP2 AQPs or appeared to be lost or reduced in eudicots (Fig. S3). Here, the NIP3 entries showed a clear dichotomy in the signatures of GSSG and FAAR selectivity filter residues (Fig. 5), which ultimately dictates their solute selectivity. This event may have resulted in gaining selectivity that would allow the NIP3 members to permeate specific saccharides, as shown in this work.

In conclusion, our phylogenomic analyses of sequences comprising 3157 AQPs (Figs. 5 and S3) propose that the *HvNIP2;1* acquired a unique solute specificity permeation as a saccharide transporter (Figs. 2, 3, S1 and S2 and Table S1) and that its structural model (Fig. 4 and Table S2) supports this concept, which is the leading notion of this work.

Experimental procedures

Materials and procedures used for fractionation and enzymatic digestion and disruption of *Pichia* cells, urea/alkali treatment of microsomal membrane fractions, solubilization from the urea/alkali-treated microsomal membrane fractions of *HvNIP2;1* and other analytical techniques are detailed in Supplementary information.

Cloning of *HvNIP2;1* and *Pichia pastoris* clone selection

Cloning of *HvNIP2;1* native cDNA (UniProtKB accession D8V828) in the pPICZ (frame B) (Invitrogen) expression vector yielding the native *HvNIP2;1*-Myc-6xHis-pPICZ-B DNA fusion, destined to be transformed in *P. pastoris*, and *Pichia* clone selection were conducted as described (46–48).

Barley *HvNIP2;1* expression in *Pichia pastoris* cells

Competent X-33 *P. pastoris* cells (Invitrogen) transformed with the linearized *HvNIP2;1*-Myc-6xHis-pPICZ-B were streaked on the YPD plates (composition defined in the EasySelect *Pichia* Expression Kit Manual) containing 100 µg/ml zeocin (InvivoGen) and incubated for 2 days at 28 °C (3). Cells from a single colony were inoculated into 2 mL of liquid BMGY media (composition defined in the EasySelect *Pichia* Expression Kit Manual) in 10 ml conical test tubes. Liquid cultures were grown for 2 days at 25 °C, transferred to 200 ml of liquid BMMY media (composition defined in the EasySelect *Pichia* Expression Kit Manual), and induced with 1% (v/v) methanol in 2-L Erlenmeyer flasks for 4 days at 25 °C maintaining 1% (v/v) methanol under shaking (120 rpm; Multitron

Broad-selective barley HvNIP2;1 aquaporin permeates sucrose

INFORS HT). Cells were harvested by centrifugation (4500g, 10 min, ambient temperature), pellets resuspended in 10% (v/v) glycerol, and stored at -80°C .

Purification of HvNIP2;1 via Immobilized Metal Affinity Chromatography (IMAC)

The SMA-solubilized preparation (49) was incubated with 0.5 to 1 ml of the Complete His-Tag Purification Resin (Roche, Indianapolis, IN, USA) equilibrated in SB and incubated for 16 to 18 h at ambient temperature. Resin with bound protein was packed in a disposable Bio-Rad column, and bound protein was eluted with 300 mM imidazole in SB at 1 ml/min flow rate at 4°C (11). Fractions (1 ml) were analysed by SDS-PAGE combined with IB, as described above. Positive fractions were pooled and concentrated to 200 μl on a Microcon Ultracel YM10 micro-concentrator (50 kDa exclusion limit, Millipore Billerica). The final preparation was aliquoted and stored with 20% (v/v) glycerol at -80°C .

Reconstitution of HvNIP2;1 in liposomes

DMPC lipids were dissolved at 10 mg/ml in chloroform, dried on a rotary evaporator under vacuum for 30 min, and rehydrated in the liposome buffer (LB) [20 mM Tris-HCl (pH 8.0) containing 100 mM KCl]. The lipid mixture was sonicated until clear and filtered through 100 nm pores (Avanti) of a uniform size using the LiposoFast Hand Extruder (Avestin). The SMA-solubilized HvNIP2;1 protein preparation and filtered DMPC liposomes were mixed at a ratio of 1:50 on a weight basis, mixed by gentle shaking at room temperatures for 15 min and the mixture was dialyzed in LB containing 50 mM MgCl_2 to disrupt the SMA polymer. The 50 mM MgCl_2 concentration was maintained in all the buffers after this step. No protein precipitation was observed during the reconstitution of HvNIP2;1 in DMPC liposomes.

Isolation of homogenous DMPC liposomes with reconstituted HvNIP2;1 through floatation on the iohexol gradient

Equal volumes (50 μl) of liposomes with reconstituted HvNIP2;1 and 80% (w/v) iohexol in the 25 mM HEPES-NaOH buffer (pH 7.5) containing 100 mM NaCl and 10% (w/v) glycerol were mixed (50, 51). The mixture was transferred to an ultra-clear polyallomer test tube (Beckman Coulter), overlaid with 350 μl of 30% (w/v) iohexol in the 25 mM HEPES-NaOH buffer (pH 7.5) containing 100 mM NaCl and 10% (w/v) glycerol, and with 100 μl of the 25 mM HEPES-NaOH buffer (pH 7.5) containing 100 mM NaCl. The mixture was centrifuged (100,000g, 4 h, four $^{\circ}\text{C}$) in the L-80XP ultracentrifuge using the SW55Ti swinging-bucket rotor (Beckman Coulter). After ultra-centrifugation, 60 μl fractions were sequentially collected from the top of the gradient to the bottom and examined by SDS-PAGE and IB. Selected fractions of liposomes with embedded HvNIP2;1 were pooled, centrifuged (10,000g, 2 min, 4°C), resuspended in LB, dialyzed for 18 h in LB at 4°C using 10-kD cut-off Slide-Alizarine dialysis cups (Thermo Fischer Scientific) and used in stopped-flow

light scattering recordings. Fractionated proteo-liposomes stored on crushed ice remained stable for up to 5 days.

Stopped-flow light scattering recordings of solute transport in proteo-liposomes with HvNIP2;1

Permeability of DMPC proteo-liposomes with reconstituted HvNIP2;1 and control liposomes lacking HvNIP2;1 was measured (46) to test the transport of 11 solutes with the DX.17 MV stopped-flow spectrophotometer (Applied Photophysics). The shrinking and re-swelling of vesicles were measured by 90° light scattering (550 nm) at 21°C , upon rapidly mixing solutions that create an outward-directed concentration gradient with test solutions in the liposome buffer [20 mM Tris-HCl (pH 8.0) containing 100 mM KCl] at the concentration of 0.2 M (340 mOsmol/kg). These measurements were repeated and extended, such that in total, the transport of a panel of 27 solutes was investigated. Traces from five individual stopped-flow acquisitions were averaged and normalized shrinking and swelling kinetics were fitted to a non-linear regression single exponential function, from which rate constants were calculated. To inhibit solute permeation, the thiol-group modifier AgNO_3 at 0.5 mM was used (52). Stopped-flow light scattering data were analyzed using Prism 9.0.0 (GraphPad Prism Software, Inc) based on two biological and two technical replicates of five averaged stopped-flow acquisitions, using non-linear regression of a one-phase decay (solute) or two-phase association (water) models. Osmotic permeability P coefficients for water were calculated based on $P_{\text{water}} = (V/A) \times \text{rate constant} / (V_w \times C_o)$ (77) and permeability coefficients of all other solutes based on $P_{\text{solute}} = (V/A) \times \text{rate constant}$ (78), where V/A is the volume to surface area ratio of liposomes, V_w is the partial molar volume of water and C_o is the external osmolarity after mixing, using the 50-nm liposome radius. P_{solute} coefficients were corrected for nonselective diffusion through lipid bilayers using control liposomes. Diameters of DMPC liposomes with reconstituted HvNIP2;1 and control liposomes were determined using the NICOMP 380 Particle Sizing System operating in a vesicle mode. The data were weighted on ± 60 liposomes.

Heterologous expression of HvNIP2;1 in X. laevis oocytes and oocyte swelling

HvNIP2;1 expression in oocytes was performed as described (46, 47). Briefly, native HvNIP2;1 DNA was inserted in the Gateway-enabled pGEMHE vector (9, 79) and complementary RNA (cRNA) was transcribed using the Ambion mMESSAGE mMACHINE kit (Life Technologies). 23 ng cRNA in 46 nl of RNA-free water or an equal volume of RNA-free water were injected in oocytes, followed by incubation in ND-96 for 24 to 48 h before measurements (9, 79). Permeability of HvNIP2;1 expressed in oocytes to solutes was investigated after the transfer of oocytes to the 5-fold diluted solution of ND96 (5 mM HEPES-NaOH buffer, pH 7.4 containing 96 mM NaCl, 2 mM KCl, 1.8 mM CaCl_2 , and 1 mM MgCl_2) supplemented with solutes at 160 mM concentration, which equaled to 200

mOsmol/kg osmolarity. 0.5 mM H[AuCl₄] (53) was used to inhibit the permeation of HvNIP2;1 expressed in oocytes.

3D molecular model of HvNIP2;1, crystal structure of SoPIP2;1, and sucrose docking and evaluations of permeation tunnels

The coordinates of a monomeric 3D HvNIP2;1 model from barley (*H. vulgare* L.) were taken from the AlphaFold Protein Structure Database (54; accession AF-D8V828-F1). In this AlphaFold HvNIP2;1 model, six residues were positioned in disallowed regions, which were all positioned at the N-terminal loop (A2, S3, N8, R10, S15, and E16) and which was not used in docking or structural evaluations of permeation pores.

The coordinates of spinach (*Spinacia oleracea* L.) SoPIP2;1 (PDB accession 2B5F) in an open state conformation (38, 39) and those of Lsi1 (PDB accession 7cjs) (42) OsNIP2;1 (PDB accession 7nl4) (43) and were taken from the Protein Data Bank.

The root-mean-square deviation values of 0.67 Å and 0.45 Å, between the HvNIP2;1 AlphaFold model and OsNIP2;1 or Lsi1, indicated high similarity at structural levels of all three AQPs, which was underlined by high sequence identities amongst three AQPs (defined in the Results section).

Docking of sucrose into HvNIP2;1 and SoPIP2;1 was performed using the HDock server (58), which is based on a hybrid algorithm of template-based modeling and *ab initio*-free docking of protein-protein and protein-ligand/DNA/RNA/systems through generating up to 100 docking poses. With each AQP, all 100 models were evaluated, and the top five models were selected (based on the position of sucrose within permeating pores) and used for analyses of sucrose dispositions.

Permeation tunnels in HvNIP2;1 and SoPIP2;1 and their properties such as lengths and diameters, were evaluated by Caver v.3.0.3 (56) (with maximum distance and desired radius parameters of 3 Å and 5 Å, respectively), embedded in the PyMol v.2.5.4 software (Schrödinger LLC), which was also used for the generation of graphics images of HvNIP2;1 and SoPIP2;1.

Phylogenomic analyses

Viridiplantae MIP sequences with matches to the PF00230 PFAM (80) were retrieved from Phytozome 12.1 (<https://phytozome.jgi.doe.gov/pz/portal.html>) (81) (Dataset S1). Archaeal, bacterial, fungal, and metazoan sequences were retrieved from the UniProt reference genomes using top hits from the EBI hmmer implementation (<https://www.ebi.ac.uk/Tools/hmmer/search/hmmer>) (82, 83). Barley and wheat sequences were curated from in-house collections and the NCBI GenBank (84). Excessively long, short, or fragmented sequences were manually removed. Two datasets were prepared, one with only Viridiplantae and a larger expanded collection of plant and archaeal, bacterial, fungal, and metazoan sequences. Jalview (85) was used to identify a 90% redundancy threshold for archaeal, bacterial, fungal, and metazoan sequences and to manually select cluster representatives. The

AlignSeqs function from DECIPHER (81) was used to align 164 selected Viridiplantae sequences. Clipkit (86) was used to trim excessively gapped sites under the gap model ($g = 0.95$). The expanded all domain dataset was prepared using hmalign where residues were assigned to the MIP PF00230 HMM from the Pfam database. The flanking unassigned residues were excluded using the hmalign trim function.

As MIP sequences were relatively short, data could not support the fit for complex, parameter-rich models to deep alignments. Thus, maximum-likelihood analyses were restricted to Viridiplantae, while the analyses of Viridiplantae, archaeal, bacterial, fungal, and metazoan sequences were limited to distance methods. Substitution model selection for the Viridiplantae and all domain data was performed using ModelTest-NG (87) with LG+G4 determined as best fit for both data under the AICc. The phylogeny was calculated using RAxML-NG v1.0.2 (88). The best-known maximum-likelihood tree was selected based on final GAMMA scores after 150 random and 150 parsimony start-tree searches. Confidence values were determined by calculating 1000 transfer bootstrap estimate replicates (89). Distance analysis was performed with FastME 2.0 (90) using the LG+G4 substitution model. Confidence values were determined by calculating 1000 bootstrap replicates.

Data availability

All data are included within the manuscript and its Supporting information.

Supporting information—This article contains Supporting information (10, 46–49, 51, 92–95).

Acknowledgments—We acknowledge Dr. Sunita Ramesh (University of Adelaide) for technical assistance.

Author contributions—M. H., S. D. T., and J. G. S.: methodology; A. V., J. G. S., S. D. T., and M. H.: investigation; M. H., S. D. T., A. V., and J. G. S.: formal analysis. M. H., J. G. S., and S. D. T.: writing—original draft.

Funding and additional information—This work was supported by funding from the Australian Research Council (DP120100900) (Australia) to M. H., and the Australian Research Council Centre of Excellence in Plant Energy Biology (CE1400008) (Australia) to S. D. T. A. V. was supported by the Sastra University scholarship Thanjavur, (Tamilnadu, India), and M. H. and J. G. S. from the University of Adelaide (Australia).

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: AQP(s), aquaporin(s); aqPn, AQP NIP-like; BA, boric acid; C_o, external osmolarity after mixing; DMPC, 1, 2-dimyristoyl-sn-glycero-3-phosphocholine; EBI, European Bioinformatics Institute; FAAR, phenylalanine-alanine-alanine-arginine residues; FHTR, phenylalanine-histidine-threonine-arginine residues; GIP, GIpF-like intrinsic proteins; GSGR, glycine-serine-glycine-arginine; HEK, Human Embryonic Kidney; HEPES, (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid); HIP,

Broad-selective barley HvNIP2;1 aquaporin permeates sucrose

hybrid intrinsic proteins; HvNIP2;1, *Hordeum vulgare* NIP2;1; IMAC, Immobilized Metal Affinity Chromatography; MEME, Multiple EM (Expectation Maximization) for Motif Elicitation; NCBI, National Center for Biotechnology Information; NPA, Asn-Pro-Ala residues; NIP, NOD26-like intrinsic proteins; NOD26, nodulin 26; PIP, plasma membrane intrinsic proteins; PDB, Protein Data Bank; P, permeability coefficient; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SIP, small and basic intrinsic proteins; SMA, styrene-maleic anhydride; SoPIP2;1, *Spinacia oleracea* PIP2;1; TIP, tonoplast intrinsic proteins; UniProt, Universal Protein Resource; V_w , partial molar volume of water; XIP, X-intrinsic proteins.

References

1. Chaumont, F., and Tyerman, S. D. (2014) Aquaporins: highly regulated channels controlling plant water relations. *Plant Physiol.* **164**, 1600–1618
2. Luang, S., and Hrmova, M. (2017) Structural basis of the permeation function of plant aquaporins. In: Chaumont, F., Tyerman, S. D., eds. *Plant Aquaporins: From Transport to Signaling*, Springer International Publishing, Cham: 1–28. Signaling and Communication in Plants
3. Tyerman, S. D., McGaughey, S. A., Qiu, J., Yool, A. J., and Byrt, C. S. (2021) Adaptable and multifunctional ion-conducting aquaporins. *Annu. Rev. Plant Biol.* **72**, 703–736
4. Danielson, J.Å., and Johanson, U. (2008) Unexpected complexity of the aquaporin gene family in the moss *Physcomitrella patens*. *BMC Plant Biol.* **8**, 45
5. Gustavsson, S., Lebrun, A.-S., Nordén, K., Chaumont, F., and Johanson, U. (2005) A novel plant major intrinsic protein in *Physcomitrella patens* most similar to bacterial glycerol channels. *Plant Physiol.* **139**, 287–295
6. Ahmed, J., Mercx, S., Boutry, M., and Chaumont, F. (2020) Evolutionary and predictive functional insights into the aquaporin gene family in the allotetraploid plant *Nicotiana tabacum*. *Int. J. Mol. Sci.* **21**, 4743
7. Ding, L., and Chaumont, F. (2020) Are aquaporins expressed in stomatal complexes promising targets to enhance stomatal dynamics? *Front. Plant Sci.* **11**, 458
8. Mukhopadhyay, R., Bhattacharjee, H., and Rosen, B. P. (2014) Aquaglyceroporins: generalized metalloid channels. *Biochim. Biophys. Acta* **1840**, 1583–1591
9. Schnurbusch, T., Hayes, J., Hrmova, M., Baumann, U., Ramesh, S. A., Tyerman, S. D., et al. (2010) Boron toxicity tolerance in barley through reduced expression of the multifunctional aquaporin HvNIP2;1. *Plant Phys.* **153**, 1706–1715
10. Chiba, Y., Mitani, N., Yamaji, N., and Ma, J. F. (2009) HvLsi1 is a silicon influx transporter in barley. *Plant J.* **57**, 810–818
11. Montpetit, J., Vivancos, J., Mitani-Ueno, N., Yamaji, N., Rémus-Borel, W., Belzile, F., et al. (2012) Cloning, functional characterization and heterologous expression of TaLsi1, a wheat silicon transporter gene. *Plant Mol. Biol.* **79**, 35–46
12. Fu, D., Libson, A., Miercke, L. J., Weitzman, C., Nollert, P., Krucinski, J., et al. (2000) Structure of a glycerol-conducting channel and the basis for its selectivity. *Science* **290**, 481–486
13. Ma, J. F., Tamai, K., Yamaji, N., Mitani, N., Konishi, S., Katsuhara, M., et al. (2006) A silicon transporter in rice. *Nature* **440**, 688–691
14. Bienert, G. P., Bienert, M. D., Jahn, T. P., Boutry, M., and Chaumont, F. (2011) Solanaceae XIPs are plasma membrane aquaporins that facilitate the transport of many uncharged substrates: characterization of SolanaceaeXIPs. *Plant J.* **66**, 306–317
15. Bienert, M. D., Muries, B., Crappe, D., Chaumont, F., and Bienert, G. P. (2019) Overexpression of X intrinsic protein 1;1 in *Nicotiana tabacum* and *Arabidopsis* reduces boron allocation to shoot sink tissues. *Plant Direct.* **3**, e00143
16. Liu, L.-H., Ludewig, U., Gassert, B., Frommer, W. B., and von Wirén, N. (2003) Urea transport by nitrogen-regulated tonoplast intrinsic proteins in *Arabidopsis*. *Plant Physiol.* **133**, 1220–1228
17. Takano, J., Wada, M., Ludewig, U., Schaaf, G., von Wirén, N., and Fujiwara, T. (2006) The Arabidopsis major intrinsic protein NIP5;1 is essential for efficient boron uptake and plant development under boron limitation. *Plant Cell* **18**, 1498–1509
18. Choi, W.-G., and Roberts, D. M. (2007) Arabidopsis NIP2;1, a major intrinsic protein transporter of lactic acid induced by anoxic stress. *J. Biol. Chem.* **282**, 24209–24218
19. Hayes, J., Pallotta, M., Baumann, U., Berger, B., Langridge, P., and Sutton, T. (2013) Germanium as a tool to dissect boron toxicity effects in barley and wheat. *Funct. Plant Biol.* **40**, 618–627
20. Wang, Y., Li, R., Li, D., Jia, X., Zhou, D., Li, J., et al. (2017) NIP1;2 is a plasma membrane-localized transporter mediating aluminum uptake, translocation, and tolerance in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 5047–5052
21. Uehlein, N., Lovisolo, C., Siefritz, F., and Kaldenhoff, R. (2003) The tobacco aquaporin NtAQP1 is a membrane CO₂ pore with physiological functions. *Nature* **425**, 734–737
22. Jahn, T. P., Möller, A. L. B., Zeuthen, T., Holm, L. M., Klaerke, D. A., Mohsin, B., et al. (2004) Aquaporin homologues in plants and mammals transport ammonia. *FEBS Lett.* **574**, 31–36
23. Loqué, D., Ludewig, U., Yuan, L., and von Wirén, N. (2005) Tonoplast intrinsic proteins AtTIP2;1 and AtTIP2;3 facilitate NH₃ transport into the vacuole. *Plant Physiol.* **137**, 671–680
24. Hwang, J. H., Ellingson, S. R., and Roberts, D. M. (2010) Ammonia permeability of the soybean nodulin 26 channel. *FEBS Lett.* **584**, 4339–4343
25. Zwiazek, J. J., Xu, H., Tan, X., Navarro-Ródenas, A., and Morte, A. (2017) Significance of oxygen transport through aquaporins. *Sci. Rep.* **7**, 40411
26. Weaver, C. D., Shomer, N. H., Louis, C. F., and Roberts, D. M. (1994) Nodulin 26, a nodule-specific symbiosome membrane protein from soybean, is an ion channel. *J. Biol. Chem.* **269**, 17858–17862
27. Lee, J. W., Zhang, Y., Weaver, C. D., Shomer, N. H., Louis, C. F., and Roberts, D. M. (1995) Phosphorylation of nodulin 26 on serine 262 affects its voltage-sensitive channel activity in planar lipid bilayers. *J. Biol. Chem.* **270**, 27051–27057
28. Diehn, T. A., Bienert, M. D., Pommerrenig, B., Liu, Z., Spitzer, C., Bernhardt, N., et al. (2019) Boron demanding tissues of *Brassica napus* express specific sets of functional Nodulin26-like intrinsic proteins and BOR1 transporters. *Plant J.* **100**, 68–82
29. Pommerrenig, B., Diehn, T. A., and Bienert, G. P. (2015) Metalloido-porins: essentiality of Nodulin 26-like intrinsic proteins in metalloid transport. *Plant Sci.* **238**, 212–227
30. Noronha, H., Araújo, D., Conde, C., Martins, A. P., Soveral, G., Chaumont, F., et al. (2016) The grapevine uncharacterized intrinsic protein 1 (VvXIP1) is regulated by drought stress and transports glycerol, hydrogen peroxide, heavy metals but not water. *PLoS One* **11**, e0160976
31. Abascal, F., Irisarri, I., and Zardoya, R. (2014) Diversity and evolution of membrane intrinsic proteins. *Biochim. Biophys. Acta* **1840**, 1468–1481
32. [preprint] Borstlap, A. C. (2018) Plant aquaporins: the origin of NIPs. *bioRxiv*. <https://doi.org/10.1101/351064v1>
33. Pommerrenig, B., Diehn, T. A., Bernhardt, N., Bienert, M. D., Mitani-Ueno, N., Fuge, J., et al. (2020) Functional evolution of nodulin 26-like intrinsic proteins: from bacterial arsenic detoxification to plant nutrient transport. *New Phytol.* **225**, 1383–1396
34. Wallace, I. S., and Roberts, D. M. (2004) Homology modeling of representative subfamilies of Arabidopsis major intrinsic proteins. Classification based on the aromatic/arginine selectivity filter. *Plant Physiol.* **135**, 1059–1068
35. Hrmova, M., Gilliam, M., and Tyerman, S. D. (2020) Plant transporters involved in combating boron toxicity: beyond 3D structures. *Biochem. Soc. Trans.* **48**, 1683–1696
36. Yu, J., Yool, A. J., Schulten, K., and Tajkhorshid, E. (2006) Mechanism of gating and ion conductivity of a possible tetrameric pore in aquaporin-1. *Structure* **14**, 1411–1423
37. De Rosa, A., Watson-Lazowski, A., Evans, J. R., and Groszmann, M. (2020) Genome-wide identification and characterisation of Aquaporins in *Nicotiana tabacum* and their relationships with other Solanaceae species. *BMC Plant Biol.* **20**, 266

38. Törnroth-Horsefield, S., Wang, Y., Hedfalk, K., Johanson, U., Karlsson, M., Tajkhorshid, E., *et al.* (2006) Structural mechanism of plant aquaporin gating. *Nature* **439**, 688–694
39. Nyblom, M., Frick, A., Wang, Y., Ekvall, M., Hallgren, K., Hedfalk, K., *et al.* (2009) Structural and functional analysis of SoPIP2;1 mutants adds insight into plant aquaporin gating. *J. Mol. Biol.* **387**, 653–668
40. Wang, H., Schoebel, S., Schmitz, F., Dong, H., and Hedfalk, K. (2020) Characterization of aquaporin-driven hydrogen peroxide transport. *Biochim. Biophys. Acta Biomembr.* **1862**, 183065
41. Kirscht, A., Kaptan, S. S., Bienert, G. P., Chaumont, F., Nissen, P., de Groot, B. L., *et al.* (2016) Crystal structure of an ammonia-permeable aquaporin. *PLoS Biol.* **14**, e1002411
42. Saitoh, Y., Mitani-Ueno, N., Saito, K., Matsuki, K., Huang, S., Yang, L., *et al.* (2021) Structural basis for high selectivity of a rice silicon channel Lsi1. *Nat. Commun.* **12**, 6236
43. van den Berg, B., Pedebos, C., Bolla, J. R., Robinson, C. V., Baslé, A., and Khalid, S. (2021) Structural basis for silicic acid uptake by higher plants. *J. Mol. Biol.* **433**, 167226
44. Schnurbusch, T., Hayes, J., and Sutton, T. (2010) Boron toxicity tolerance in wheat and barley: Australian perspectives. *Breed. Sci.* **60**, 297–304
45. Deng, F., Liu, X., Chen, Y., Rathinasabapathi, B., Rensing, C., Chen, J., *et al.* (2020) Aquaporins mediated arsenite transport in plants: molecular mechanisms and applications in crop improvement. *Crit. Rev. Environ. Sci. Technol.* **50**, 1613–1639
46. Rongala, J. (2016). Ph.D. thesis, *Characterisation of Cereal Transport Proteins Involved in Boron Toxicity Tolerance*. The University of Adelaide
47. Venkataraghavan, A. (2017). MSc thesis, *Transport Characteristics of a Plant Aquaporin from the Nodulin-26 like Intrinsic Protein (NIP) Family*. SASTRA (Shanmugha Arts Science Technology & Research Academy) University and The University of Adelaide
48. Nagarajan, Y. (2010). MSc thesis, *A Silicon Transporter Involved in Abiotic Stress Tolerance in Cereal Plants: Molecular Cloning and Recombinant Protein Expression*. Flinders University and the University of Adelaide
49. Knowles, T. J., Finka, R., Smith, C., Lin, Y.-P., Dafforn, T., and Overduin, M. (2009) Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. *J. Am. Chem. Soc.* **131**, 7484–7485
50. Periasamy, A., Shadiac, N., Amalraj, A., Garajová, S., Nagarajan, Y., Waters, S., *et al.* (2013) Cell-free protein synthesis of membrane (1,3)- β -D-glucan (curdlan) synthase: co-translational insertion in liposomes and reconstitution in nanodiscs. *Biochim. Biophys. Acta* **1828**, 743–757
51. Nagarajan, Y., Rongala, J., Luang, S., Singh, A., Shadiac, N., Hayes, J., *et al.* (2016) A barley efflux transporter operates in a Na⁺-dependent manner, as revealed by a multidisciplinary platform. *Plant Cell* **28**, 202–218
52. Dordas, C., Chrispeels, M. J., and Brown, P. H. (2000) Permeability and channel-mediated transport of boric acid across membrane vesicles isolated from squash roots. *Plant Physiol.* **124**, 1349–1362
53. Niemietz, C. M., and Tyerman, S. D. (2002) New potent inhibitors of aquaporins: silver and gold compounds inhibit aquaporins of plant and human origin. *FEBS Lett.* **531**, 443–447
54. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589
55. Needleman, S. B., and Wunsch, C. D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**, 443–453
56. Jurcik, A., Bednar, D., Byska, J., Marques, S. M., Furmanova, K., Daniel, L., *et al.* (2018) CAVER Analyst 2.0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. *Bioinformatics* **34**, 3586–3588
57. Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., *et al.* (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–208
58. Yan, Y., Tao, H., He, J., and Huang, S.-Y. (2020) The HDock server for integrated protein-protein docking. *Nat. Protoc.* **15**, 1829–1852
59. Agre, P., Sasaki, S., and Chrispeels, M. J. (1993) Aquaporins: a family of water channel proteins. *Am. J. Physiol.* **265**, F461
60. Deshmukh, R. K., Vivancos, J., Ramakrishnan, G., Guérin, V., Carpentier, G., Sonah, H., *et al.* (2015) A precise spacing between the NPA domains of aquaporins is essential for silicon permeability in plants. *Plant J.* **83**, 489–500
61. Deshmukh, R., Sonah, H., and Belanger, R. R. (2020) New evidence defining the evolutionary path of aquaporins regulating silicon uptake in land plants. *J. Exp. Bot.* **71**, 6775–6788
62. Nawaz, M. A., Azeem, F., Zakharenko, A. M., Lin, X., Atif, R. M., Baloch, F. S., *et al.* (2020) In-silico exploration of channel type and efflux silicon transporters and silicification proteins in 80 sequenced viridiplantae genomes. *Plants* **9**, 1612
63. Tsukaguchi, H., Shayakul, C., Berger, U. V., Mackenzie, B., Devidas, S., Guggino, W. B., *et al.* (1998) Molecular characterization of a broad selectivity neutral solute channel. *J. Biol. Chem.* **273**, 24737–24743
64. Tsukaguchi, H., Weremowicz, S., Morton, C. C., and Hediger, M. A. (1999) Functional and molecular characterization of the human neutral solute channel aquaporin-9. *Am. J. Physiol.* **277**, F685–F696
65. Yool, A. J., Stamer, W. D., and Regan, J. W. (1996) Forskolin stimulation of water and cation permeability in aquaporin 1 water channels. *Science* **273**, 1216–1218
66. Yasui, M., Hazama, A., Kwon, T. H., Nielsen, S., Guggino, W. B., and Agre, P. (1999) Rapid gating and anion permeability of an intracellular aquaporin. *Nature* **402**, 184–187
67. Byrt, C. S., Zhao, M., Kourghi, M., Bose, J., Henderson, S. W., Qiu, J., *et al.* (2017) Non-selective cation channel activity of aquaporin AtPIP2;1 regulated by Ca²⁺ and pH. *Plant Cell Environ.* **40**, 802–815
68. Kourghi, M., Nourmohammadi, S., Pei, J. V., Qiu, J., McGaughey, S., Tyerman, S. D., *et al.* (2017) Divalent cations regulate the ion conductance properties of diverse classes of aquaporins. *Int. J. Mol. Sci.* **18**, E2323
69. Qiu, J., McGaughey, S. A., Groszmann, M., Tyerman, S. D., and Byrt, C. S. (2020) Phosphorylation influences water and ion channel function of AtPIP2;1. *Plant Cell Environ.* **43**, 2428–2442
70. Raven, J. A. (1980) Short- and long-distance transport of boric acid in plants. *New Phytol.* **84**, 231–249
71. Borgnia, M. J., Kozono, D., Calamita, G., Maloney, P. C., and Agre, P. (1999) Functional reconstitution and characterization of AqpZ, the E. coli water channel protein. *J. Mol. Biol.* **291**, 1169–1179
72. Dean, R. M., Rivers, R. L., Zeidel, M. L., and Roberts, D. M. (1999) Purification and functional reconstitution of soybean nodulin 26. An aquaporin with water and glycerol transport properties. *Biochemistry* **38**, 347–353
73. Routray, P., Li, T., Yamasaki, A., Yoshinari, A., Takano, J., Choi, W. G., *et al.* (2018) Nodulin intrinsic protein 7;1 Is a tapetal boric acid channel involved in pollen cell wall formation. *Plant Physiol.* **178**, 1269–1283
74. Mitani, N., Yamaji, N., and Ma, J. F. (2008) Characterization of substrate specificity of a rice silicon transporter, Lsi1. *Pflug. Arch. Eur. J. Physiol.* **456**, 679–686
75. Rivers, R. L., Dean, R. M., Chandy, G., Hall, J. E., Roberts, D. M., and Zeidel, M. L. (1997) Functional analysis of nodulin 26, an aquaporin in soybean root nodule symbiosomes. *J. Biol. Chem.* **272**, 16256–16261
76. Hrmova, M., and Gilliam, M. (2018) Plants fighting back: to transport or not to transport, this is a structural question. *Curr. Opin. Plant Biol.* **46**, 68–76
77. Zhao, M., Tan, H., Scharwies, J., Levin, K., Evans, J. R., and Tyerman, S. D. (2017) Association between water and carbon dioxide transport in leaf plasma membranes: assessing the role of aquaporins. *Plant Cell Environ.* **40**, 789–801
78. Verkman, A. S., Dix, J. A., and Seifter, J. L. (1985) Water and urea transport in renal microvillus membrane vesicles. *Am. J. Physiol.* **248**, F650–F655
79. Wege, S., Qiu, J., Byrt, C., Houston, K., Waugh, R., Gilliam, M., *et al.* (2021) A single residue deletion in the barley HKT1;5 P189 variant restores plasma membrane localisation but not Na⁺ conductance. *Biochim. Biophys. Acta Biomembr.* **1863**, 183669
80. Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., *et al.* (2021) Pfam: the protein families database in 2021. *Nucleic Acids Res.* **49**, D412–D419
81. Wright, E. S. (2015) DECIPHER: harnessing local sequence context to improve protein multiple sequence alignment. *BMC Bioinformatics*. **16**, 322
82. Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., and Finn, R. D. (2018) HMMER web server: 2018 update. *Nucleic Acids Res.* **46**, W200–W204

Broad-selective barley HvNIP2;1 aquaporin permeates sucrose

83. Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., *et al.* (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641
84. Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., *et al.* (2013) GenBank. *Nucleic Acids Res.* **41**, D36–D42
85. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191
86. Steenwyk, J. L., Iii, T. J. B., Li, Y., Shen, X.-X., and Rokas, A. (2020) ClipKIT: a multiple sequence alignment trimming software for accurate phylogenomic inference. *PLoS Biol.* **18**, e3001007
87. Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., and Flouri, T. (2020) ModelTest-NG: a new and scalable tool for the selection of dna and protein evolutionary models. *Mol. Biol. Evol.* **37**, 291–294
88. Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019) RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**, 4453–4455
89. Lemoine, F., Domelevo Entfellner, J.-B., Wilkinson, E., Correia, D., Dávila Felipe, M., De Oliveira, T., *et al.* (2018) Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature* **556**, 452–456
90. Lefort, V., Desper, R., and Gascuel, O. (2015) FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol. Biol. Evol.* **32**, 2798–2800
91. Pei, J., Kim, B.-H., and Grishin, N. V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res.* **36**, 2295–2300
92. Kaneko, T., Kitamura, K., and Yamamoto, Y. (1973) Susceptibilities of yeasts to yeast cell wall lytic enzyme of *Arthrobaeter luteus*. *Agric. Biol. Chem.* **37**, 2295–2302
93. Hasler, L., Walz, T., Tittmann, P., Gross, H., Kistler, J., and Engel, A. (1998) Purified lens major intrinsic protein (MIP) forms highly ordered tetragonal two-dimensional arrays by reconstitution. *J. Mol. Biol.* **279**, 855–864
94. Karlsson, M., Fotiadis, D., Sjövall, S., Johansson, I., Hedfalk, K., Engel, A., *et al.* (2003) Reconstitution of water channel function of an aquaporin overexpressed and purified from *Pichia pastoris*. *FEBS Lett.* **537**, 68–72
95. Luang, S., Hrmova, M., and Ketudat Cairns, J. R. (2010) High-level expression of barley beta-d-glucan exohydrolase HvExoI from a codon-optimized cDNA in *Pichia pastoris*. *Protein Expr. Purif.* **73**, 90–98